CHAPTER 2

Varieties of higher-order theory

David M. Rosenthal

1. Introductory

A touchstone of much modern theorizing about the mind is the idea, still tacitly accepted by many, that a state's being mental implies that it's conscious. This view is epitomized in the dictum, put forth by theorists as otherwise divergent as Thomas Nagel (1979: 174) and Daniel Dennett (1991:132), that the appearance and reality of mental states coincide.

Traditionally this claim was cast not in terms of a state's being conscious, but in terms of one's being conscious of the state. Thus Descartes writes that "no thought can exist in us of which we are not conscious at the very moment it exists in us."1 These remarks echo Aristotle's claim that, since thoughts take on the forms of their objects, thoughts actually have themselves as objects, from which he concludes that, whenever we perceive or think, we perceive that we do.2

Today we typically cast things not in terms of what states we're conscious of, but what states are conscious, tout court. But doubtless this is meant to capture the same phenomenon.3 If an individual is in a mental state but is in no way whatever conscious of that state, we would not intuitively count it as a conscious state. So a state's being conscious consists of one's being conscious of it in some suitable way; perhaps, for example, we must be immediately conscious of it. It is this equivalence of a state's being conscious with one's being conscious of it in some suitable way that points toward a higher-order theory of what it is for a mental state to be conscious. We can explain a state's being conscious in terms of a higher-order state's being directed on that state because a state's being conscious consists in one's being conscious of it.

It is occasionally held that explaining a state's being conscious in terms of one's being conscious of that state is circular, since it explains consciousness in terms of consciousness (e.g., Goldman 1993:366). But this is a mistake. We
understand what it is to be conscious of something independently of what it is for a state to be conscious. We are conscious of things when we sense them, perceive them, or have thoughts about them as being present to us. Circularity would threaten only if these kinds of state could not themselves occur without being conscious states. Not only does it beg the question against such a theory to assume that; we have ample independent reason, both from everyday experience and experimental findings, to hold that such states do occur without being conscious. To keep things clear, we can refer to being conscious of something as transitive consciousness and to a mental state's being conscious as state consciousness (Rosenthal 1990/1997).

There are, however, a variety of ways in which one might seek to characterize the higher-order awareness of our conscious states. We are aware of things by sensing them, perceiving them, and having thoughts about them; which of these is operative when we are aware of our conscious states? And is the higher-order awareness distinct from the target state one is aware of, or might it be that a state's being present be a part of that state itself? Might a state's being conscious consist, moreover, simply in a disposition for a higher-order state to occur, rather than in the actual occurrence of that higher-order state? Finally, if being in a conscious state is a matter of one's being aware of oneself as being in the relevant state, might conscious states consist simply in that higher-order awareness, without any first-order target at all? It is these questions that I address in what follows.

2. Higher-order sensing

Sensing things is one way to be conscious, or aware, of those things. So one possibility, is that we sense our conscious states. This appeal to inner sense has considerable intuitive force. One reason such a model is inviting has to do with the intuitive immediacy that Descartes urged characterizes the way we're conscious of our conscious states. We seem to be immediately aware of the things we sense. Nothing seems subjectively to intervene between the things we sense. Nevertheless, something seems to intervene between our consciousness of the states we sense. This is the higher-order awareness distinct from the target state one is aware of, or might that higher-order awareness be part of that state itself? Might a state's being conscious consist, moreover, simply in a disposition for a higher-order state to occur, rather than in the actual occurrence of that higher-order state? Finally, if being in a conscious state is a matter of one's being aware of oneself as being in the relevant state, might conscious states consist simply in that higher-order awareness, without any first-order target at all? It is these questions that I address in what follows.

Mental capacity. And, even in creatures like ourselves, with higher mental functions, state consciousness seems intuitively closely tied to sensory functioning. It may even seem as though sensing is the only way we are transitively conscious of things. If so, inner sense is the only option for a higher-order theory.

Sensing always involves mental qualities, in virtue of which a creature senses various perceptible properties, and the qualities that occur when sensing is conscious present a formidable challenge for any theory of consciousness. When qualitative states are conscious, there is something qualitative that it's like for the creature to be in those states. The states seem subjectively to "light up." A theory of consciousness must give some account of what such lighting up amounts to as well as some credible story about how it arises.

Higher-order sensing may seem especially appealing in this connection. If the higher-order states in virtue of which qualitative states are conscious are themselves qualitative, perhaps that will help explain what it is for conscious qualitative states to light up and how they do. A qualitative state's lighting up would consist in one's being conscious of that state by way of a higher-order qualitative state.

But despite these apparent advantages, such a theory faces insuperable difficulties. It seemed that being aware of qualitative states by way of higher-order states that themselves exhibit mental qualities would explain the first-order states' lighting up. But it's unclear how such higher-order mental qualities might help, and even that there could be any such higher-order qualities. Higher-order qualities can't help explain the lighting up unless they are themselves conscious. But, if a state's being conscious consists in its being sensed, the higher-order qualities would be conscious only if there were, in turn, third-order sensations that sensed those second-order sensations. The threat of regress looms. Moreover, the only mental qualities we're ever conscious of are those of first-order conscious states. We're never conscious of distinct, higher-order qualities, even when we become aware of being conscious of the first-order states, as we do when we focus introspectively on those states.

An even greater difficulty arises about what those higher-order qualities might be. Perhaps they're the same in each case as the first-order qualities, so that conscious seeing, for example, occurs when we visually sense our first-order visual states. But that cannot be. Mental qualities are keyed to the properties they enable a creature to perceive. The range of qualities that characterize visual sensations, for example, resemble and differ in ways that make it possible to discriminate physical objects in respect of their perceptible color and spatial properties. These mental qualities plainly aren't the same as the perceptible properties to which they enable visual access. Whatever the property an
object has in virtue of which we call it red, that property cannot be the same as any property of the visual sensations by means of which we sense objects as being red.

So, if we do actually sense our visual sensations, it must be that the higher-order mental qualities in virtue of which we do so are distinct from the first-order mental qualities that enable visual access to colored physical objects. But it’s wholly unclear what qualities these might be. Nor is it clear how unknown qualities could help explain what it is for conscious qualitative states to light up. Higher-order sensing cannot make good on its promise to do justice to the intuitive immediacy and basic mental character of consciousness.

3. Higher-order perceiving

Sensing has no conceptual content. But perceiving not only has qualitative character, as does sensing, but has conceptual content as well; so it’s convenient to think of perceiving as conceptualized sensing. Just as its conceptual content distinguishes perceiving from mere sensing, so its qualitative character distinguishes it from mere thinking. No mental quality need occur if I merely think that there is a red object in front of me, whereas perceiving a red object involves some distinctive mental quality.

So the difficulty about higher-order qualities also undermines the hypothesis that we are aware of our conscious states by perceiving them, rather than just sensing them. Perception always exhibits mental qualities, and there are no suitable higher-order qualities for higher-order perception to exhibit. So there are no higher-order perceptions, properly so called.

Still, the absence of higher-order qualities may seem to be more decisive against higher-order sensing than against higher-order perceiving. Qualitative character is the only distinctively mental way we have to characterize sensing; so without higher-order qualities there is no higher-order sensing. Perceiving, because it exhibits both conceptual content and qualitative character, has mental properties apart from its mental quality. So it’s open to argue that, despite the absence of higher-order qualities, perceiving still provides the best model of the higher-order awareness we have of our conscious states. Thus William Lycan writes that a theory can appeal to higher-order perceiving without claiming that such higher-order awareness "is like external perception in every single respect" (1996:28; cf. Lycan, this volume: §5).

Since higher-order awareness of conscious states cannot resemble perceiving in respect of mental quality, it must be in respect of its other mental properties, which are intentional. So to sustain higher-order perceiving, we must show that the higher-order awareness is, in ways that matter, more like perceiving than like thinking. Lycan advances a number of respects in which he believes this is so. For one thing, he urges, we cannot have a thought about something without already being aware of it; so we would have to be perceptually aware of our mental states before having thoughts about them. But we have thoughts about many things we have never perceived. Since mental states are arguably just the kinds of things we cannot perceive, they are doubtless examples of this.

Lycan also argues that first-person appearances support a perceptual model. The phenomenology of being in conscious states, he urges, represents those states as present to us, and we seem able to attend to those states in the way we can attend to things we perceive. But we attend to objects of thought no less than those of perception. And thoughts can also represent things as present to us. Perceiving does always represent things as being to some degree present, whereas thoughts often do not; so higher-order awareness resembles only some cases of thinking, whereas it resembles all cases of perceiving. But that doesn’t show that it’s more like perceiving than like the relevant cases of thinking.

Lycan urges that we have considerable voluntary control over which perceptual states in our sensory fields are conscious, and that such voluntary control is more characteristic of perceptual awareness than of the awareness that comes from thinking about things. One might doubt that we have all that much voluntary control over our higher-order awareness. But, whatever the case about that, we can also direct and focus our thought processes, perhaps even more readily than we can our perceiving.

Lycan argues that our higher-order awareness monitors what states we are in, as perceiving does with various physical properties of things; it’s less obvious, by contrast, that we could monitor things by having thoughts about them. Perceptual monitoring relies on sense organs, or other suitable mechanisms in the case of enterceptive and proprioceptive perceiving, to discern how things are. If higher-order awareness were like perceiving, it too would require some such mechanism, which presumably would be a matter of cortical connections. And such connections could subserve such monitoring equally well if, independently of this issue, such higher-order awareness were more like thought than like perception.

A perceptual model, Lycan maintains, can capture the way our awareness of our mental states comes in degrees, which a model based on thoughts can’t. Lycan’s examples suggest that he actually has in mind degrees of attention; we attend more to some states than to others. But thoughts vary in how focused
and attentive they are no less than perceptions. Lycan claims that higher-order awareness is like perceiving in that we regard it as a reliable source of information. But that holds at least as well for the thoughts we have about small numbers, simple shapes, logical connections, and the everyday behavior of commonsense objects.

Any higher-order theory must explain how it is that we can become aware of exquisitely fine-grained differences among our qualitative states. This looks problematic for a higher-order theory based on thoughts, since we arguably don't have concepts corresponding to all the different mental qualities we can be conscious of. Lycan urges that higher-order awareness resembles perceiving in this respect, since perceiving is dedicated to discerning fine differences in perceptible properties.

But thinking distinguishes among properties in at least as fine-grained a way as does perceiving. And, though we plainly don't have distinct concepts for all our conscious mental qualities, that doesn't show that we lack the conceptual resources needed to capture all those qualitative variations, since we can readily capture them using comparative concepts. We describe shades of red, for example, as more or less yellowish, lighter or darker, more or less saturated, and the like. And we do this with mental qualities no less than with the perceptible properties those qualities enable us to perceive. Our conceptual resources are, after all, sufficient to make us aware of all the fine-grained variations among our conscious mental qualities.

There is compelling experimental evidence that we actually are aware of our mental qualities in respect of such comparative aspects. We're aware of far more fine-grained differences among mental qualities when they occur together than when they occur one at a time (see Raffman 1995). It might be argued that this tells against any higher-order theory. Why wouldn't the higher-order states in virtue of which we're conscious of mental qualities be independent of whether qualitative states occur together or not? And if they are, we would be conscious of qualitative states with the same fineness of grain whichever way they occur.

But there is no basis for assuming that we are conscious of our mental qualities in the same way whatever qualities accompany them. And this result provides evidence that we aren't, but rather are conscious of mental qualities at least partly in comparison with one another, as exemplifying particular mental qualities to a greater or less degree than some accompanying quality. And this in turn suggests that our higher-order awareness of mental qualities are thoughts, since thoughts are more versatile than perceptions in characterizing things comparatively.

Lycan urges that we have purely recognitional concepts (see Loar 1997), which apply to our sensations not by way of ties with other concepts, but solely in virtue of one's ability to recognize what type of sensation one has. And he urges that a perceptual model will better accommodate such concepts than a model based on thoughts.

But it's unlikely that any concepts for qualitative states are purely recognitional in this way. Rather, our concepts for mental qualities connect in crucial ways with our concepts for the physical properties that those mental qualities enable us to perceive; indeed, our concepts for mental qualities very likely derive from our concepts for perceptible properties. We don't have individual concepts for each type of mental quality, nor for each corresponding perceptible property. Still, we can conceptually single out every mental quality by its location within a suitable quality space, just as we can for the corresponding perceptible property. We distinguish properties of both types as being more or less similar to and different from other properties in their quality space. So our concepts for mental qualities are intimately tied both to concepts for other mental qualities and to concepts for perceptible properties (Rosenthal 1999a, forthcoming: Ch. 7).

Lycan's arguments rely on our folk-psychological conceptions of the relevant mental functioning. But there is also a folk-psychological consideration that suggests that thoughts are a better model for higher-order awareness than perception. The things perceiving gives us access to are physical objects or states of affairs; we perceive red objects, the growling of our stomachs, damage to our bodies, and the positions of our limbs, and we do so by way of characteristic sensations. But it is folk-psychologically odd also to speak of perceiving those sensations. No such oddness attends the idea that we have thoughts about those sensations and feelings; we can have thoughts about anything.

Folk psychology aside, the perceptual model has caused considerable theoretical mischief. Hume's famous problem about the self is due to his never "perceiving anything but the perceptions" (178/1939:634); he assumes that any awareness of the self would be perceptual (Rosenthal 2004). And some, like John Searle (1992:97), deny we are ever conscious of our mental states at all because they assume such awareness would have to be perceptual.

On balance, then, neither folk psychology nor empirical findings sustain higher-order perceiving, and the initial problem about higher-order qualities remains. The perceptual model, it seems, does not withstand scrutiny.
4. Dispositional higher-order thoughts

Sensing and perceiving are not, however, the only ways we are conscious of things. We are also conscious of things by having thoughts about them as being present. If I think, independently of any sensory input, of an object as being present, that's a way of being conscious of it. The requirement that the thought represent the object as being present echoes the ordinary case of perceiving, but no perceptual input is needed.

We can conclude, then, that our higher-order awareness of conscious states is in some way a matter of having thoughts about those states. Still, there are different versions of theories that appeal to such higher-order thoughts (HOTs).

Theorists who otherwise have little use for higher-order theories sometimes acknowledge that HOTs may be suitable to explain introspective consciousness (e.g., Block 1995:235). We are introspectively conscious of a state when we are not simply aware of that state, but aware of it in a deliberate, attentively focused way. Ordinary, nonintrospective consciousness, by contrast, occurs when one is aware of being in the state but not in this deliberate, attentively focused way. Because we consciously focus on introspected mental states, introspection involves actually being aware that we are conscious of those states.

So it's reasonable to explain being introspectively conscious of a state not as simply having a HOT about that state, but as having a HOT about it that is itself conscious. And if a state isn't conscious, introspectively or otherwise, there is no HOT. That suggests that a state's being conscious in the everyday, nonintrospective way results from something in between these two.

The natural conclusion is that a mental state's being nonintrospectively conscious consists in its being accompanied by a HOT when that HOT is not, itself, a conscious thought. When no HOT occurs, the target state isn't conscious; when there's a conscious HOT, the target is introspectively conscious. So when a HOT occurs that isn't conscious the target is conscious but not introspectively so.

A HOT model seems especially appealing for introspective consciousness because HOTs are then conscious; so we are aware of their occurrence. Since we aren't conscious of HOTs except when we introspect, the model is subjectively less inviting for nonintrospective consciousness; if HOTs aren't conscious, it seems subjectively that none occur. So it might seem that we must explain nonintrospective consciousness by appeal to something other than HOTs, or indeed higher-order awareness of any sort.

But there is another alternative. States that are conscious but not introspectively so fall in some way between introspectively conscious states and states that aren't conscious at all. So we need something that falls between having a conscious HOT and having no relevant HOT. Still, we needn't countenance nonconscious HOTs for the intermediate case. Perhaps a state is nonintrospectively conscious if, rather than being accompanied by some HOT, it is simply disposed to be accompanied by a conscious HOT. This dispositional alternative makes room for the temptation to avoid positing HOTs that aren't conscious. And it fits with the pretheoretic idea that a state's being conscious is at least partly the dispositional properties of its being available for introspective access.

Peter Carruthers (2000) has defended such a dispositional HOT theory, though not by appeal to the foregoing line of argument. He explicitly recognizes the possibility of nonconscious HOTs, but argues that the number of actual HOTs needed to capture all the phenomenological detail and subtle variation in our conscious experience at any moment would be prohibitively large. It is implausible, he claims, "that so much of our cognition should be occupied with formulating and processing the vast array of higher-order thoughts necessary to render our experience conscious at each moment of our waking lives" (2000:221).

Carruthers recognizes that one could respond to this "objection from cognitive overload" by urging, with Dennett (1991:Ch. 11) that our conscious experience is not as rich in detail as it seems. But it's not clear that this reply helps. Dennett credibly argues that our sense of a rich, detailed phenomenological scene is illusory, since phenomenology suggests far more detail than we can actually discriminate. But the consciousness of our experience is a matter not of what we can discriminate, but of how our qualitative experience seems to us. And, as Dennett concedes, it does seem to be richly detailed.

Still, it's not easy to know how to assess the objection from cognitive overload. Cortical resources are far greater than needed for most, perhaps all, of our cognitive processing. So sufficient cortical resources are doubtless available to accommodate HOTs for all our conscious states. The consciousness of experience seems effortless, whereas cognitive processing requires some conscious effort and attention; so we would not have guessed that experience's being conscious calls for substantial cognitive resources. But we should be wary of letting such commonsense considerations affect our appraisal of a theory.

Carruthers urges that occurring HOTs would make us conscious of all the fine-grained distinctions among our visual perceptions only if we had "a concept for each just-discriminable shade of colour" (2000:222). But, as we saw
earlier, we are conscious of such discriminations largely in comparative terms; so we needn’t have distinct concepts for each shade.

Like Lycan, Carruthers holds that our higher-order concepts for qualitative states are purely recognitional, and he sees this as a difficulty for a theory based on occurrent HOTs. Just as our perceptual awareness of physical green guides the application of our recognitional concept of that color, so our higher-order recognitional concepts, he claims, must also apply by way of some independent, nonconceptual awareness of the thing recognized. But a theory on which we are conscious of our qualitative states solely by way of occurrent HOTs allows no room for an independent, nonconceptual awareness that could guide the application of the recognitional concepts that figure in those HOTs. So the application of these concepts would be groundless on such a theory, and would not parallel that of our first-order recognitional concepts (this volume: § 3).

But such a parallel would be surprising if the relevant concepts were indeed purely recognitional. Some mental occurrence must guide the recognizing of things, and the only mental occurrence that might do that for the recognizing of physical colors is our sensations or perceptions of those colors. But sensations of green could guide not only our recognition of physical green, but also our recognition of the corresponding mental quality of green, which characterizes such sensations; one needn’t in addition have some independent awareness of that sensation. Carruthers claims that reflection reveals that our recognition of sensations relies on some independent awareness of those sensations. But reflection is the unusual situation in which we have some third-order, introspective awareness, and in that special case our second-order awareness is indeed independent of our introspecting.

It’s in any case unlikely, as noted earlier, that the concepts that figure in occurrent HOTs are purely recognitional. Rather, our concepts single out mental qualities by locating them within the quality space distinctive of the relevant sensory modality. And, since this mimics the way our concepts for perceptible physical properties locate those properties within a corresponding quality space, it captures the parallel that actually does obtain between the higher- and first-order concepts.

Carruthers favors a recognitional model of our higher-order awareness because he thinks it helps avoid difficulties about conscious qualitative states. But the appeal of a recognitional model doubtless also reflects the continuing hold of the traditional thesis that mind is transparent to consciousness. Since recognizing is factive, being aware of something by recognizing it can’t go wrong. So, if our consciousness of our mental states is recognitional, that consciousness will automatically get things right. But consciousness does not always represent our mental states accurately. Consciousness seems infallible because it never shows itself to be mistaken and it’s tempting to think that there’s no other way to know what mental states one is in. But consciousness is not the only way to determine what mental state one is in, and there is sometimes compelling independent evidence that goes against what consciousness tells us.

Carruthers argues that there could have been little evolutionary pressure for the routine generation of HOTs for all the thoughts and experiences in daily practical activity. True; but our daily experience is seldom if ever conscious in respect of all the detail relevant to such practical activity. Armstrong’s (1980:60) familiar example of a person driving long distance on psychological automatic pilot helps here. It is unlikely that the person’s relevant visual states literally fail to be conscious; rather, they simply aren’t conscious in respect of the rich detail needed for the task of driving. Occurrent HOTs would fail, in this case, to represent that rich detail. This often happens with activities that require concentrated effort: perceptual states that figure in an activity are seldom conscious in respect of all the detail relevant to carrying out the activity.

Indeed, it is in any case puzzling what evolutionary pressure there could have been for mental states to be conscious, whatever the explanation of their being conscious. Evolutionary pressure on mental functioning operates only by way of interactions that such functioning has with behavior. And mental functioning interacts with behavior solely in virtue of its intentional and qualitative properties. If a mental state’s being conscious does consist in its being accompanied by a higher-order state, that higher-order state would contribute to the overall causal role, but this contribution would very likely be minimal in comparison with that of the first-order state. So there could be little adaptive advantage in states’ becoming conscious.  

In any case, evolution often occurs without clear adaptive advantage; there are side effects whenever a feature is selected. Doubtless adaptive pressures favored high cortical mass, and that may have resulted in creatures with far more cortical resources than needed for everyday functioning. And such excess cortical area might then have come to subserve states that represent other cortical states. So creatures could well have come to have the actual HOTs needed for their mental states to be conscious, even though their being conscious is not itself especially adaptive.

Edmund T. Rolls (this volume: § 2) has made a compelling case that HOTs might, in some kinds of case, make a pivotal contribution to overall adaptive value. Plans for action sometimes involve a sequence of first-order hypothetical thoughts, and it may well sometimes be useful to review these hypotheticals to evaluate and, if necessary, correct them. Such evaluating and correcting,
Rolls argues, would likely involve thinking about those first-order hypotheti­cal thoughts, which would require HOTs that can reveal the syntactic connec­tions among those thoughts. And these HOTs would make one conscious of the first-order thoughts.

This proposal reflects the pretheoretic idea that consciousness is somehow important for the critical evaluation of rational thinking. But the need to evaluate one's thinking critically is likely to produce evolutionary pressure for HOTs only if there is no easier way to get satisfactory results. And revising plans of action can result instead simply from having additional first-order thoughts about the relevant situations and the connections that hold among them. That mechanism for revision is especially plausible for our distant, less cerebral an­cestors, on whom such evolutionary pressures could have had an effect. Nor does the revising of plans require identifying difficulties in one's first-order thinking by having HOTs about that thinking; even nonconscious conflicts often cause mental tension that in turn elicits compensatory adjustments in our thoughts and plans.

Carruthers's arguments against actual HOTs to one side, it's not obvious that dispositions for HOTs can help. The principle advantage of higher-order theories is that they explain how it is that we're transitively conscious of our conscious states. But being disposed to have a thought about something doesn't make one conscious of that thing, but only potentially conscious of it. A state's being conscious in the everyday, nonintrospective way does dispose one to be introspectively conscious of that state. But a higher-order theory must also explain ordinary, nonintrospective consciousness. Nor does a dispositional theory have any advantage in explaining why a nonintrospectively conscious state exposes one to introspect that state; we can expect that an occurrent HOT that isn't conscious exposes one to become conscious of that HOT.

It's also not initially obvious how dispositions could help with the difficulty about cognitive overload. Dispositions are themselves states, no less than actual HOTs; so dispositions to have HOTs should be no less cognitively demanding than actual HOTs. Again, it may be only the absence of conscious cognitive effort that makes dispositions seem intuitively preferable.

Carruthers's answer to both worries, about becoming conscious of con­scious states and avoiding cognitive overload, appeals to the particular the­ory of intentional content he endorses. A state's intentional content, he urges, is in part a matter of what other mental states and what behavior that state is disposed to cause. This causal potential takes a special form when a state is suitably connected to a psychological subsystem capable of discriminating among types of states, which Carruthers thinks of as the mind-reading system.

A state then has the potential to cause in that mind-reading system a HOT about the state itself. And, by itself, that causal potential gives the state a cer­tain higher-order content, in addition to its ordinary content in virtue of which it represents features of nonmental reality. Conscious experiential and inten­tional states thus have dual content; an experience of red, for example, will also have the content seems red or is an experience of red. And it will "have these [higher-order] contents categorically, by virtue of the powers of the HOT consumer system, in advance of any HOT being tokened" (2000:242). This higher-order content comes automatically with the availability of the states to the mind-reading system.

Cognitive overload is thus avoided, Carruthers urges, since generating this higher-order content makes no cognitive demands beyond those incurred in the states' simply having first-order contents and their being connected to the mind-reading system. And these higher-order contents explain how it is that we're aware of our conscious states, as Carruthers puts it, how states "can come to acquire the properties of subjectivity and what-it's-likeness distinctive of phenomenal consciousness" (2000:242; Carruthers's emphasis).

One might prefer a theory of consciousness not to be hostage to such a heavy load of other controversial views. But that aside, there are difficulties for this account. Each first-order state has some higher-order content in virtue of the state's availability to the mind-reading system, and that higher-order content results in one's being conscious of the state, so that there's something it's like for one to be in the state. But, if it's sufficient for a state to be conscious simply that it have some particular first-order content and that it be available to the mind-reading system, it's unclear how that state, or any state of its type, could ever occur without being conscious. It cannot be its first-order content is sometimes different, since we type states by way of their content. And, in any case, every suitably connected first-order state will have such higher-order content. Nor can it be that the mind-reading system sometimes shuts down, since other states remain conscious.

So it must be that individual states are sometimes available to the mind-reading system and sometimes not. But it's implausible that such shifts in availability would occur, since the first-order states themselves presumably occur in systems with fairly stable connections to the mind-reading system. And making such shifts would doubtless be at least as cognitively demanding as simply having occurrent HOTs. A theory cast in terms of occurrent HOTs, by contrast, has no difficulty with a state's shifting between being conscious and not; occurrent HOTs presumably come and go, just as other occurrent thoughts do.
Intentional content, on Carruthers's view, is a matter of a state's causal potential in respect of other mental states and behavior. The specifically higher-order content in virtue of which a state is conscious consists in that state's having the causal potential to affect something external to itself, namely, the mind-reading system.

But Carruthers acknowledges that conscious first-order states "have these [higher-order] contents categorically, ... in advance of any HOT being tokened." It is a state's having this higher-order content which explains there being something it's like for one to be in that state, and thus that state's being conscious. So it is the actual occurrence of a state with higher-order content, not merely a disposition to have a higher-order state, that explains a state's being conscious. As with most theories of content, it's a dispositional matter according to Carruthers that these states have their higher-order content, but it's an occurrence state with higher-order content that explains consciousness. It's the theory of content that's dispositional, not the theory of consciousness.9

Conscious first-order states have, Carruthers holds, higher-order content "in advance of any HOT being tokened." What, then, would an occurrence HOT add? Since first-order states are already nonintrospectively conscious without any occurrence HOT, the only possibility is that the an occurrence HOT would result in the first-order state's becoming introspectively conscious. But when we introspect, we are aware of our being conscious of the introspected state; so the occurrence HOT is itself also conscious. Despite Carruthers's recognition that HOTs need not be conscious, the only role he leaves open for occurrence HOTs is the role conscious HOTs have in introspection.

5. Intrinsic higher-order thoughts

If a state is conscious in virtue of some higher-order awareness of that state, it's natural to assume that this higher-order awareness consists in the occurrence of a distinct state with suitable higher-order content. Absent some reason to the contrary, distinct mental functions call for distinct states. And distinct higher-order states also explain how states shift between being conscious and not, which would be puzzling if the higher-order were built into the first-order target itself.

Nonetheless, a number of theorists have urged, following Franz Brentano, that the higher-order content in virtue of which we are conscious of our conscious states is intrinsic to those states (Gennaro, this volume, 1996; Kriegel 2003; Natsoulas 1999). As Brentano put it, all mental acts "apprehend [themselves] incidentally" (1973/1874:128). Every conscious state, on this view, is about itself as well as some nonmental reality.

Carruthers's view actually seems to be a version of such an intrinsic theory. The higher-order content in virtue of which a state is conscious, on his view, belongs categorically to the very state that is conscious. Though the higher-order content is a matter of the state's having connections with the mind-reading system, that content is a property of the target state one is conscious of, not a distinct higher-order state.

When a state is conscious, we are conscious of that state. But, except for the special case of introspective consciousness, we are not also conscious of being conscious of the state; it seems subjectively that we are conscious of only one state. So, if one relied on consciousness to reveal mental functioning, one would conclude that, when a state is conscious, there aren't two states but only one. This consideration has led some, such as Brentano, Karen Neander (1998), and Joseph Levine (2001:105, 168), to reject a higher-order theory altogether. But it has led others, such as Gennaro (this volume) and Natsoulas (1999), to argue instead that the higher-order content in virtue of which a state is conscious is intrinsic to that very state.

This argument assumes that consciousness reveals everything about our mental functioning, or at least everything relevant to the issue at hand. But we know that this isn't so, since there are many mental states that aren't conscious. Consciousness does reveal the phenomenological data that a theory of consciousness must do justice to. But to save these phenomena, we need only explain why things appear to consciousness as they do; we need not also suppose that these appearances are always accurate. A theory of consciousness explains how things appear to consciousness.

When a state is conscious, one is conscious of that state, though not typically aware of being conscious of it. But that doesn't support a theory on which only a single state occurs, nor does it undermine a higher-order theory. Rather, it shows only that the higher-order awareness of the state one is conscious of is rarely conscious. We needn't suppose that there's only one state, but merely that there's only one conscious state.

There is a related objection to the view that the higher-order awareness is due to a distinct state, an objection that does not rely on how things appear to consciousness. If the higher-order awareness is distinct from its target, the two might occur independently. It's theoretically unproblematic if the target occurs without the higher-order awareness; then the target simply isn't conscious. But, if the states are distinct, the higher-order awareness might also occur without any target. And, even if there's a target, the higher-order awareness might fail
to represent it accurately. Since it might seem that there is no good answer to how things would be phenomenologically in these kinds of case, perhaps this possibility tells against a higher-order theory (Levine 2001:108; Neander 1998:420).

One might advance an intrinsic theory as a way to meet this difficulty (Gennaro, this volume; Kriegel 2003). If the higher-order awareness is actually part of its target, it plainly can't occur without the target. And perhaps it also cannot then misrepresent that target.

It's not obvious that absent or misrepresented targets actually lead to any difficulty, phenomenologically or otherwise (see §6). But an intrinsic theory would in any case have no advantage in meeting whatever difficulties do occur. The distinction between an absent target and a misrepresented target is in an important way arbitrary. Suppose my higher-order awareness is of a state with property P, but the target isn't P, but rather Q. We could say that the higher-order awareness misrepresents the target, but we could equally well say that it's an awareness of a state that doesn't occur. The more dramatic the misrepresentation, the greater the temptation to say the target is absent; but it's plainly open in any such case to say either. The two kinds of case, moreover, should occasion the same kinds of phenomenological perplexities, if any. A higher-order awareness of a P state without any P state would be subjectively the same whether or not a Q state occurs. The first-order state can contribute nothing to phenomenology apart from the way we're conscious of it.

On an intrinsic theory, we do have a nonarbitrary reason to say that the target is never absent; whatever state the higher-order awareness is part of counts as the target. But it's not on that account obvious why misrepresentation could not occur. Simply being intrinsic to the target does not guarantee that the higher-order awareness will be accurate. And, since misrepresentation would occasion the same difficulties, if any, that would occur for an absent target, an intrinsic theory has no advantage in this connection over a theory on which the higher-order awareness is distinct from its target.

On Carruthers's dispositional theory, our higher-order awareness is intrinsic to its target, since that target has higher-order content in virtue of its tie to the mind-reading system. That tie enables the mind-reading system to "generate[] recognitional concepts of experience, riding piggy-back on the first-order contents of experience" (2000:241). But that tie between target and mind-reading system cannot protect that higher-order content against error. However accurate the mind-reading system may be, the connection any particular state has to that system can go wrong, and with it the mind-reading system's capacity to recognize that state and hence the state's higher-order content. A dispositional theory has no advantage in ensuring accuracy of higher-order content.

We're conscious of our conscious states in a way that seems, subjectively, to be direct and unmediated. And one might urge that this supports an intrinsic theory, since nothing would mediate between the higher-order awareness and its target if that awareness is actually part of the target. But the datum we need to explain is not actual immediacy, but rather subjective immediacy. And these needn't go together, at least not unless we assume that mental functioning is transparent to consciousness. It could be that nothing actually mediates between our higher-order awareness and its target even though it subjectively seems as though something does.

More likely, something might mediate even though we are subjectively unaware of anything doing so. Typically we aren't conscious of any HOTs; it doesn't subjectively seem as though there are any. So, if we're conscious of our conscious states by way of of distinct HOTs that aren't conscious, it will seem as though nothing mediates between our consciousness of those states and the states themselves. And, when those distinct HOTs are conscious, it will still seem as though nothing mediates between them and their targets if we're conscious of no inference from which the HOTs seem to arise. The HOTs will then seem subjectively to be spontaneous and uncaused. A theory that appeals to distinct HOTs can thus do justice to phenomenological immediacy, whereas an intrinsic theory must somehow explain why the actual lack of mediation issues in the corresponding appearance. It is likely that an intrinsic theory will appeal here to the very same considerations as a theory cast in terms of distinct HOTs.

Not only do the advantages claimed for an intrinsic theory fail to hold up; such a theory faces several important challenges. For one thing, it must explain why, if the higher-order awareness is intrinsic to the target, that higher-order awareness is about the target only in respect of its other, first-order content.10

An intrinsic theory must explain what happens when a state goes from being nonintrospectively conscious to being introspectively conscious. Does the target state come to have third-order content in virtue of which one comes now to be aware of its second-order content? Or does one instead become aware of that second-order content by way of a distinct third-order state? Neither answer is theoretically satisfying. It would be surprising if one's becoming introspectively conscious of a state were a matter of that state's actually changing, by taking on new, third-order content. But, if a distinct third-order state is responsible for introspective consciousness, why not a distinct second-order state for ordinary, nonintrospective consciousness?
The need to explain introspection recalls the initial point that an intrinsic theory may have difficulty giving an informative explanation of how states shift between being conscious and not. On an intrinsic theory, a state's coming to be conscious consists in its actually changing its content, and its ceasing to be conscious consists in a loss of intrinsic content. It would be theoretically more satisfactory to assume that such shifts leave the state unchanged.

Underlying all these issues is a pressing need for an independent way of individuating states, which doesn't beg the question between intrinsic and extrinsic higher-order awareness. It might seem that individuating mental states is inevitably arbitrary, and so wouldn't preclude either position. But there is a compelling consideration that undermines individuating in a way that makes the states of higher-order awareness intrinsic to their targets.

Intentional states differ not only in content, but in mental attitude as well. One can hold toward any particular content a variety of distinct attitudes, such as believing, desiring, anticipating doubting, wondering, and many others. And we individuate intentional states so that no state exhibits more than one mental attitude.11

But intentional states don't make one conscious of the things they are about unless those states exhibit a suitable attitude toward that content. Wondering, doubting, and desiring about something do not make one conscious of that thing. For an intentional state to make one conscious of the thing it's about, it must involve an assertoric attitude towards its content.

Consider a conscious case of an intentional state, such as wondering or doubting, which exhibit some nonassertoric attitude. That state is conscious in virtue of some higher-order awareness of it. Since we have ruled out higher-order sensing and perceiving, only some higher-order intentional content will do, and one must hold an assertoric attitude toward that higher-order content. So, when the conscious state is an intentional state with a nonassertoric attitude, the higher-order awareness must be distinct from its target itself, since we individuate intentional states so that no single state has two distinct attitudes. Considerations of mental attitude decisively undermine an intrinsic higher-order theory cast in terms of HOTs.

Brentano's version of an intrinsic theory is evidently perceptual; his prime example of a conscious state is hearing, and he argues that a conscious act of hearing apprehends itself (1973/1874:128). And hearing something does make one conscious of it. But, as we saw earlier, all perceiving requires some mental quality, and the mental qualities special to audition enable us to hear only sounds, not auditory states themselves. Perceptual modality tells against an intrinsic higher-order perception theory just as mental attitude tells against an intrinsic HOT theory.

6. Higher-order thoughts and the intentional stance

The advantage of a higher-order theory is that it explains how it is that we're conscious of our conscious states. We are conscious of those states in virtue of some higher-order awareness. That higher-order awareness, moreover, explains not simply that we are conscious of those states, but also the particular way we are conscious of them. As with any case of being aware of something, the representational character of a higher-order awareness determines just how one is conscious of the state that the awareness represents.

This has implications for cases in which targets are absent or inaccurately represented. If I consciously take something I see to be a cow when it's actually a horse, phenomenologically it's as though I consciously see a cow. Similarly, if I am conscious of myself as being in a P state, it's phenomenologically as though I'm in such a state whether or not I am. If I'm not in a P state, that will make a difference to my overall mental functioning, just as it may make a difference to my interactions with my environment if the thing I take to be a cow is actually a horse. But the phenomenology is determined solely by the way I'm aware of things, whether perceived physical objects or my own mental states.

Typically we see things accurately, and it's also likely that consciousness ordinarily represents correctly what mental states we are in. But misrepresentation of such states can happen, (see, e.g., Nisbett & Wilson 1977), and it is an advantage of a higher-order theory that it accommodates such occurrences.

That consciousness sometimes represents us as being in states we have independent reason to think we aren't in is an important theme of Dennett's (1991). When we see wallpaper decorated with repeating, identical objects or Warhol's repeating, photographic portraits of Marilyn Monroe, it seems subjectively that we see all the repeating tokens at once with equal resolution. But that cannot be; we see most of those tokens parfoveally, and parafoveal resolution is far below that of foveal vision. Here as elsewhere, we're aware of our representations in ways that diverge from what we know those representations must be (1991:354; see Ch. 11 passim).

In stressing this kind of occurrence, Dennett's discussion is congenial to higher-order theories. But there is another respect in which it isn't. Suppose I see somebody I don't know, but my memory of a friend causes me consciously to mis perceive the person I see as my friend. Dennett famously urges that there
is no fact of the matter about whether the memory contaminates my current perception before or after that perception becomes conscious. On a higher-order theory, by contrast, there is precise moment when the perception becomes conscious, even if we can't easily determine when that is; it becomes conscious just when the higher-order awareness occurs.

Dennett's denial of a precise moment for the contamination seems to conflict with his recognition that consciousness sometimes mischaracterizes the way we represent things. If the way we are aware of a representation mischaracterizes that representation, there must be two distinct states, the representation and our awareness of it. But then the occurrence of this distinct awareness of a representation should provide the exact moment at which the representation becomes conscious.

But there is no conflict between Dennett's denial of an exact moment of consciousness and his recognition that consciousness sometimes mischaracterizes our representations. Dennett sees the representations that consciousness mischaracterizes not as the commonsense states of folk psychology, but as subpersonal events of content fixation. Though we are conscious of ourselves as being in various folk-psychological states, the representational states that actually occur are not those folk-psychologically taxonomized states, but merely those subpersonal events of content fixation. It is such subpersonal events, not folk-psychological states, which consciousness mischaracterizes when we seem to see many Marilyns at once with equal resolution, and similarly for other such cases.

These subpersonal events "are precisely locatable in both space and time" (1991:113). But that doesn't, according to Dennett, enable us to resolve the puzzles about timing, since he holds that those subpersonal events do not correspond to the folk-psychologically taxonomized mental states we're conscious of ourselves as being in. There are, strictly speaking, no first-order folk-psychological states, only subpersonal events of content fixation and our taking of ourselves to be in folk-psychologically taxonomized states (1991: Ch. 10). When I take myself to perceive somebody, I am in effect conscious of myself as perceiving that person (Rosenthal 2000). But there is no folk-psychological perceiving whose contamination by memory occurs at some particular moment, only the the interplay of subpersonal events, which do not correspond in any straightforward way to the folk-psychological states we interpret ourselves as being in.

We can thus see Dennett's view as a kind of higher-order theory on which there are acts of higher-order interpretation, but not their apparent first-order, folk-psychological targets. Thus Dennett holds that there is no way that things seem apart from the way that they seem to seem, no "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" (1991:132). The only seeming we can sensibly speak of is higher-order seeming.

This eliminativism in respect of first-order folk-psychological targets fits comfortably with Dennett's well-known view that intentional states are simply real patterns of behavior discernible from the intentional stance (1981). But that view aside, there seems little reason to deny that suitably grouped events of content fixation often constitute folk-psychological states. And if so, not only do HOTs occur, but also the first-order states those HOTs represent us as being in.

7. Distinct higher-order thoughts

The foregoing considerations all point to a higher-order theory cast in terms of distinct, occurrent HOTs (Rosenthal 1986, 1990/1997, 2002, forthcoming). Such a theory has all the advantages of other higher-order theories, and avoids their shortcomings.

Many of the difficulties that affect higher-order theories come from trying to determine the nature of the higher-order awareness by appeal to folk-psychological considerations or to how things seem subjectively. But we are rarely conscious of that higher-order awareness; so subjective impressions and folk psychology can have little useful to tell us. We can best think of that awareness as being a theoretical posit, designed to explain the phenomenal appearances of our conscious mental lives. It is success in that explanatory task that will establish the existence of that higher-order awareness and tell us about its nature. By that criterion, distinct, occurrent HOTs do the best job.

That HOTs are theoretical posits doesn't mean that we are never aware of them in an intuitively direct way; indeed, that's just what happens when we introspect. But we learn in the first instance that HOTs occur by developing a suitable theory to explain the subjective appearances of consciousness.

When qualitative states are conscious, there is something qualitative that it's like for one to be in those states; when they aren't conscious, there is nothing it's like. A higher-order theory must explain how being aware of a state results in there being something qualitative that it's like for one to be in such states. Higher-order sensing or perceiving initially seemed best suited to do so, since the qualitative character of such higher-order states might explain why being conscious of a state results in its lighting up qualitatively.
Since there are no higher-order qualities, that strategy cannot work. But higher-order qualities couldn't help in any case. A higher-order quality would explain why there is something qualitative that it's like for one to be in a particular state only if the higher-order quality is conscious. But, since higher-order qualities would be conscious only when we introspect, they couldn't help with nonintrospective consciousness, which is our principle concern. And when we do introspect, we aren't conscious of any such higher-order qualities, but only of the qualities of the first-order states. Higher-order qualitative states cannot explain why there is something qualitative that it's like for one to be in conscious, first-order qualitative states.

But it may seem even less likely that HOTs could help. HOTs, like other thoughts, have no qualitative properties; their only mental properties are intentional. How, then, could being conscious of a state by having a HOT about it result in there being something qualitative that it's like for one to be in that state?

It is important to be clear about just what would count as an answer to this question. One might hold that a answer would be satisfactory only if it resulted in its tracing a purely rational connection between the occurrence of a HOT and there being something qualitative that it's like for one to be in the target state. Thus Levine argues that we can explain why water boils at 212°F at sea level in ways that make it inconceivable that it wouldn't, and we should demand no less for a satisfactory answer to the question about conscious qualitative character.

The explanation of water's boiling at 212°F involves, as Levine notes, a "rich elaboration of" chemical theory (2001:79). But, as the history of chemistry testifies, such theory is hardly a purely rational matter, having to do only with what's conceivable. It's inconceivable, relative to chemical theory, that water would fail to boil at 212°F, but chemical theory is a matter of how things are, not just what we can conceive. Chemical theory is not only surprising and unexpected, but something we can conceive to be false.

If we had a parallel theory for qualitative states, we could doubtless explain why such states phenomenologically light up under certain conditions. And relative to that theory, it would seem inconceivable that it could have been otherwise. But can cannot demand that such a theory should antecedently seem rational, or a matter simply of what we can and cannot conceive. Like chemistry, that theory will come from examining how things actually are.

Once we see that such a theory could not be a matter simply of what's conceivable, we can develop, in advance of such a theory, a good theoretical hunch about how conscious qualitative character can be explained. And there is reason to think that HOTs actually do make a difference to what it's like qualitatively for one to be in particular states, and hence that they may well be responsible for there being something qualitative that it's like.

Somebody who hasn't learned to distinguish the gustatory sensations of wines or the sounds produced by oboes and clarinets may well be conscious of these distinct mental qualities as though they were the same. And it sometimes happens that these mental qualities come then to seem distinct to such a person by that person's learning words for the distinct mental qualities. This is typically in a context in which one consciously concentrates on those mental qualities, so the new words are used for the mental qualities, not the physical stimuli, as Carruthers (2000:240) and others have claimed. As one learns the new words, a subjective difference emerges between mental qualities that had previously been subjectively indistinguishable.

What might explain this change? Learning new words makes a difference only if one learns their meanings, which means learning to deploy the corresponding concepts in thoughts. And, since the concepts apply to mental qualities, the thoughts in which those concepts figure will be about mental states that exhibit those qualities; they will be HOTs. So learning how to have HOTs about states as exhibiting distinct mental qualities results somehow in the subjective emergence of a conscious qualitative difference between them; HOTs do make a difference to what it's like qualitatively for us to be in various mental states. And, if having new HOTs can make a difference in what it's like qualitatively for one to be in a particular state, having a HOT can presumably make the difference between there being nothing qualitative that it's like for one and there being something that it's like for one.

Carruthers claims that this account might be right about the causal antecedents of there being something it's like for one, but not about what constitutes there being something it's like (2000:240). But it's unclear what kind of constituting Carruthers has in mind. The theory doesn't attempt a conceptual analysis of what it's like, any more than Carruthers's does. But the theory does argue that there is something qualitative that it's like for one to be in a state whenever one has a suitable HOT that one is in that state and the HOT characterizes the state in qualitative terms. If so, having such a HOT does constitute there being something qualitative that it's like for one to be in the state. The theory claims not that HOTs cause qualitative consciousness to occur, but that the having of suitable HOTs is what it is for qualitative states to be conscious.

Carruthers maintains that identifying the having of a suitable HOT with there being something it's like for one to be in the target state means giving up on a reductive explanation of qualitative consciousness (this volume: §4). But whatever one's views on reductive explanation, this is a theoretical identifica-
tion, supported both by data and theoretical considerations, and it thus carries substantial explanatory force.

It might be thought that new conscious qualities emerge not because we come to be conscious of existing qualities in new ways, but by what psychologists call perceptual learning. Learning new words might actually result in new mental qualities coming to occur, rather than our simply becoming conscious of mental qualities that were already there. But this is unlikely, since, unlike perceptual learning, the effect is dramatic and rapid.

If mental qualities were automatically conscious, new conscious qualities would imply new qualities, and perceptual learning would be the only possible. But mental qualities need not be conscious. What it is to be a particular mental quality is a matter of its position in the quality space of the relevant modality. And we can construct such quality spaces by reference to the perceptual discriminations a creature can make (Rosenthal 1999a, 1999b), independent of whether the relevant qualitative states are conscious.

Levine objects that we "have a more determinate and substantive conception of [mental qualities] ... , a conception that is not exhausted, or adequately captured by the formal description of a location in a similarity space" (2001:107). Perhaps so; but, as with other commonsense phenomena, our conception of the properties of our qualitative states may well not be true to the actual nature of these properties.

HOTs explain how we're conscious of our conscious states. When those states are qualitative, HOTs characterize their qualitative properties in terms of how much they resemble and differ from other qualities in the relevant modality. So a particular quality often seems different to consciousness when it occurs with other qualities, since we can then be conscious of it in comparison with others. And, because HOTs result in our being conscious of ourselves as being in various qualitative states, the occurrence of a HOT about a qualitative state will be subjectively the same whether or not that state occurs.

By the same token, the subjective appearances will be the same whether the target causes the HOT or the HOT arises independently. Doubtless targets are often causally implicated in the occurrence of HOTs, though other factors must also figure, since otherwise the target would always cause a HOT and hence always be conscious. But what matters to the phenomenal appearances is simply what state the HOT makes one conscious of oneself as being in.

Since HOTs are distinct from the states they are about, it's possible for a HOT to occur without its target. And, since having a HOT results in one's being conscious of oneself as being in a particular state, a HOT's accompanying its target is subjectively indistinguishable from a HOT's occurring in the absence of that target. When a target doesn't occur, one isn't in the state one is conscious of oneself as being in, but that's not a problem. All that matters to one's being in a conscious state is what it's like for one, and what it's like for one in that case is that one is in the state in question. Being in a conscious state is not being in that state and being conscious of being in it, but simply being conscious of oneself as being in the state.

Language is likely not necessary for thinking, and the conceptual resources for HOTs are not too demanding for infants and nonhuman animals to have. A HOT has the content that one is in a particular state, a state we would characterize as mental, though the HOT need not do so. And the reference HOTs make to a self need involve nothing more than a distinction between oneself and everything else (Rosenthal 2004).

A higher-order theory is the only type of theory that can explain how it is that we're conscious of our conscious states. It seems that a higher-order theory based on distinct, occurrent HOTs can best accomplish that task.

Notes

1. 1964-1975: VII, 246 (Fourth Replies). Also: "the word 'thought' applies to all that exists in us in such a way that we are immediately conscious of it" (1964-1975: VII, 160 [Geometrical Exposition of Second Replies]). (Translations and emphasis mine.)
2. For the first claim, see 1984: II, Metaphysics XII, 7,1072b20-23; for the second, see 1984: II, Nicomachean Ethics IX, 9, 1170a31-2; 1984:1, de Anima III, 2,425b 12-20.
3. The late 19th century shift to describing mental states as being conscious or not is very likely due to the growing recognition at that time that many mental states fail to be conscious. We try to describe things in a way that draws some contrast. Saying that we're conscious of all mental states draws a clear contrast, since we're not conscious of everything, whereas saying that all mental states are conscious leaves the intended contrast obscure, since we don't describe other comparable things as conscious.
4. I'll use 'conscious' and 'aware' and their cognates as equivalent.
5. The term derives from Kant (1998/1787:174 [A22/B37]); Locke had earlier written of "internal Sense" (1975:105 [II, i, 4]).
6. I appeal here to Lycan (this volume: §6), which he generously made available in advance of publication.
7. The relevant accessibility here is specifically introspective, as against the more general informational availability that underlies Dennett's contention that "consciousness is cerebral celebrity" (1993:929). In that way it also differs from Block's notion of access consciousness, on which a state's content is "posed to be used as a premise in reasoning... and for [the] rational control of action and... speech" (1995:231).
8. So it’s not a problem for higher-order theories, as Dretske (1995:117) urges, if they can’t explain such adaptative advantage.

9. As noted earlier, Carruthers claims that, because the concepts of mental qualities that figure in HOTs are purely recognitional, there must be an independent, nonconceptual awareness of the qualities such HOTs are about. His view provides for such an independent awareness, since it holds that the higher-order content that conscious qualitative states have in virtue of their connections with the mind-reading system is experiential content. Such content is “analog relative to a certain conceptual repertoire,” since it “admit[s] of significantly more variations than there are concepts to classify them” (2000:134; Carruthers’ emphasis); so Carruthers holds that it isn’t conceptual. Such experiential contents also allow for there to be distinct higher-order contents for each mental quality, which Carruthers believes is needed.

Carruthers denies that this is an inner-sense view, since he regards such analog content as a species of intentional content (2000:232). But since analog content is experiential content, the view does appeal to higher-order perception (cf. Carruthers, this volume, esp. §3). And it in any case doubtful that experiential content is wholly nonconceptual.

I am grateful to Robert Lutzker for pressing me on a number of the issues discussed in this section.

10. Brentano actually assumed that the higher-order awareness was about the entire state, so that the higher-order content was also conscious (1973/1874:129), though rarely attended to. But that content is actually seldom conscious.

11. As Brentano himself in effect observed, in noting that we individuate intentional states by reference to the mental act being performed (1973/1874:127). Plainly two distinct attitudes would issue in two distinct mental acts.

References


Carruthers, Peter (this volume). HOP over FOR, HOT theory.


Dennett, Daniel C. (1993). The message is: There is no medium. Philosophy and phenomenological research 53,4 (December), 919-931.


Gennaro, Rocco J. (this volume). Higher-order thoughts, animal consciousness, and misrepresentation: A reply to Carruthers and Levine.


Lycan, William G. (this volume). The superiority of HOP to HOT.


Rolls, Edmund T. (this volume). A higher order syntactic thought (HOST) theory of consciousness.


Higher-Order Theories of Consciousness

An Anthology

Edited by

Rocco J. Gennaro
Indiana State University