DAVID M. ROSENTHAL

TWO CONCEPTS OF CONSCIOUSNESS

(Received 8 July, 1985)

No mental phenomenon is more central than consciousness to an adequate understanding of the mind. Nor does any mental phenomenon seem more stubbornly to resist theoretical treatment.

Consciousness is so basic to the way we think about the mind that it can be tempting to suppose that no mental states exist that are not conscious states. Indeed, it may even seem mysterious what sort of thing a mental state might be if it is not a conscious state. On this way of looking at things, if any mental states do lack consciousness, they are exceptional cases that call for special explanation or qualification. Perhaps dispositional or cognitive states exist that are not conscious, but nonetheless count as mental states. But if so, such states would be derivatively mental, owing their mental status solely to their connection with conscious states. And perhaps it makes sense to postulate nonconscious versions of ordinary mental states, as some psychological theories do. But if consciousness is central to mentality in the way this picture insists, any such states are at best degenerate examples of mentality, and thus peripheral to our concept of mind.

This picture is both inviting and familiar. But there are other features of the way we normally think about mind which result in a rather different conception of the relation between consciousness and mentality. We often know, without being told, what another person is thinking or feeling. And we sometimes know this even when that person actually is not aware, at least at first, of having those feelings or thoughts. There is nothing anomalous or puzzling about such cases. Even if it is only seldom that we know the mental states of others better than they do, when we do, the mental states in question are not degenerate or derivative examples of mentality. Moreover, conscious states are simply mental states we are conscious of being in. So when we are aware that somebody thinks or feels something that that person is initially unaware of thinking or feeling, those thoughts and feelings are at first mental states that are not also conscious states. These considerations suggest a way of

Philosophical Studies **49** (1986) 329–359.
© 1986 *by D. Reidel Publishing Company*

looking at things on which we have no more reason to identify being a mental state with being a conscious state than we have to identify physical objects with physical objects that somebody sees. Consciousness is a feature of many mental states but, on this picture, it is not necessary or even central to a state's being a mental state. Consciousness seems central to mentality only because it is so basic to how we know our own mental states. But how we know about things is often an unreliable guide to their nature.

These two alternative pictures of the connection between consciousness and mentality have different implications about what sort of explanation is possible of what it is for a mental state to be a conscious state. If we take the view that consciousness is not a necessary feature of mental states, then we cannot define mental states as conscious states. Accordingly, we must seek some other account of what makes a state a mental state. But once we have an account of mentality that does not appeal to consciousness, we can then try to explain what conscious states are by building upon that very account of mentality. In particular, it then makes sense to try to formulate nontrivial necessary and sufficient conditions for a mental state to be a conscious state. On this conception of mentality and consciousness, it is open for us to proceed sequentially in this way, first defining mentality and then consciousness.

No such procedure, however, is possible if instead we adopt the view that being a mental state is at bottom the same as being a conscious state. For we cannot then explain what makes conscious states conscious by appeal to a prior account of mentality, since on that view mentality presupposes consciousness itself. Any attempt to explain consciousness by formulating necessary and sufficient conditions for a mental state to be conscious will thus automatically fail. If consciousness is already built into mentality, any such explanation will be uninformative. If not, then, on the present view, the conception of mind on which our explanation of consciousness is based will unavoidably be radically defective. It is plain that there is no third way; nothing that is not mental can help to explain consciousness. So, if consciousness is essential to mentality, no informative, nontrivial explanation of consciousness is possible at all. Moreover, since we cannot then proceed sequentially, explaining mentality first and then consciousness, the gulf that seems to separate mind and consciousness from the rest of reality will appear impossible to bridge. Thomas Nagel succinctly expresses this view when he

writes that "[c]onsciousness is what makes the mind-body problem really intractable."[1]

Although it seems effectively to preclude our giving any informative explanation of consciousness, the view that consciousness is essential to all mental states does have apparent advantages. For one thing, that view, which has strong affinities with the Cartesian view of mind, fits well with many of our common-sense intuitions about the mental. And perhaps that view even does greater justice to those intuitions than a view of the mind on which not all mental states are conscious. These two competing pictures of mind and consciousness seem to present us, therefore, with a difficult choice. We can opt to save our presystematic intuitions at the cost of being unable to explain consciousness. Or we can hold open the possibility of giving a satisfactory explanation, but risk being less faithful to our common-sense intuitions about what mental states are.

One reaction to this quandary is simply to accept the more Cartesian of the two pictures, and accept that an illuminating explanation of consciousness will simply prove impossible. This traditional response is now hard to credit. Mind and consciousness are continuous with other natural phenomena of which we can give impressively powerful explanations. And it is difficult to believe that a singularity in nature could exist that would utterly and permanently resist all attempts to explain it. For these reasons, some more recent writers have chosen, instead, simply to abandon common-sense intuitions about mind when they conflict with our explanatory goals. Physics does not aspire to reconstruct all our presystematic intuitions about the things around us. Why, proponents of this eliminativist approach ask, should the science of mind proceed differently?[2]

But we should, wherever possible, seek to explain our common-sense intuitions rather than just explain them away. And we should hesitate to jettison our presystematic conceptions of things, whether mental or physical, unless efforts to do justice to them have decisively failed. Indeed, even physical theories must square as much as possible with our common-sense picture of physical reality. In what follows, I argue that we need embrace neither the Cartesian nor the eliminativist stance toward the consciousness of mental states. Instead, we can both be faithful to our presystematic intuitions about consciousness and mind and, at the same time, construct useful, informative explanations of those phenomena. In section I, I develop

the two pictures sketched above. In particular, I articulate the two different definitions of mentality itself that comprise the core of those two pictures. I then use the non-Cartesian concept of mind and consciousness to construct a systematic and theoretically satisfying explanation of what it is for a mental state to be conscious – an explanation, that is, of what it is that distinguishes conscious from nonconscious mental states. I show further how the definition of mentality central to each of the two pictures determines a distinct conception of consciousness, and how the Cartesian concept of consciousness makes any informative explanation of consciousness impossible. In sections II and III, then, I go on to argue that the non-Cartesian explanation can save the phenomenological appearances and explain the data of consciousness as well as the more familiar Cartesian picture. And I argue there that the standard considerations that favor the Cartesian view are baseless. And in section IV I conclude with some observations about consciousness and our knowledge of the mental, and about the actual significance of the insights that underlie the Cartesian picture.

I

All mental states, of whatever sort, exhibit properties of one of two types: intentional properties and phenomenal, or sensory, properties. Something has an intentional property if it has propositional content, or if it is about something. Sensory properties, by contrast, are less homogeneous. Examples are the redness of a visual sensation and the sharp painful quality of certain bodily sensations. Some mental states may have both intentional and phenomenal properties. But whatever else is true of mental states, it is plain that we would not count a state as a mental state at all unless it had some intentional property or some phenomenal property.

Something close to the converse holds as well. For one thing, only mental states can have phenomenal properties. Although we use words such as 'red' and 'round' to refer to properties of physical objects as well as to properties of mental states, we refer to different properties in the two cases. The introspetible redness of a visual sensation is not the same property as the perceptible redness of a tomato, for example, since each can occur in the absence of the other. Moreoever, mental states are not objects at all, and therefore cannot have the same properties of shape and color that physical physical objects have. Indeed, we do not even use quality words the same way when

we talk about mental states and about physical objects. We speak interchangeably about red sensations and sensations of red, but it makes no nonmetaphorical sense to talk about tomatoes of red. Similar considerations apply to properties that are special to bodily sensations. Knives and aches may both be dull, but the dullness of a knife, unlike that of an ache, has to do with the shape of its edge. Phenomenal properties, properly so called, are unique to mental states.[3]

Things are slightly less straightforward with intentional properties, since items other than metal states can exhibit in intentionality. Speech acts and works of art, for example, can be about things and can have propositional content. But except for mental states themselves, nothing has intentional properties other than those modes and products of behavior which express intentional mental states. So it is reasonable to hold that these modes and products of behavior derive their intentionality from the mental states they express. As Roderick M. Chisholm puts it, "thoughts are a 'source of intentionality' – i.e., nothing would be intentional were it not for the fact that thoughts are intentional."[4] So, even though intentional properties belong to things other than mental states, they do so only derivatively. Accordingly, all mental states have intentional or sensory properties, and sensory properties belong only to mental states and intentional properties nonderivatively to mental states alone. We thus have a compelling basis for defining mental states as just those states which have either intentionality or phenomenal quality.

There are, however, objections to this way of delineating the distinctively mental which seem to favor a mark of the mental based on consciousness instead. For one thing, a mark of the mental that relies solely on intentional and phenomenal properties may seem to underplay the special access we have to our own mental states. Even if all mental states do exhibit intentional or sensory features, one might urge that the more revealing mark of the mental would somehow appeal, instead, to that special access. On such a mark, what is essential to mental states would not be intentional or sensory character, but consciousness itself. Moreover, if we take the possession of either sensory or intentional properties to be definitive of the mental, we must then explain why we regard this disjunctive mark as determining the mental. Why do we construe as a single category the class of states that have one or the other of these two kinds of properties? It does not help to note that that some mental states, for example, perceptual states, have both sorts of characteristic. Despite the existence of such mongrel cases, it seems unlikely that pure

phenomenal states, such as pains, have anything interesting in common with pure intentional states, such as beliefs. And we can avoid this difficulty if, instead, we take consciousness to be what makes a state a mental state. Finally, the characteristically mental differences among kinds of mental states are all differences in what intentional or sensory properties those states have. So those properties may seem to figure more naturally in an account of how we distinguish among types of mental state than in an account of how mental states differ from everything else. These various considerations all suggest that an account of mind in terms of consciousness may be preferable to an account that appeals to intentionality and phenomenal character.

Moreover, defining mentality in terms of consciousness need not involve any circularity. We can say what it is to be a conscious state in a way that does not explicitly mention being mental. A state is conscious if whoever is in it is to some degree aware of being in it in a way that does not rely on inference, as that is ordinarily conceived, or on some sort of sensory input. Conscious states are simply those states to which we have noninferential and nonsensory access.

People do, of course, have many more beliefs and preferences, at any given time, than occur in their stream of consciousness. And the nonconscious beliefs and preferences must always have intentional properties. But this need not be decisive against taking consciousness as our mark of the mental. For we can construe beliefs and preferences as actual mental states only when they are conscious. On other occasions we could regard them to be merely dispositions for actual mental states to occur; in those cases, we can say, one is simply disposed to have occurrent thoughts and desires.

Consciousness is intuitively far more crucial for sensory states than for intentional states. This disparity is something we must explain if we take consciousness as the mark of all mental states. Construing nonconscious beliefs and preferences not as actual mental states but as mere dispositions to be in such states helps us give a suitable explanation. Sensory states normally result from short-term stimulations; so we have little reason to talk about our being disposed to be in particular types of sensory state. By contrast, we are often disposed to be in intentional states of various kinds. Since we are typically not conscious of being thus disposed, the tie between consciousness and mentality may at first sight seem less strong with intentional than with sensory states. But that tie may apply equally to both sorts of state if we count only nondispositional states as mental states, properly

speaking. For when we focus on short-term, episodic intentional states, the common-sense intuition that mental states must be conscious is no less compelling than it is in the case of phenomenal states.

The two marks of the mental just sketched are independent of each other, and both lay claim to long and well-established histories. Thus writers with Cartesian leanings have generally favored some mark based on consciousness, while those in a more naturalist, Aristotelian tradition have tended to rely instead on some such mark as intentionality or sensory character. For it is a roughly Aristotelian idea that the mental is somehow dependent on highly organized forms of life, in something like the way in which life itself emerges in highly organized forms of material existence. And this idea suggests that one should try to delimit the mental in terms of the various distinctively mental kinds of functioning and, thus, by reference to the intentional and phenomenal characteristics of mental states. To the Aristotelian, such a mark has the advantage of inviting one to conceive of the mental as continuous with other natural phenomena. Thus Aristotle's own account of psychological phenomena gives great prominence to sense perception, thereby stressing the continuity between the mental and the biological.

The Cartesian tradition, by contrast, conceives of the mental as one of the two jointly exhaustive categories of existence, standing in stark opposition to everything physical. And on this view it is tempting to select some single essential feature, such as consciousness, to be the mark of the mental. For this kind of mark will stress the sharp contrast between mental and physical, and play down the differences among types of mental states compared to how different all mental states are from everything else. In this spirit, Descartes takes nonperceptual, propositional states to be the paradigm of the mental and, notoriously, has great difficulty in explaining how perception can involve both mental and bodily states.

Although both marks of the mental have enjoyed widespread acceptance, it is crucial which mark we adopt if our goal is to give an explanation of consciousness. Conscious states are simply mental states we are conscious of being in. And, in general, our being conscious of something is just a matter of our having a thought of some sort about it. Accordingly, it is natural to identify a mental state's being conscious with one's having a roughly contemporaneous thought that one is in that mental state. When a mental state is conscious, one's awareness of it is, intuitively, immediate in some way. So we can stipulate that the contemporaneous thought one has is not mediated

by any inference or perceptual input. We are then in a position to advance a useful, informative explanation of what makes conscious states conscious. Since a mental state is conscious if it is accompanied by a suitable higher-order thought, we can explain a mental state's being conscious by hypothesizing that the mental state itself causes that higher-order thought to occur.

At first sight it may seem that counterexamples to this explanation are rife. Although we are usually, when awake, in some conscious mental state or other, we rarely notice having any higher-order thoughts of the sort this explanation postulates. Typically, mental states occur in our stream of consciousness without our also having any evident thought that we are in those states. But such cases are not counterexamples unless we presuppose, contrary to the present explanation, that all mental states are conscious states. For otherwise, there will be no reason to assume that the higher-order thoughts that our explanation posits would, in general, be conscious thoughts. On this explanation, a mental state is conscious if one has a suitable second-order thought. So that second-order thought would itself be a conscious thought only if one also had a third-order thought that one had the second-order thought. And it begs the question against that account to assume that that these higher-order thoughts are usually, or even often, conscious thoughts. If a mental state's being conscious does consist in one's having a suitable higher-order thought, there is no reason to expect that this thought would ordinarily be a conscious thought. Indeed, we would expect, instead, that the third-order thoughts that confer consciousness on such second-order thoughts would be relatively rare; it is hard to hold in mind a thought about a thought that is in turn about a thought. So the present account correctly predicts that we would seldom be aware of our second-order thoughts, and this actually helps confirm the account.

It is important to distinguish a mental state's being conscious from our being introspectively aware of that state. Higher-order thoughts are sometimes invoked to explain introspection, which is a special case of consciousness.[5] But introspection is a more complex phenomenon than the ordinary consciousness of mental states. Intuitively, a mental state's being conscious means just that it occurs in our stream of consciousness. Introspection, by contrast, involves consciously and deliberately paying attention to our contemporaneous mental states. As Ryle remarks "introspection is an attentive operation and one which is only occasionally performed, whereas consciousness is supposed to be a constant element of all mental processes."[6]

Normally when mental states occur in one's stream of consciousness, one is unaware of having any higher-order thought about them. But when we are reflectively or introspectively aware of a mental state, we are aware not only of being in that mental state; we are also aware that we are aware of being in it. The Cartesian picture of mind and consciousness thus tacitly conflates a mental state's being conscious with our being introspectively aware of it. For on that picture the consciousness of a mental state is inseparable from that mental state. So reflective awareness, which is being aware both of a mental state and of one's awareness of that state, will be inseparable from awareness of the state which is not thus reflective. Here our common-sense intuitions diverge from the Cartesian view that consciousness is essential to mental states, since the two kinds of awareness plainly do differ.

Introspection is consciously and deliberately paying attention to mental states that are in our stream of consciousness. So, whatever else one holds about consciousness, it is natural to explain introspection as one's having a conscious higher-order thought that one is in the mental state that one is introspectively aware of. So, if these higher-order thoughts all had to be conscious, we could invoke them only to explain introspective consciousness. For only when we are introspectively aware of a mental state are we also aware of our higher-order thoughts. But higher-order thoughts are not automatically conscious, any more than other mental states are. They are conscious only when we have a yet high-order thought that we have such a thought. So there is no difficulty about using higher-order thoughts to explain not only reflective or introspective awareness, but also what it is for a mental state just to be in our stream of consciousness without our also consciously focusing on it. Introspective awareness of a particular mental state is having a thought that one is in that mental state, and also a thought that one has that thought. Having a conscious mental state without introspectively focusing on it is having the second-order thought without the third-order thought. It may seem slightly odd that each of these hierarchies of conscious mental states has a nonconscious thought at its top. But whatever air of paradox there seems to be here is dispelled by the common-sense truism that we cannot be conscious of everything at once.

One might urge against the present account that higher-order thoughts are unnecessary to explain the consciousness of mental states. Intuitively, a mental state is conscious if it is introspectible. And one might conclude from this

that, to explain such consciousness, we need not posit actual higher-order thoughts, but only dispositions to have such thoughts. A mental state is conscious, on this suggestion, if one is disposed to think that one is in that state.[7] But there are several difficulties with such a dispositional account. For one thing, the consciousness of mental states is phenomenologically something occurrent. Since consciousness does not appear to be dispositional, it is ad hoc simply to posit a disposition that comes and goes as needed. We cannot, of course, save all the phenomenological appearances, but we should prefer to do so when we can. Moreover, it is unclear what explanatory work a disposition to have a higher-order thought would do, except when one actually had that thought, and the disposition would then be superfluous.

In any case, the present account readily enables us to explain the intuition that a state's being conscious means that it is introspectible. To introspect a mental state is to have a conscious thought about that state. So introspection is having a thought about some mental state one is in and, also, a yet higher-order thought that makes the first thought conscious. It is a feature of our experience that, when a mental state is conscious, we can readily come to have a conscious thought about that mental state. On the present account, we do not come to have a new thought about that mental state; we simply come to be conscious of a thought we already had, albeit nonconsciously. Higher-order thoughts are mental states we can become aware of more or less at will. A state's being conscious therefore amounts to its being introspectible. Only if being unaware of a higher-order thought meant that one simply did not have that thought would we have reason to try to make do with dispositions, rather than the actual thoughts themselves.

On the present account, conscious mental states are mental states that cause the occurrence of higher-order thoughts that one is in those mental states. And, since those higher-order thoughts are distinct from the mental states that are conscious, those thoughts can presumably occur even when the mental states that the higher-order thoughts purport to be about do not exist. But such occurrences would not constitute an objection to this account. It is reasonable to suppose that such false higher-order thoughts would be both rare and pathological. Nor would they be undetectable if they did occur. We can determine the presence of nonconscious mental states by way of their causal connections with behavior and stimuli, and with other mental states, both conscious and not. Similarly, we can detect the absence of mental

states by virtue of the causal connections they would have with such other events.

By itself, the present account of consciousness does not imply a materialist or naturalist theory of mind. Indeed, the account is compatible with even a thoroughgoing Cartesian dualism of substances. But it does square nicely with materialist views. For the account holds that what makes conscious mental states conscious is their causing higher-order thoughts that one is in those mental states. And the materialist can reasonably maintain that this causal pattern is due to suitable neural connections.

Moreover, the materialist can argue that intentional and sensory properties are themselves simply special sorts of physical properties. For one thing, arguments that these mental properties are not physical properties usually rely on the unstated, and question-begging assumption that anything mental is automatically nonphysical. Independent support for this supposition is seldom attempted. Even more important, however, the characteristics that are supposed to show that intentional or sensory properties are not physical turn out, on scrutiny, to be characteristics that various indisputably physical properties also exhibit.[8] So even if no developed, satisfactory account of these properties is presently at hand, there is no reason to doubt that accurate accounts will be forthcoming that are compatible with a thoroughgoing naturalist view of mind. Together with the present explanation of the consciousness of mental states, this should make possible a reasonably comprehensive naturalist theory of mind.

It is a welcome benefit of the present account that it is hospitable to naturalist theories, but this is not its main strength. Rather, its principal advantage is just that it enables us to explain what it is for a mental state to be a conscious state. The present explanation, moreover, has precise empirical consequences that one could reasonably hope to test. For it implies not only that conscious mental states are accompanied by distinct higher-order thoughts, but also that some causal mechanism exists that connects conscious mental states to the corresponding higher-order thoughts.

Such an explanation is possible only if we adopt the non-Cartesian view that intentional and sensory character are jointly the mark of the mental. If, instead, we were to follow the Cartesian tradition in regarding consciousness itself as the key to mentality, no account of consciousness in terms of higher-order thoughts could succeed. For then one would have to deny that

a mental state could occur without its being conscious. As Descartes put it, "no thought can exist in us of which we are not conscious at the very moment it exists in us."[9] But, if all mental states are conscious, and a higher-order thought exists for every conscious mental state, serious, insurmountable difficulties immediately ensue. For one thing, there would be denumerably many distinct higher-order thoughts corresponding to every conscious mental state. No mental state could be conscious without being accompanied by a higher-order thought. But that thought would itself have to be conscious, and so a yet higher-order thought would be necessary. This regress would never halt. It strains credulity to suppose that human beings can have infinitely many conscious thoughts at a particular time. And even if we could, it is hardly sensible to explain a mental state's being conscious by way of such an infinite series.

Even more damaging consequences follow for an account in terms of higher-order thoughts if all mental states are conscious states. As noted above, we are not normally aware of the higher-order thoughts that, on such an account, make mental states conscious. But, if all mental states were conscious, we would be aware of any higher-order thoughts that we have. So we could not explain why we typically seem not to have such thoughts by saying that they are simply not conscious thoughts. By requiring that all mental states be conscious states, the Cartesian conception of mentality rules out our explaining consciousness by reference to higher-order thoughts.

If consciousness were what makes a state a mental state, therefore, any account that represents that consciousness as being due to a connection that conscious mental states have with some other mental state would be radically misguided. For that other mental state would then itself have to be conscious, and we would have to invoke yet another mental state to explain its being conscious. A vicious regress would thus be unavoidable. So long as we hold that all mental states are conscious, we can prevent that regress only by maintaining that the consciousness of a mental state is not a relation that state bears to some other mental state, but rather an intrinsic property. Moreover, if consciousness is what makes mental states mental, it will be viciously circular to explain that consciousness in terms of a relation that conscious mental states bear to other mental states. An explanation in terms of other mental states would appeal to states we know to be conscious. It is plain that we cannot explain or analyze consciousness at all unless we can do so in terms of some sort of mental phenomenon. So, if consciousness is what

makes a state a mental state, consciousness will not only be an intrinsic, nonrelational property of all mental states; it will be unanalyzable as well. It will, as Russell disparagingly put it, be "a pervading quality of psychical phenomena."[10] Indeed, if being mental means being conscious, we can invoke no mental phenomenon whatever to explain what it is for a state to be a conscious state. Since no nonmental phenomenon can help, it seems plain that, on the Cartesian concept of mentality, no informative explanation is possible of what it is for a mental state to be conscious.

Since consciousness is a matter of our noninferential and nonsensory knowledge of our mental states, it is tempting to describe the issue in terms of such notions as incorrigibility, infallibility, and privacy.[11] But the foregoing obstacles to explaining consciousness do not derive from any such epistemic matters. Rather, they result simply from the Cartesian idea that all mental states are conscious states.

On the Cartesian concept of mentality and consciousness, consciousness is essential to mental states. It is therefore a nonrelational property of those states that is very likely unanalyzable as well. That this conception prevents us from explaining consciousness in any useful way is the most compelling reason we can have for adopting, instead, a non-Cartesian mark of the mental. But there are other reasons as well to prefer a non-Cartesian mark. For one thing, it is impossible to conceive of a mental state, whether or not it is conscious, that lacks both intentional and sensory properties. So, even though it may not always be easy to imagine one's being in a mental state that is not conscious, intentional and sensory properties are evidently more central to our concept of a mental state than consciousness is. So, even though the characteristically mental differences among mental states are, as noted earlier, a function of their intentional and sensory properties, those properties are not only important for explaining how we distinguish among the various types of mental state. They are also necessary for explaining how mental states differ from everything nonmental.

The Cartesian might concede that we can have no notion of a mental state that has neither intentional nor phenomenal character, but go on to insist that we also can have no idea of what it would be like to be in a nonconscious mental state even if it does have intentional or sensory properties. But knowing what it would be like to be in such a state is not relevant here. Knowing what it is like to be in a state is knowing what it is like to be aware of being in that state. So, if the state in question is not a conscious mental

state, there will be no such thing as what it is like to be in it, at least in the relevant sense of that idiom. This does not show, however, that intentional and phenomenal states cannot lack consciousness. Conscious states resemble and differ in respect of their intentional or phenomenal features. Accordingly, nonconscious mental states will simply be states that resemble and differ from one another in exactly these ways, but without one's being noninferentially aware of their existence and character.

Indeed, it is indisputable that inner states that resemble and differ in just these ways do occur outside our stream of consciousness. Many sorts of mental state, such as beliefs, desires, hopes, expectations, aspirations, various emotions, and arguably even some bodily feelings such as aches, often occur in us without our noticing their presence. And the only thing that makes these states the kinds of states they are is the intentional and phenomenal properties they have. So we must explain what it is for these states to be mental not by reference to consciousness, but by appeal to their having phenomenal or intentional character. As noted above, we can deny that some of these mental phenomena are properly speaking mental states at all, and instead construe them as mere dispositions for mental states to occur. But states of these sorts often have a strong effect on our actual behavior, and even influence the course and content of our stream of consciousness. These mental phenomena must presumably be nondispositional states at least on those occasions when they exercise such causal influence. So the only reason to regard them as mere dispositions would be a question-begging concern to sustain the theory that all mental states are conscious states.

Perhaps the Cartesian will counter that, even if nonconscious intentional states are unproblematic, the idea that a mental state could have sensory character and yet not be conscious is simply unintelligible. For it may seem that the very idea of a nonconscious state with sensory qualities is, in effect, a contradiction in terms. What seems to make intelligible the idea of a mental state's having phenomenal qualities at all is our immediate awareness of how such states feel, or what they are like for those who are in them. This issue will receive extended consideration in section III. For now, however, it is enough to note that, even if we understand what it is for a state to have sensory quality only because we are familiar with cases in which we are conscious of being in such states, it hardly follows that nonconscious sensory states cannot occur. That we understand a kind of phenomenon by way of a particular kind of case does not show that cases of other sorts are impossible.

II

On the Cartesian view, consciousness is definitive of the mental. This concept of mentality implies that consciousness cannot be a relational characteristic of mental states, and that it may well be inexplicable as well. The difficulty in explaining consciousness on that view actually results from the Cartesian strategy for dealing with mental phenomena. The main strength of the Cartesian picture is that it closely matches our presystematic, common-sense intuitions. But it achieves this close match by building those intuitions into our very concepts of mind and consciousness. And this automatically trivializes any explanation we might then give of them. We cannot very well give non-question-begging accounts of intuitions that we incorporate definitionally into our very concepts. Explanations based on the Cartesian conceptions of mind and consciousness thus rely heavily, and ineliminably, on interdefinition of such terms as 'mind', 'consciousness', 'subjectivity', and 'self'. Such interdefinition may be useful in marking out a range of interconnected phenomena, but it cannot do much to help explain the phenomena thus delineated.

On the non-Cartesian concept, by contrast, consciousness is not essential to mental states, and thus consciousness may well be an extrinsic characteristic of whatever mental states have it. The Cartesian achieves its close match with common sense at the cost of ruling out any useful explanation. No such trade-off is necessary on the non-Cartesian picture. The non-Cartesian has no trouble in giving a theoretically satisfying explanation of consciousness. And it is possible to show that this account enables us to save the phenomenological appearances at least roughly as well as the Cartesian can. Moreover, objections to an account cast in terms of higher-order thoughts can be convincingly met. In this section and the next I consider some of the most pressing of these objections, and also argue that such an account does do justice to the phenomenological data. In the present section I take up various general questions about the adequacy of the non-Cartesian account; in section III I address issues that pertain specifically to sensory qualities and to subjectivity.

One especially notable feature of our presystematic view of consciousness which the Cartesian conception seems to capture perspicuously is the close connection between being in a conscious state and being conscious of oneself. An account in terms of higher-order thoughts has no trouble here. If a mental state's being conscious consists of having a higher-order thought that one is in that mental state, being in a conscious state will imply having a thought about

oneself. But being conscious of oneself is simply having a higher-order thought about oneself. So being in a conscious mental state is automatically sufficient for one to be conscious of oneself.

Any reasonable account of consciousness will presumably insist on this connection. But the Cartesian can say little that is informative about why the connection should hold. An account that appeals to higher-order thoughts has no such difficulty. Moreover, there is a well-motivated reason why the higher-order thought that the non-Cartesian invokes must be a thought about oneself. To confer consciousness of a particular mental state, the higher-order thought must be about that very mental state. And the only way for a thought to be about a particular mental state is for it to be about somebody's being in that state. Otherwise, the thought would just be about that type of mental state, and not about the particular token of it. So, in the case at hand, the higher-order thought must be a thought that one is, oneself, in that mental state.[12]

Having a thought that one is, oneself, in a particular mental state does not by itself presuppose any prior conception of the self, or of some sort of unity of consciousness. Rather, the present view allows us to explain these conceptions as themselves actually arising from our being in conscious mental states. For we can construe the second-order thoughts as each being a thought to the effect that whatever individual has this very thought is also in the target mental state. And, if a fair number of these thoughts are conscious thoughts, it is plausible to suppose that a sense of the unity of consciousness will, in time, emerge.

If one held the Cartesian view that all mental states are conscious, invoking higher-order thoughts would issue in the vicious regress noted in secion I. So, if one is tempted by both these moves, one might try to adjust things in order to avoid that outcome. The most promising way to do so would be simply to insist that the higher-order thoughts in virtue of which we are conscious of conscious mental states are actually part of those conscious states themselves. Every conscious mental state would then be, in part, about itself, and our knowledge that we are in such states would be due to that self-reference. Metaphorically, we would then conclude that a mental state's knowing itself is, in Ryle's apt metaphors, a matter of its being "self-intimating" (158) or "self-luminous" (159).[13]

This line of reasoning is particularly inviting, since it suggests that the Cartesian can, after all, give some nontrivial explanation of the consciousness

of mental states. Conscious mental states are conscious, on this account, because they are about themselves. And this self-reference is intrinsic; it does not result from some connection those states have with other mental states. But anything that would support the view that conscious mental states are conscious because they know, or are in part about, themselves would provide equally good evidence that consciousness is due to an accompanying higher-order thought. Moreover, we have no nonarbitrary way to tell when one mental state is a part of another. Accordingly, there is no reason to uphold the idea that our awareness of conscious states is a part of those states other than a desire to sustain the Cartesian contention that all mental states are conscious states. Moreover, if conscious states have parts in this way, the question arises whether all the parts of such states must be conscious, or only some. If all, then the awareness of the mental state will have to be conscious. A regress would thus arise that is exactly parallel to that which arose when we construed the awareness of conscious mental states as due to a distinct higher-order thought. The only advantage of an account on which that awareness is a part of the conscious mental state is if the awareness is a nonconscious part of the conscious state. This reinforces the conclusion that there is no nonarbitrary way to distinguish this view from an account in terms of higher-order thoughts. And it undercuts the idea that the Cartesian can formulate an informative explanation of consciousness along these lines. Since the Cartesian explanation would work only if the part of each conscious state that makes it conscious were itself conscious, the regress is unvoidable.

One reason that consciousness seems intrinsic to our sensory states is that it is difficult to isolate that consciousness as a distinct component of our mental experience. When we try to focus on the consciousness of a particular sensory state, we typically end up picking out only the sensory state we are conscious of, instead. As Moore usefully put it, consciousness is "transparent," or "diaphanous."[14] Since efforts to pick out consciousness itself issue instead in the states we are conscious of, it is tempting to conclude that the consciousness is actually part of those states. But the present account gives a better explanation of the diaphanous character of consciousness. We normally focus on the sensory state and not on our consciousness of it only because that consciousness consists in our having a higher-order thought, and that thought is usually not itself a conscious thought.

There is a strong intuitive sense that the consciousness of mental states is somehow reflexive, or self-referential. But we need not invoke the idea that

conscious states are conscious of themselves to explain this intuition. For a mental state to be conscious, the corresponding higher-order thought must be a thought about oneself, that is, a thought about the mental being that is in that conscious state. So, as noted above, we can construe that thought as being, in part, about itself. For it is reasonable to regard the content of the thought as being that whatever individual has this very thought is also in the specified mental state. The sense that something is reflexive about the consciousness of mental states is thus not due to the conscious state's being directed upon itself, as is often supposed. Rather, it is the higher-order thought that confers such consciousness that is actually self-directed.

The foregoing objections have all challenged whether an account based on higher-order thoughts can do justice to various ways we think about consciousness. But one might also question whether higher-order thoughts are enough to make mental states conscious. Here a difficulty seems to arise about mental states that are repressed. By hypothesis such states are not conscious. But it might seem that mental states can be repressed even if one has higher-order thoughts about them. Higher-order thoughts could not then be what makes mental states conscious. A person who has a repressed feeling may nonetheless take pleasure, albeit unconscious pleasure, from having that repressed feeling. But to take pleasure in something we must presumably think that it is so. So that person will have a higher-order thought about the repressed feeling.[15] Moreover, it appears intuitively that the feeling cannot remain unconscious unless the pleasure taken in it also does. And this suggests that, contrary to the present account, a higher-order thought can confer consciousness only if that thought is itself already conscious.

But genuine counterexamples along these lines are not all that easy to come up with. Despite the foregoing suggestion, one can take pleasure in something without having any actual thought about it. I cannot, of course, take pleasure in something I disbelieve or doubt,[16] but that does not imply the actual occurrence of a thought that it is so. Indeed, we frequently form no actual thought about the things in which we take pleasure. Sometimes, by 'thought', we mean only to speak of propositional contents, as when I talk about some thought you put forth. Taking pleasure is a propositional mental state; so taking pleasure in something does involve a thought about it, in the sense of a proposition. But it hardly follows that one also has a thought, in the sense of a particular kind of mental state. Having a thought in that sense is the holding of an assertive mental attitude, which need not occur when one

takes pleasure in something. So, in the foregoing example, we have no reason to suppose that the person would actually have any higher-order thought about the repressed feeling.

The difference between taking pleasure in a mental state and having an actual thought about it is crucial for the present account of consciousness. It is natural to hold that being aware of something means having a thought about it, not taking pleasure in it. And one can take pleasure in something without knowing what it is that gives one pleasure. One may have no idea why one feels good, or be mistaken about why. One may even be unaware of feeling good at all if one is sufficiently distracted or other factors interfere. So taking pleasure in something is compatible with being unaware of that thing. Such considerations also apply to putative counterexamples based on other sorts of higher-order mental states. For example, repressed feelings are, presumably, always accompanied by higher-order desires not to be in them. But desires that something not be so do not, in general, imply any awareness that it is.

Conceiving of nonconscious mental states on the model of the repressed cases is doubly misleading. For one thing, it ignores factors that in such cases presumably block consciousness. Moreover, it conceals a tacit Cartesian premise. For it suggests that consciousness is the norm: unless exceptional pressures intervene, a mental state will automatically be conscious. Consciousness is to be presupposed unless some external factor prevents it. Thus, on this model, we can explain the forces that interfere with consciousness, but consciousness itself may very likely be inexplicable.[17]

III

Whatever one holds about intentional states, it may seem altogether unacceptable to try to explain the consciousness of sensory states by way of higher-order thoughts. Consciousness seems virtually inseparable from sensory qualities, in a way that does not seem so for intentional properties. Indeed, as noted at the end of section I, it may seem almost contradictory to speak of sensory states' lacking consciousness. This intimate tie between sensory quality and consciousness seems to hold for all sensory states, but appears strongest with somatic sensations, such as pain. Saul A. Kripke succinctly captures this intuition when he insists that "[f]or a sensation to be *felt* as pain is for it to *be* pain"[18] and, conversely, that "for [something] to exist

without being *felt as pain* is for it to exist without there *being any* pain."[19] And, more generally, Kripke seems to insist that for something to be a sensation of any sort it must be felt in a particular way (*NN* p. 146).

Since consciousness seems more closely tied to sensations than to intentional states, it is tempting to consider a restricted from of the Cartesian view, on which all sensations are conscious but not all intentional states are.[20] This restricted thesis would still allow one to explain consciousness in terms of higher-order thoughts; no regress would arise, because then those thoughts could themselves be nonconscious. But the Cartesian view holds not only that all mental states are conscious, but also that consciousness is an intrinsic property of mental states. And if it is, an explanation in terms of higher-order thoughts is impossible, and all the problems about giving an informative explanation of consciousness will arise. So even if not all mental states are conscious, it is important to see whether consciousness is intrinsic to those which are.

We can, however, explain our tendency to associate consciousness and sensory qualities without having to suppose that consciousness is intrinsic to sensory states, or even that all sensory states are conscious. We are chiefly concerned to know what bodily sensations we and others have because they are highly useful indicators of bodily and general well being. People cannot tell us about their nonconscious sensations, and bodily sensations usually have negligible effect on behavior unless they are conscious. So nonconscious sensations are not much use as cues to such well being, and we thus have little, if any, interest in pains or other somatic sensations, except when they are conscious.

Things are different with other sorts of mental states, even perceptual sensations. It is often useful to know somebody's thoughts, emotions, and perceptual sensations, even when that person is unaware of them. Moreover, when mental states are not conscious, our interest in knowing about them is greatest with propositional states, less with emotions, less still with perceptual sensations, and far the least with somatic sensations. Strikingly, our sense that consciousness is intrinsic to mental states increases accordingly. The less useful it is to know about a particular kind of mental state even when the person is unaware of it, the more compelling is our intuition that that kind of mental state must be conscious. This correlation is telling evidence that, even in the case of pains and other somatic sensations, the idea that being mental entails being conscious is just a reflection of our usual interests, and not a

matter of the meanings of our words or of the nature of the mental itself.

Some of our idiomatic ways of describing somatic sensations do entail consciousness. Something's hurting, for example, implies awareness of the hurt. And perhaps one cannot correctly say that somebody is in pain unless that person knows it. Phrases such as 'what a sensation is like' and 'how a sensation feels' reinforce this impression, since they refer both to a sensory quality and to our awareness of it, and seem thus to yoke the two together. But when one is in pain or when something hurts, we not only are in a sensory state, but are also aware that we are. And our idiomatic descriptions of these situations have no bearing on whether that very kind of sensory state may sometimes occur without one's being aware of it. Perhaps we would then withhold from such states the epithet 'pain'. But those states would still resemble and differ from other nonconscious states in just those ways in which conscious pains resemble and differ from other conscious sensory states. And that is what it is for a state to have sensory qualities. The intuitive simplicity of those qualities might tempt one to hold that consciousness also is simple and, hence, an intrinsic characteristic of sensory states. But it is question begging to suppose that the apparent simplicity of sensory qualities tells us anything about the nature of our consciousness of them.

Examples of sensory states that sometimes occur without consciousness are not hard to come by. When a headache lasts several hours, one is seldom aware of it for that entire time. Distractions occur, and one pays attention to other things, or just forgets for a bit. But we do not conclude that each headache literally ceases to exist when it temporarily stops being part of our stream of consciousness, and that such a person has only a sequence of discontinuous, brief headaches. Rather, when that happens, our headache is literally a nonconscious ache.[21] The same holds even more vividly for mild pains and minor bodily discomforts. So, to insist that nonconscious states are just not mental states, or that they cannot have sensory qualities, is not, as Kripke seems to urge (e.g., *NN* 152–3), the elucidation of decisive and defensible presystematic intuitions, but only the tacit expression of the Cartesian definition of mind.

Indeed, an account in terms of higher-order thoughts actually helps explain the phenomenological appearances. If a sensory state's being conscious is its being accompanied by a suitable higher-order thought, that thought will be about the very quality we are conscious of. It will be a thought that one is in a state that has that quality. So it will indeed be impossible to describe

that consciousness without mentioning the quality. An account in terms of higher-order thoughts actually helps explain why the qualities of our conscious experiences seem inseparable from our consciousness of them.

Moreover, we typically come to make more fine-grained discriminations as we master more subtle concepts pertaining to various distinct sensory qualities. Experiences from wine tasting to hearing music illustrate this process vividly. An account in term of higher-order thoughts explain the bearing these concepts have on our very awareness of sensory differences. If consciousness is intrinsic to sensory states, the relevance of concepts remains mysterious. The Cartesian might just deny that sensory differences exist when we are unaware of them. But it will be even more difficult to explain how learning new concepts can actually cause sensory qualities to arise that previously did not exist.

Perhaps the strongest objection to an account in terms of higher-order thoughts is that there are creatures with conscious sensations whose ability to have thoughts at all may be in doubt. Infants and most nonhuman animals presumably have a relatively rudimentary ability to think, but plainly do have conscious sensations. But one need not have much ability to think to be able to have a thought that one is in a particular sensation. Infants and nonhuman animals can discriminate among external objects, and master regularities pertaining to them. So most of these beings can presumably form thoughts about such objects, albeit primitive thoughts that are very likely not conscious. No more is needed to have thoughts about one's more salient sensory experiences. Infants and nonhuman animals doubtless lack the concepts required for drawing many distinctions among their sensory states. But, as just noted, one can be aware of sensory states and yet unaware of many of the sensory qualities in virtue of which those states differ.

The common tendency to link the ability to think with the ability to express thoughts in speech may account for the doubt we can fall into about whether infants and nonhuman animals can think at all. But the capacity for speech is hardly necessary for thinking. It is often reasonable to interpret nonlinguistic behavior, of other people and of non-language-using creatures alike, in terms of the propositional content and mental attitude we take it to express. Such behavior is convincing evidence of the occurrence of intentional states.

Forming higher-order thoughts about one's own propositional mental states takes a lot more than having such thoughts about one's sensations. For

one thing, the concept of a mental state with propositional content is more complex than the concept of a sensory experience. And picking out particular mental states demands an elaborate system of concepts, whereas referring to salient sensory experiences does not. An account in terms of higher-order thoughts fits well with these points. Infants and most nonhuman species lack the ability to have the more complex higher-order thoughts needed to make intentional states conscious, though they presumably can form higher-order thoughts about their sensory states. And, though these beings plainly have conscious sensations, we have little reason to suppose that their intentional states are also conscious. Indeed, these considerations help explain why we associate consciousness so much more strongly with sensory than with intentional states. Conscious of sensory states arises far more readily, since higher-order thoughts about them are far easier to have.

Some animal species, however, lack the ability to think at all. And this may seem to support Nagel's contention that conscious "experience is present in animals lacking language and thought" (167, n. 3). But being a conscious creature does not entail being in conscious mental states. For an organism to be conscious means only that it is awake, and mentally responsive to sensory stimuli (cf. Ryle, pp. 156–7). To be mentally responsive does require that one be in mental states. And to be mentally responsive to sensory stimuli may even mean that one is in some way conscious of the objects or events that are providing such stimulation. But a creature can be in mental states without being in conscious mental states, and can be conscious of external or bodily events without also being aware of its own mental states.

Conscious experiences, as Nagel has stressed, manifest a certain subjectivity. We each experience our sensory states in a way nobody else does, and from a point of view nobody else shares. It is notoriously difficult to articulate these differences. But we understand their occurrence reasonably well, and it is far from clear that such subjectivity causes any problem for the present account.

One way differences arise in sensory experiences is from variations in sense organs, or other aspects of physical makeup. Experiences also vary from individual to individual because of such factors as background and previous experience. When these factors diverge markedly, aspects of our sensory experiences may as well. When the individuals belong to distinct species, this effect may be quite dramatic. But hard as it is to pin down precisely what these differences amount to, they do not bear specifically

on the consciousness of the experiences in question. Rather, the variations are due to differences in the mental context in which the experiences occur, or, when biological endowment is at issue, they are actual differences in the sensory qualities of those experiences. Nagel holds that "the subjective character of experience" is a matter of what "it is like to *be* a particular organism – [what] it is like *for* that organism" (p. 166). But the present account can accommodate this idea. What it is like to be a particular conscious individual is a matter of the sensory qualities of that individual's conscious experiences, and the mental context in which those experiences occur. The consciousness of those experiences, by contrast, is simply that individual's being aware of having the experiences.[22]

According to Nagel, "[a]ny reductionist program has to be based on an analysis of what is to be reduced. If the analysis leaves something out, the problem will be falsely posed" (p. 167). Indeed, no account that is even "logically compatible with" the absence of consciousness could, Nagel contends, be correct (p. 166;cf. "Panpsychism," p. 189). And the present account is reductionist, since it seeks to explain conscious mental states ultimately in terms of mental states that are not conscious. But that account aims only at explaining consciousness, and not also at conceptual analysis. And satisfactory explanations do not, *pace* Nagel, require full analyses of the relevant concepts. Explanation, in science and everyday context alike, must generally proceed without benefit of complete conceptual analyses.

Nagel's language is strongly evocative of that sense we have of ourselves which can make it appear difficult to see how, as conscious selves, we could find ourselves located among the physical furniture of the universe. When we focus on ourselves in this way, there seems to be nothing more basic to our nature than consciousness itself. If nothing were more basic to us than consciousness, there would be nothing more basic in terms of which we could explain consciousness. All we could do then is try to make consciousness more comprehensible by eliciting a sense of the phenomenon in a variety of different ways. Analyzing concepts would be central to any such project, and Nagel's demand for conceptual analysis would then make sense. But consciousness could be essential to our nature only if all mental states are conscious states. If a fair number of our mental states are not conscious, we cannot define our mental natures in terms of consciousness, and there will be nonconscious mental phenomena in terms of which we can explain consciousness itself.

The puzzled cognitive disorientation that can result from reflecting on the gulf that seems to separate physical reality from consciousness makes any noncircular explanation of consciousness seem inadequate. How could any explanation of consciousness in terms of nonconscious phenomena help us to understand how consciousness can exist in the physical universe, or how physical beings like ourselves can have conscious states? But no other explanation can do better with these quandaries so long as an unbridgeable gulf seems to divide the conscious from the merely physical. To understand how consciousness can occur in physical things, we must dissolve the intuitive force of that gulf. And we can do so only by explaining the consciousness of mental states in terms of mental states that are not conscious. For the stark discontinuity between conscious mental states and physical reality does not also arise when we consider only nonconscious mental states. And once we have explained consciousness by reference to nonconscious mental states, we may well be able also to explain nonconscious mental states in terms of phenomena that are not mental at all.

IV

The central place consciousness has in our conception of the mental is doubtless due in large measure to the way we know about mind in general, and in particular about our own mental states. We get most of that knowledge, directly or indirectly, from introspection. And we have introspective access to mental states only when they are conscious. Since our chief source of knowledge about the mind tells us only about conscious mental states, it is natural to infer that consciousness is an important feature of mental phenomena.

But stronger claims are sometimes made about the epistemic status of introspection. Introspection may seem particularly well adapted to its subject matter, since most of our knowledge of mind derives from introspection, and all introspective knowledge is about mind. This close fit may tempt some to hold that introspection is a privileged source of knowledge that is somehow immune from error. If so, perhaps introspection reveals the essential nature of those states. And, since introspection tells us only about conscious mental states, perhaps consciousness is itself a part of that essential nature. But inviting as these Cartesian conclusions may be, they are without foundation. Introspection is simply the having of conscious thoughts that one is in partic-

ular mental states. Those thoughts can by themselves no more reveal the essences of those states than having a conscious perceptual thought that a table is in front of one can reveal the essence of the table. Nor can we infer anything from the close fit between introspection and its subject matter. Sight is an equally well adapted to knowing about colored physical objects. But there are other ways to know about those objects. And even though sight informs us only about illuminated objects, we can hardly conclude that only illuminated objects are colored.

Introspective apprehension seems to differ, however, from perceptual knowledge in a way that undermines this analogy. Perception is never entirely direct. Some causal process always mediates, even in ostensibly direct perception, between our perceptual experience and what we perceive. Introspection, by contrast, may seem wholly unmediated. And if it is, there would be no way for error or distortion to enter the introspective process. There would thus be no difference between how our mental states appear to us and how they really are. Mental states would have no nonintrospectible nature, and introspection would be an infallible and exhaustive source of knowledge about the mind. Nagel evidently endorses this view when he claims that "[t]he idea of moving from appearance to reality seems to make no sense" in the case of conscious experiences (174). Kripke too seems to hold that introspection is different from perception in this way. Thus he writes:

> although we can say that we pick out [physical] heat contingently by the contingent property that it affects us in such and such a way, we cannot similarly say that we pick out pain contingently by the fact that it affects us in such and such a way ("IN" p. 161; cf. *NN* pp. 150–2).

We could not, Kripke contends, have been aware of our pains in a way different from the way we actually are.

Introspection is the reflective awareness of our mental states. So, the only way introspective apprehension of those states might be entirely unmediated would be for consciousness to be a part, or at least an intrinsic property, of such states. For nothing could then come between a mental state and our being conscious of it, nor between our being thus conscious and our also having reflective consciousness. But as noted in section II, that view is indefensible. Accordingly, consciousness must be a relational property, for example, the property of being accompanied by higher-order thoughts. And some causal process must therefore mediate between mental states and our

awareness of them – in Kripke's example, between a pain and "the fact that it affects us in such and such a way." And, since mental states might have been connected causally to different high-order thoughts, we might have been aware of mental states differently from the way we are. The appearance of mental states will not, therefore, automatically coincide with their reality. Indeed, since how mental states appear is a matter of our introspective awareness of them, their appearance and reality could be the same only if our consciousness of mental states were a part or an intrinsic property of those states.

These considerations notwithstanding, we do rely heavily on introspection in picking out and describing mental states. And introspection tells us about nothing except conscious mental states. So even though consciousness is not what distinguishes mental states from everything else, it is reasonable to hold that it is by reference to a range of conscious states that we fix the extension of the term 'mental'. Similarly, even though the various kinds of mental state can all occur nonconsciously, it is also reasonable to suppose that we fix the extensions of our terms for different kinds of mental state by way of the conscious cases. As Kripke and Hilary Putnam have stressed, what fixes the extension of a general term can turn out to be distinct from what is essential to the items in that extension.[23] And, just as the way we know about things is not, in general, a reliable guide to their nature, so the way we pick out things is not, either. So, even if we fix the extensions of terms for mental states by way of the conscious instances, we could still discover that the states so determined are not all conscious, and that what is actually essential to all such states is just their sensory or intentional properties.

The idea that we fix mental extensions by way of the conscious cases plainly supports the non-Cartesian picture. But it also helps explain why consciousness seems so crucial to our mental concepts. And it even enables us to explain why we group sensory and intentional states together as mental states, despite its seeming that they have little intrinsic in common. We do so because in both cases we fix extensions by way of states to which we have noninferential and nonobservational access.

Kripke contends that what fixes the reference of terms for sensory states cannot diverge from what is essential to those states (*NN* pp. 149–54; "IN" pp. 157–61). Thus, he insists, "[i]f any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon *is* pain."[24] But what fixes the extension of 'pain' must coicide with what is essential to

pains only if it is necessary that pains affect us in the way they do. And this would be necessary only if consciousness were intrinsic to them. It therefore begs the question to base the Cartesian picture on an insistence that what fixes the extensions of mental terms cannot diverge from the essences of mental states. Kripke offers no independent support for that insistence.

Relative to what we now know about other natural phenomena, we still have strikingly scant understanding of the nature of the mental. So introspection looms large as a source of information, just as sense perception was a more central source of knowledge about physical reality before the flourishing of the relevant systematic sciences. But, since not all knowledge about mind is derived from introspection, we have no more reason to suppose that mental states have no nonintrospectible nature than that the nature of physical objects is wholly perceptible. Nor, therefore, have we any reason to hold that the essences of mental states must be what fixes the extensions of mental terms. It is reasonable to conclude that whatever temptation we have to accord absolute epistemic authority to introspection derives solely from our relative ignorance about the mind. Only because we now know so little about mental processes does it make sense to suppose that, in the case of mental states, appearance and reality coincide. Accordingly, we have no reason to continue to favor that picture, or to reject an explanation of consciousness based on higher-order thoughts.[25]

NOTES

[1] 'What is it like to be a bat?', The Philosophical Review, LXXXIII, 4 (October 1974): 435–50; reprinted in Nagel's Mortal Questions (Cambridge and New York: Cambridge University Press, 1979); 165–180, p. 165. Page references to Nagel will be to Mortal Questions and, unless otherwise indicated, to that article.

[2] Richard Rorty and Paul M. Churchland, among others, have championed this view; see Rorty's Philosophy and the Mirror of Nature (Princeton: Princeton University Press, 1979), Part I, and Churchland's Scientific Realism and the Plasticity of Mind (Cambridge: Cambridge University Press, 1979).

[3] In a review of Frank Jackson's Perception [The Journal of Philosophy, LXXXII, 1 (January 1985): 28–41] I argue in detail that words for such qualities have this double use. On this point see also G. E. Moore, 'A reply to my critics', in The Philosophy of G. E. Moore ed. Paul Arthur Schilpp (LaSalle, Illinois: Open Court, 1942): 535–677, pp. 655–8; and Thomas Reid, Essays on the Intellectual Powers of Man, ed. Baruch A. Brody (Cambridge, MA: The M.I.T. Press, 1969), II, xvi, p. 244.

[4] 'Chisholm-Sellars correspondence on intentionality', in Minnesota Studies in the Philosophy of Science, II, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell (Minneapolis: University of Minnesota Press, 1958): 521–39, p. 533. In 'Intentionality,'

Midwest Studies in Philosophy, X (1985), I argue that this claim is defensible if it is construed in strictly causal terms.

[5] See, e.g., David M. Armstrong, A Materialist Theory of the Mind (New York: Humanities Press, 1968), pp. 92–115 and 323–7 and 'What is consciousness?', in Armstrong's The Nature of Mind and Other Essays (Ithaca, NY: Cornell University Press, 1980): 55–67, pp. 59–61; David Lewis, 'An argument for the identity theory,' The Journal of Philosophy, LXIII, 1 (January 6, 1969): 17–25, p. 21 and "Psychophysical and Theoretical Identifications," Australasian Journal of Philosophy, 50, 3 (December 1972): 249–58, p. 258; and Wilfrid Sellars, "Empiricism and the Philosophy of Mind," in Sellars' Science, Perception and Reality (London: Routledge & Kegan Paul and New York: Humanities Press, 1963): 127–96, pp. 188–9 and 194–5.

[6] Gilbert Ryle, The Concept of Mind (London: Hutchinson and Company, 1949), p. 164.

[7] Allen Hazen has urged this line especially forcefully, in correspondence. Also, see Kant's claim that the representation 'I think' must be able to accompany all other representations (K.d.R.V., B131–2; cf. B406, though Kant insists that the representation 'I think' is a nonempirical (B132) or transcendental (B401, A343) representation.

[8] For a detailed argument to this effect, see my 'Mentality and neutrality,' The Journal of Philosophy, LXXIII, 13 (July 15, 1976): 386–415, sec. I.

[9] Fourth Replies, Oeuvres de Descartes, ed. Charles Adam and Paul Tannery (Paris: J. Vrin, 1964–75) [henceforth "AT"] VII, 246 [see The Philosophical Works of Descartes, ed. Elizabeth S. Haldane and G. R. T. Ross (Cambridge: Cambridge University Press, 1931) (henceforth "HR") II, 115]. See also the Geometrical Exposition of the Second Replies: "the word 'thought' applies to all that exists in us in such a way that we are immediately conscious of it" [AT VII, 160 (see HR II, 52)], and elsewhere, e.g.: First Replies, AT VII, 107 (HR II, 13), Letter to Mersenne, AT III, 273 [see Descartes: Philosophical Letters, ed. Anthony Kenny (Oxford: Oxford University Press, 1970), p. 90], Fourth Replies, AT VII, 232 (HR II, 105) and AT VII, 246 (HR II, 115), and Principles I, ix, AT VIII–1, 7 (HR I, 222).

[10] Bertrand Russell, The Analysis of Mind (London: George Allen & Unwin Ltd. and New York: Humanities Press, 1921), p. 9.

[11] See, e.g., Rorty's argument "that incorrigibility is the best candidate" for a satisfactory mark of the "sorts of entities [that] make up the content of the stream of consciousness" ["Incorrigibility as the Mark of the Mental," The Journal of Philosophy, LXVII, 12 (June 25, 1970): 399–424, pp. 406–7; cf. also Rorty's Philosophy and the Mirror of Nature, e.g., pp. 88–96, and Part I, passim.

[12] As Hector-Neri Castaneda and G. E. M. Anscombe have pointed out, believing something of oneself must involve the mental analogue of the indirect reflexive construction, represented here by 'oneself' [Castaneda: 'On the Logic of Attributions of Self-Knowledge to Others,' The Journal of Philosophy, LXV, 15 (August 8, 1968): 439–56 and elsewhere; Anscombe: 'The first person,' in Mind and Language, ed. Samuel Guttenplan (Oxford: Oxford University Press), 1975: 45–65]. Even when George believes that somebody who turns out to be George is F it may not be true that George believes that he, himself, is F. For example, George may truly believe somebody is F while wrongly believing that that person is not George, himself. Or he may not even believe of himself that he is George. Unlike token-reflexive constructions, these terms involve anaphora. But the clauses that contain them are grammatical transforms of sentences that do contain genuine token reflexives.

[13] See Franz Brentano, Psychology from an Empirical Standpoint, tr. Antos C. Rancurello et al. (London: Routledge & Kegan Paul and New York: Humanities Press, 1973), pp. 129–30.

[14] G. E. Moore, 'The Refutation of Idealism', in his Philosophical Studies, (London: Routledge & Kegan Paul, 1922): 1–30, pp. 20 and 25.

[15] I am grateful to Georges Rey and Eric Wefald for independently raising this point.
[16] On this point, see Robert M. Gordon's illuminating 'Emotions and knowledge,' The Journal of Philosophy, LXVI, 13 (July 3, 1969): 408–13 and 'The aboutness of emotions,' American Philosophical Quarterly, 11, 1 (January 1974): 27–36.
[17] It is noteworthy that this very attitude appear even in Freud's own writings. Freud does, indeed, "energetically den[y] the equation between what is psychical and what is conscious" ['Some elementary lessons in psycho-analysis,' in The Complete Psychological Works of Sigmund Freud, tr. and ed. James Strachey (London: The Hogarth Press, 1966–74) (henceforth "Works"), XXIII: 279–86]. And he understood that to do so one must define the mental in terms of phenomenal and intentional character; thus he insisted that "all the categories which we employ to describe conscious mental acts ... can be applied" equally well to unconscious mental states ('The Unconscious,' Works, XIV: 166–215, p. 168). Moreover, he maintained that "[t]he psychical, whatever its nature may be, is itself unconcious" ('Some elementary lessons,' p. 283), and so, "[l]ike the physical, the psychical is not necessarily in reality what it appears to us to be" ("The Unconscious," p. 171; Cf. 'Some elementary lessons', p. 282 and The Interpretation of Dreams, Works, V, p. 613). But despite all this, Freud operated with a surprisingly Cartesian concept of consciousness. Consciousness, he wrote, is a "unique, indescribable" quality of mental states ('Some elementary lessons,' p. 282), and "the fact of consciousness" "defies all explanation or description" (An Outline of Psycho-Analysis, Works, XXIII: 141–208, p. 157). In thus regarding consciousness as unanalyzable, Freud seems to have uncritically accepted the core of the Cartesian doctrine he strove to discredit. (To dissociate the present account from Freud's views I eschew here the colloquial 'unconscious mental state' in favor of the somewhat awkward term 'nonconscious'.)
[18] 'Identity and necessity,' in Identity and Individuation, ed. Milton K. Munitz (New York: New York University Press, 1971): 135–64 [henceforth "IN"], p. 163, n. 18; emphasis original here and elsewhere.
[19] Naming and Necessity (Cambridge, MA: Harvard University Press, 1980) [henceforth "NN"] p. 151. Compare Reid, Essays, II, xvi, p. 243: "When [a sensation] is not felt, it is not. There is no difference between a sensation and the feeling of it; they are one and the same thing.
[20] Even Freud does not hold that feelings can strictly speaking be unconscious, though he sees no difficulty about unconscious intentional states (The Ego and the Id, Works, XIX: 3–68, pp. 22–3; cf. An Outline, p. 197).
[21] Thus it is a parody for Wittgenstein to suppose that all we could mean by an unconscious toothache, e.g., is "a certain state of decay in a tooth, not accompanied by what we commonly call toothache" [The Blue and Brown Books (Oxford: Basil Blackwell, 1958), p. 22].
[22] On difficulties in Nagel's discussion and, especially, his notion of a point of view, see my 'Reductionism and knowledge,' in How Many Questions?, ed. Leigh S. Cauman, Isaac Levi, Charles Parsons, and Robert Schwartz (Indianapolis: Hacket Publishing Company, 1983): 276–300.
[23] NN 54–9, "IN" 156–61, and Putnam, 'The meaning of "meaning",' in Putnam's Philosophical Papers, vol. 2 (Cambridge: Cambridge University Press, 1975): 215–71, pp. 223–35.
[24] NN 153; cf. "IN" 162–3. Again, cf. Reid, Essays, II, xvi, p. 243: A "sensation can be nothing else than it is felt to be. Its very essence consists in being felt." Cf. also J. J. C. Smart: "[t]o say that a process is an ache is simply to classify it with other processes that are felt to be similar" ['Materialism,' The Journal of Philosophy, LX, 22 (October 24, 1963): 651–62, p. 655. It is striking that this Cartesian claim about our mental concepts should be shared by theorists who, in other respects, diverge as sharply and as thoroughly as do Smart and Kripke. That it is so shared suggests that this claim

may underlie much of what is, in different ways, unintuitive about each of those theories.

[25] I am greatly indebted to many friends and colleagues for comments on earlier versions of this paper, most especially to Margaret Atherton, Adam Morton, and Robert Schwartz.

The Graduate School and University Center of the City University of New York,
Ph.D. Program in Philosophy,
New York, NY 10036-8099,
U.S.A.