Much of the perplexity that motivates modern discussion of the nature of mind derives indirectly from the striking success of physical explanation. Not only has physics itself advanced at a remarkable pace in the last four centuries; every hope has been held out that, in principle, all science can be understood and ultimately studied in terms of mechanisms proper to physics. Seeing all natural phenomena as explicable in terms appropriate to physics, however, makes the mental seem to be a singularity in nature. Chemistry and biology may well be reducible to physics, but the same seems hardly possible for the mental. The gulf between mind and physics seems too great to bridge, and the success of physics guarantees its standing. The place of mind in nature is thereby rendered problematic. This line of reasoning has tempted thinkers since Descartes to see the mind as not only independent of other natural phenomena, but as even somehow lying outside the natural order itself.

A variety of particular problems about how the mental fits with the rest of nature have been widely discussed in recent years. Less often noticed, however, is that similar problems appear to affect our understanding of the concept of the self in relation to the natural order. For something to be, or have, a self, two conditions seem intuitively necessary. There must be some sort of unity in the mental life of that being. In addition, to have or be a self, one must be distinguishable from other beings, and in particular from other beings of the same or of a relevantly similar sort. Intuitively, nothing can be a self unless it functions in some suitably unified way and unless there is a reasonably clear contrast between it, considered as a self, and other things distinct from it. It is perhaps plausible to see the
conceptual resources of modern physics as adequate to explain the physical unity and individuality of the macroscopic objects around us, including the biological integrity of living organisms such as ourselves. But it may well seem that, however well physics can accomplish those tasks, its conceptual resources are simply insufficient to explain the special functional unity and individuality involved in something’s being a self.

The difficulty emerges especially vividly when we contrast the way a common sense view of things represents the place of the self in nature with the way this might be represented by the view of the world suggested by physics. On the common sense view, the world consists of macroscopic objects of various kinds, which behave in the ways they do because of their intrinsic properties. This sort of view leaves it open to regard those organisms which have mental abilities as simply special sorts of organisms. There may still be a problem in saying exactly what mental abilities are and what properties and capacities are special to such organisms. But the problem is not that of explaining how the mental can arise and exist in nature, but the problem of articulating just what the mental is and trying to determine what is special to those organisms that have mental states. On a common sense view of this sort, the special unity and individuality that characterize selves also poses no special problem. Organisms generally exhibit a remarkable degree of functional unity because of their biological makeup. This functional unity results in a clear contrast between individual organisms and their surrounding environment, including other individuals. Organisms that have more elaborate abilities and properties will function in ways that are more distinctively unified and more effectively resist impingement from the environment. Such organisms will therefore be selves to a greater degree than less elaborate organisms. Various psychological and social abilities will presumably be the culmination of such functioning.

By contrast, modern physics conceives of natural phenomena as due to the behavior and interaction of nonmacroscopic entities governed by mathematically formulable laws. Such laws must generate and explain the relative autonomy of biological organisms by appeal to the nomological behavior of interacting parts, ideally at the biochemical and biophysical levels of analysis. The emergence of selves among the furniture of physical reality needs special explanation, therefore, as does the existence of mind itself. Many inanimate objects exhibit functional unity and individuality, and yet are not selves. The common sense view accounts for this by appeal to the elaborateness and kind of unity of function. But from the viewpoint of physics, distinctions of elaborateness and kind of functional unity will at best seem artificial and ad hoc. The attempt to explain the functional unity and individuality of selves on the basis of the intrinsic properties of macroscopic objects may strike one as too easy; the concept of the self seems
almost to be built into the common sense view of things. But by the same token the attempt to explain such unity and individuality in terms proper to physics may well seem simply hopeless.

It is natural to conclude that the self cannot be understood at all on the basis of physical considerations, and must instead be accounted for in terms of the mental. If neither mind nor self can be located within the natural order, the temptation will be great to understand each in terms of the other. Moreover, only the mental seems available as an explanation of the self once the physical has been excluded. The most dramatic example of this line of reasoning is Descartes’ identification of the self with a being whose essence is simply that it thinks. But an account of the self in terms of the mental need not invoke mental substances; it can rely instead on mental states.

Though avoiding commitment to mental substances has evident advantages, it is far from clear that an account of the self cast solely in terms of mental states can succeed. In section I, I consider the three most common approaches to such an account, and argue that each fails. These approaches rely, respectively, on propositional mental states, sensory mental states, and desires. In section II, I turn to an examination of the intuitive connection that holds between being a self and being a center of consciousness. I argue that this Cartesian idea also cannot do justice to our conception of the self. Finally, I argue in section III that a better account of the self can be constructed that is based not on propositional, sensory, or desiderative mental states, but on the emotions. And I urge that the close connection that holds between the emotions and our concept of the self indicates both why a purely mental account of the self is likely to fail and what sort of account holds promise for success.

I.

Among attempts to account for the self on the basis of mental states, that which conceives of the self in terms of propositional mental states seems to many to be the most intuitively compelling. Propositional mental states—"thoughts," for short—can be characterized in two distinct respects: their propositional content and the mental attitude held toward that propositional content. Nothing about the propositional content of thoughts seems to give us any special grip on the concept of the self. Even if your mental states are distinctively yours and mine are distinctively mine, the propositional content of our thoughts is that aspect of them which can be
shared. If I say that \( p \) and you understand me, we have thoughts that have the same propositional content; I, if sincere, believe that \( p \) and you understand me to say that \( p \). And if you believe me and I was sincere, we end up both believing that \( p \); what we believe, that is, the propositional content of our beliefs, is the same, even though we each have our own occurrent and dispositional beliefs. Propositional content does not help us distinguish among selves; it helps us bridge the gap between them.

The mental attitudes one holds toward the propositional content of thoughts, by contrast, does intuitively reveal a connection between mental states and the self. Typical of such attitudes are believing, doubting, suspecting, wondering, guessing, disbelieving, anticipating, and assuming. These and other attitudes provide a kind of mental frame for the propositional content of one’s thoughts, analogous to the illocutionary force of speech acts. Just as the illocutionary force of an utterance determines what sort of speech act it is, so the mental attitude of a thought determines the kind of mental act it is. Propositional content is what can be shared among distinct mental acts; my mental acts at different times, my mental act of surmising now and my mental act of being confident later, your mental act of doubting and mine of affirming, all these can have their propositional content in common. The distinct types of mental attitude, and the distinct occasions of their being held serve to differentiate these distinct mental acts. Without some suitable mental attitude, however loose or noncommittal, no mental act is possible; propositional content alone can supply no more than the object of that attitude.

These considerations suggest a particular conception of the self. On this view, a self somehow unites thoughts, by holding a variety of mental attitudes towards various propositional contents. The self is what holds the mental attitudes and, because it holds various attitudes, they are unified in a single functioning mind. Without such a self, mental acts would simply exist on their own, unrelated to one another except to the extent of sometimes sharing their propositional content. In this situation, no two thoughts would be more closely united than any others. Accordingly, we would be unable to distinguish between somebody’s having two contradictory thoughts and two beings having thoughts that contradicted one another. Our ordinary conception of mental acts rules out their existing in this disconnected way. But to make sense of that ordinary notion we must presuppose that selves exist and that they are able to unite mental acts by underlying the relevant mental attitudes.

The foregoing argument may convince us that an acceptable account of mental acts requires some notion of a unifying self. It is less clear that the argument helps us develop a conception of such a self, or understand what a self is. Rather, the needed conception of the self must be supplied from some other source. The inability of the argument to help here becomes
obvious when one asks what would be left out of an account of mental acts that made no reference to the self. It is difficult to see how to avoid giving the question-begging answer that what is left out is the unity of mental acts with respect to distinct selves.

In addition, the foregoing argument sheds little if any light on what sort of thing a self is. Presumably, considerations of the sort just advanced convinced Descartes that the self is a thinking thing. But, as Hobbes, in effect, pointed out, overt actions presuppose unification in a self no less than mental acts. An act of walking belongs to a self no less than a mental act of doubting; it may or may not be the same being that walks and doubts, or walks and eats, just as it may or may not be the same self that doubts and suspects. Argument would therefore be necessary to establish the privileged position of mental acts in regard to the nature of the self. It is not easy to see what sort of argument could accomplish this other than one that establishes a distinction between mental and physical substances. For simply regarding mental states as nonphysical does not warrant seeing the self as more centrally bound up with mental activities than with overt actions such as walking.

An account of the self based on the need to unify mental acts also fails to be informative about how a self could do such a thing, and what sort of thing a unifying self would have to be. This difficulty does not affect a conception of the self based, at least in part, on the need to unify overt actions. An act of walking belongs to the same agent as an act of eating, say, because both involve a single body. Short of a commitment to unextended, thinking substances, no parallel move is possible with mental acts. Moreover, it is unclear that even the postulation of unextended substances would help. Such substances are definitionally designed to be substrata of mental states and, in particular, for thoughts. But simply saying that a particular entity discharges some function does not itself make it intelligible how that function is discharged. Eating and walking can involve a single body by involving movements of that body. The idea that unextended substances can unify distinct mental acts is not modeled on such unity of walking and eating, but on the unity that a physical substratum is thought to confer on distinct properties of physical objects, such as shape and color. But such unification by substrata does not explain any unification that is otherwise in need of explanation. Such considerations presumably are the sort that led Kant to conclude that the unity of our mental acts does not involve any empirical tie among them, but is rather a transcendental unity.

The difficulty in understanding how a self could succeed in unifying mental acts, and what sort of entity such a self would have to be, causes problems for the individuating of selves, so conceived. If it is unclear, for example, how my mental acts are bound together, then it will remain at least as unclear why yours and mine are bound into two distinct unities.
The unifying and individuating functions of the self are thus correlative. Without some clear account of why your mental states are not unified with mine, I cannot understand why we are distinct selves. For a conception of the self based on mental states allows no other method of individuating. I return to this difficulty below.

The effort to understand the self as that which binds mental acts together has a distinctively *a priori* flavor. We do not, on this conception, encounter the self in anything we experience, but must instead presuppose it to have an adequate grasp of our thinking. Largely because of dissatisfaction with this kind of *a priori* model, many post-Cartesian thinkers have sought to isolate the mental determinants of the self not in our propositional mental states, but rather in our perceptual and bodily sensory states. Sensory states not only differ from propositional states in lacking conceptual content; they also differ in having, unlike thoughts, some qualitative or phenomenal character, such as the redness of a visual sensation or the painful character of pains. Both thoughts and sensations are described in ways that parallel our descriptions of particular sorts of nonmental states and events. Both mental acts and speech acts are described by means of propositional clauses governed by verbs of mental attitude or illocutionary force. Our descriptions of sensory states, on the other hand, parallel the nonmental terminology we use to talk about the perceptible properties of physical objects and processes, and perceptible events that affect our bodies. Bodily sensations, such as pains, can be throbbing, stabbing, or dull, just as wounds to our bodies can be. Perceptual sensations can be red, triangular, loud, or salty, just as the perceptible physical objects and processes are that, in standard circumstances, cause corresponding sensations. For present purposes, the ordinary distorting effects of perspective, distance, and intervening medium can be disregarded. Likewise, difficult problems about the precise status of properties such as color, sound, smell, and taste, when these are conceived of as properties of physical objects, can here be set aside.

Though we can correctly characterize both physical objects and sensations in sensory terms, for example, as being red, the redness of a sensation is not the same property as the redness of a physical object, but a mental analogue. A table can be red, at least in common sense terms, and its redness is a first-order property of the table. But a sensation is not an entity, on a par with a table or a person; it is rather a state of a sentient being. So if a visual sensation is red, the redness is a second-order property; the sensory state itself is the first-order property of the sentient being. Moreover, if somebody has a sensation of red, there need be nothing relevant in the sentient being that has the physical color of red. The relations between physical and mental qualities are complicated and notoriously difficult to articulate. But even if some reductionist effort to define
one kind of property in terms of the other can succeed, it is important that, short of some such reduction, the two sorts of property are distinct.

Like thoughts, sensations are intuitively tied together, and it is tempting, therefore, to suppose that the self is somehow responsible for this tie. But the unity of sensations seems also to be less abstruse and more readily accessible to analysis than the unity of mental acts. Visual sensations, for example, are unified by virtue of their belonging to a single visual field, and similarly with other sensory modalities. Moreover, if a red and triangular sensation belongs to a particular visual field, intuitively this is because that sensation bears discernible sensory relations to other sensations in that visual field. For example, the sensation may be visually adjacent and to the left of some other visual sensation. These considerations suggest that sensations are unified in the sensory fields that contain them because of factors connected with the sensory content of those sensations.

Similarly, an auditory sensation and a visual sensation can seem both to occupy the same, or a similar, location in their respective sensory fields. This is not because some overarching or neutral phenomenal field exists, which subsumes or envelopes the special fields of the various sensory modalities. Rather, we correlate locations in our auditory field with those in our visual field because we associate the sensations that occur in various locations in one field with sensations that occur in various locations in the other field. Such association of auditory with visual sensations is natural, since we ordinarily take sensations of the two modalities to have common, nonmental causes. Correlations of phenomenal places in the field of one sensory modality with places in the field of another result from associations of the various sensory contents that occur at the relevant places. In any case, since sensations of distinct modalities can seem to be located in corresponding phenomenal places, we can join together the relevant sensory fields into a single experiential field, including perceptual sensations of all modalities. Again, this merging of fields is intimately bound up with the qualitative content of the particular sensations involved; for it is one aspect of the qualitative content of a sensation that it occurs at a particular position in the relevant field.

The unity of sensory mental states into an experiential field suggests a second way of accounting for the self in terms of mental states. One can regard the self as a matter of the unity of this experiential field; this view would echo Hume's claim that the self is "nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement." Such an account would have an advantage, intuitively, over an account that relies on the unity of mental acts. For, if the foregoing considerations are correct, we presuppose the unity of mental acts and the consequent need for a self to unify them because of a priori arguments about the nature of mental acts.
The unity of sensations, however, seems instead to be based on empirically discernible characteristics of these mental states. It is not surprising that seventeenth- and eighteenth-century rationalists should have generally adopted the first approach toward understanding the self, whereas empiricists of the same period should have tended toward the second.

Once we turn from the unifying to the individuating role of the self, however, an account in terms of the unity of sensory mental states is far less promising. The unity of such states does seem to be partly a function of their sensory qualities. But the kind of unity that results from considering sensory qualities is inadequate to explain why sensory mental states belong to distinct selves. If two sentient organisms are similarly located and have similar sense organs and cognitive apparatus, their sensations will presumably be strikingly similar. The sensations of one sentient being would, in these circumstances, be qualitatively as close to those of another as they are to any others that belong to the same sentient being. The unity of sensations therefore seems to be no more helpful or informative about the way mental states belong to distinct selves than the unity of propositional mental states was. The operation of the self must therefore be presupposed in the unification of sensory mental states, just as it is in the unity of mental acts.

Indeed, the apparent advantage of relying on the unity of sensory mental states proves to make the problem of individuating all the more pressing. The selves that are supposed to unify propositional mental acts are conceived on the model of substrata, which operate in an unknown way and whose existence is established by a priori considerations. Such reasoning may fail to explain how distinct selves are individuated. But such a model does no better in explaining how selves unify mental acts; on this model, one simply stipulates that the self can unify mental acts into distinct, individual groups. But, if one conceives of the self as depending, instead, on the occurrence of many sensations in a single perceptual field, the unifying and individuating functions of the self will be on different footings. Mental states will then be unified by the operation of the perceptual field, rather than by any presupposed substratum. But, in the absence of such a substratum, the individuating of distinct selves will be all the more inexplicable. The ability of the model based on sensations to avoid substrata in unifying mental states makes the subsequent individuating of selves seem impossible, unless some other model is invoked to supplement the perceptual field. Perhaps reasoning of this sort is responsible for the pronounced tendency towards solipsism that results from the adoption of a model of the self based solely on the unity of the perceptual field. Even invoking some sort of individuating substrata will not fully dispel this tendency toward solipsism. For, as noted above, unless the operation of an underlying self is understood in unifying mental states, it will remain
mysterious how they can be unified into different groups corresponding to
distinct selves.

It may seem that desires might help in giving an acceptable account of
the self. The way in which mental acts such as believing, doubting, and
suspecting are unified is obscure partly because of the abstract and intel­
lectual character of these mental acts. At best, beliefs and doubts have an
oblique connection with nonmental reality, a connection mediated by the
occurrence of other mental states that involve perception and action. Those
which involve perception seem not to help. If their sensory field unifies
them, this unity is powerless to explain how distinct selves can be indi­
viduated; if something other than perceptual fields is responsible for such
unity, an account of the self based on these mental states has no advantage
over an account based on thoughts. Desires resemble thoughts in suggest­
ing a need for an underlying self needed to unite them; like thoughts,
desires involve both a mental frame and some propositional content. Unity
is presumably to be achieved by a single self holding the relevant desidera­
tive attitudes. But in this case, the unifying function may seem more
accessible than in the case of suspicions, doubts, and beliefs. For that
which underlies the unity of desires involves the notion of agency; desires
belong to a single self because a single agent hold these desires.

The concept of agency cannot help, however, in the attempt to give an
account of the self in strictly mental terms. If the concept of agency in­
volves overt actions, then agency would be sufficient to individuate distinct
selves. For the physical behavior involved in action would be a matter of
movements of distinguishable bodies, and we could then individuate selves
by reference to these bodies. But, if agency is understood in strictly mental
terms, no such method of individuation will be available. For, if overt
action is ruled out, there is nothing to the concept of agency beyond the
unity of a variety of desiderative and aversive mental acts, together with
other interacting propositional attitudes and perceptions. And by itself, the
unity of desiderative mental states is no more comprehensible than that of
other propositional attitudes. The situation is parallel to that of perceptual
mental states. If perceptions are regarded as achievements, so that seeing
and hearing things imply the presence of nonmental objects that are seen
and heard, then there is no problem about individuating selves by reference
to perception. Different selves can be thought of, then, as having distinct
causal relations to various nonmental objects, which in turn can be inde­
pendently individuated. But, if perceptions are understood in strictly men­
tal terms, so that nonmental objects of perception are excluded from
consideration, such mental states cannot help individuate selves.
II.

It might be objected that the discussion has thus far not touched on the central insight that motivates a purely mental account of the self, an insight that might well be invoked to deal with the difficulties raised above. Intuitively, the unity of mental states consists in their belonging to, or being immediately accessible to, single centers of consciousness. The idea of centers of consciousness explains much of the appeal of the Cartesian identification of the self with the mind. For centers of consciousness presumably connect all the mental states that exist in or for a particular self; the unity of the self, on this view, consists in the presence of a variety of mental states to a single center of consciousness. If this line of reasoning is correct, the failures noted above in the accounts based on thoughts, sensations, and desires should come as no surprise. Without the operation of centers of consciousness, such mental states cannot be unified in any comprehensible or effective way, nor will it be intelligible how they can belong to distinct selves.

The notion of centers of consciousness is doubtless similar to the idea of individual points of view, which Thomas Nagel has forcefully and eloquently argued is necessary for an adequate account of subjectivity. But it is far from obvious what subjective points of view or centers of consciousness might amount to. In particular, if we are to understand the unifying and individuating functions of the self in terms of such centers of consciousness, we shall need some grasp of how centers of consciousness operate in performing these functions. Intuitively, mental states are unified by being present to a single center of consciousness, and distinct selves are individuated because mental states are present to distinct centers of consciousness. But without an account of what it is for a mental state to be present to a particular center of consciousness, these explanations will be no less idle than explanations that appealed just to the mental states themselves.

Two models of the consciousness of mental states are available that might help give some account of what is involved in a mental state's being present to a center of consciousness. On the first, all mental states are automatically conscious states, and such consciousness consists in every mental state's somehow being transparent to itself. This model is implicit in much that has been written in the Cartesian tradition. All mental states know themselves, on this account, and all are therefore, in Ryle's useful phrases, "self-intimating" or "self-luminous." Substantial difficulties face this model of consciousness, difficulties that seem decisive against it. But in any case this model is unable to help with the present problem. For even
if mental states know themselves or are otherwise self-transparent, this does not explain how two mental states can be present to one center of consciousness, or two others can be present to two distinct centers of consciousness. Each individual mental state may be self-luminous without thereby belonging to one self, rather than another. This model of the consciousness of mental states may initially appear to help because it echoes the Cartesian doctrine that all the mental states of a Cartesian mind are transparent to that mind. But that doctrine, once again, presupposes rather than explains the way in which mental states are present to distinct, individual selves. And by itself, the self-luminosity of individual mental states cannot supply the unity necessary for the concept of the self.

The second model of the consciousness of mental states relies on the idea that being conscious of something means being aware of it, and being aware of something amounts to having a thought about it. If I have a thought about a tree, I am aware and therefore conscious of the tree; if I have a thought about a mental state, I am correspondingly aware and therefore conscious of that mental state. This model seeks to capture the immediacy that characterizes our consciousness of our own mental states by stipulating that the higher-order thought one has about the conscious mental state be noninferentially arrived at. But we have noninferential thoughts about the mental states of others, as well as about our own. So this model specifies that the higher-order thought must involve reference to oneself; it must be a thought that one is, oneself, in the mental state in question. It is not problematic for this account that we are generally unaware of the occurrence of such higher-order thoughts. If the consciousness of mental states consists in having a higher-order thought about them, those higher-order thoughts will themselves not be conscious thoughts unless one also has a still higher-order thought about those thoughts. Suitable care in developing this kind of account allows one to do intuitive justice to the phenomena of conscious mental states at least as well as an account based on the Cartesian notion of intrinsic self-transparency.

One might expect, however, that such a view of consciousness could do little if anything to capture or explain the unity of centers of consciousness. For, on the current account, the consciousness of mental states is an extrinsic property they have, and extrinsic properties may seem too accidental to hold out promise of an adequate account of such unity. But the intrinsic character of the consciousness of mental states, on the Cartesian model just considered, provided no help in understanding the unity of the self. And the conception that relies on higher-order thoughts actually fares better on this score. For the higher-order thought that confers consciousness on other mental states represents itself as belonging to the same self as the mental state that it is about. Higher-order thoughts thereby succeed in yolking themselves together with the mental states that they make con-
scious. Moreover, higher-order thoughts are possible that, by being about two or more mental states, would bring unity to them.

Even though an account of consciousness that invokes higher-order thoughts induces more unity among mental states than a Cartesian model can, it still falls far short of any serious success. The idea that mental states are united by all being present to a single center of consciousness cannot be explained by piecemeal appeal to various higher-order thoughts unifying themselves and one or two other mental states. Moreover, the unifying effect of such higher-order thoughts seems, again, to rely on an already presupposed self. For higher-order thoughts are able to unite several mental states by representing them as belonging to the same self that the higher-order thought itself belongs to. If the concept of the self is required to form the requisite sorts of higher-order thoughts, the occurrence of such thoughts will not be able to serve as the source of this concept.

Combining features of the two models just considered suggests a way in which one might be misled into thinking that the consciousness of mental states sheds light on the unity of the self. On the account that appeals to higher-order thoughts, such thoughts can unite mental states that they are about. On the Cartesian view, all mental states must be conscious. If one combines these two doctrines, the result seems to be an ascending hierarchy of higher-order thoughts, each level conferring consciousness on the level immediately beneath it. Since each higher-order thought can reasonably be about two or three mental states on the next level down, the picture is suggested of an ascending hierarchy that tends to converge. At each higher level, fewer thoughts will be needed to confer consciousness of those below. It is tempting to think that, if the hierarchy is traced back sufficiently, there may be one, or at most a few, higher-order thoughts. This would induce enough unity on the lower-order mental states to explain how mental states are unified in particular selves. But the mongrel account just constructed cannot be developed into a defensible view. Only if consciousness is intrinsic to mental states does it make sense to insist that all mental states are conscious; an account of consciousness that relies on higher-order thoughts cannot combine with an account that represents all mental states as conscious save on pain of an infinite regress. And it seems phenomenologically beyond dispute that conscious thoughts are not present to us at the higher reaches of the hierarchy just envisaged.

In the absence of an alternative account of the consciousness of mental states, such consciousness will not help us understand what it is for a number of mental states to be united by being present to a single center of consciousness. And, without some explanation of what being present to a single center of consciousness involves, we will be unable to explain, in turn, the unifying and individuating functions of the self. Indeed, the very idea that centers of consciousness are responsible for the unity of conscious
mental states reveals some of the solipsistic aspect noted in section I of mentalistic accounts of the self. It seems obvious to introspection that we have conscious mental states, and that are somehow unified by belonging to a single self. But there seems also to be little to say in explanation of this unity, except that one's conscious mental states are in any case clearly not states of anything other than oneself. Intuitively, the unity of such states is less a matter of their belonging to a single thing, and being thereby bound together, as it is a matter of their not belonging to anything to which we could possibly think of them as belonging. They are bound together by all of them seeming together to belong to nothing. These considerations suggest that, at bottom, the idea of a center of consciousness that unites conscious mental states is simply the solipsistic idea that, in Wittgenstein's words, "[t]he subject does not belong to the world; rather it is a limit of the world."\(^{12}\)

Much as it invites such solipsistic conceptions, it may seem that a Cartesian thinking substance has a special claim to be considered as fundamental to our concept of the self. Even though a Cartesian mind can do little to explain the unifying and individuating functions of the self, it has a particular advantage which may compensate for these disabilities. For it is arguable that what is essential to something's being a self is that its activities are in some relevant way self-directed. On a common-sense view of things, something can count as a self if its activities are largely self-determined, and if it is able to maintain itself against challenges from the environment. It can be maintained that a Cartesian mind, or a center of consciousness, exhibits these properties to a preeminent degree. The non-physical nature of Cartesian minds guarantees their complete invulnerability to environmental assaults, and their independence of physical things, and of each other, makes their activities self-determined. Cartesian minds are also self-directed in the additional sense that their essential activity is thinking, and on the Cartesian model all the thinking of a particular mind must be transparent to that mind.

But the pure form in which Cartesian minds exhibit such self-determination and autonomy with respect to the environment hardly helps overcome the difficulties of being unable to explain the unity and individuality of these minds. For, without some way to distinguish different minds, and minds from other things, the self-determination and autonomy of such minds are idle versions of independence. Once again, the independence of the self is needed to form a clear grasp of such self-determination and autonomy; for unless we can clearly distinguish minds from other things and from each other, self-determination and autonomy will be empty. The self-knowledge and self-transparency of the mind does not help here. In order to understand what is involved in a mental state's being transparent to, or known by, a mind, one must first understand what it is for the mental
state to belong to the mind. It will not work to explain belonging to a mind in terms of being known by it, for presumably belonging to a mind is a necessary condition of being known by it. As noted above, moreover, these Cartesian doctrines invite the solipsistic reasoning that is destructive to a proper conception of the self. Indeed, it is a short step from the pure self-transparency of the Cartesian mind to the phenomenalist view that the world around us, being constructed from mental qualities, is equally transparent to the mind.

III.

The argument of the last two sections hold out little promise for a successful account of the self cast in strictly mental terms. In this concluding section, I urge that the best hope for such an account lies in an appeal to a type of mental state thus far not considered, namely, the emotions. The argument of this section is not intended to contest the bleak outlook for a purely mental account of the self. Indeed, some of the strengths of an account based on the emotions will be directly related to the likelihood that no strictly mental account can succeed. Rather, the current section seeks simply to argue that an account based on the emotions does far greater justice to the self than an account that appeals to thoughts, sensations, desires, or centers of consciousness.

The principal failings of those accounts of the self considered above concerns their inability to explain how selves can be individuated. Each of the foregoing accounts put great stress on the unity of selves, and each such account does some intuitive justice to this unity. But these accounts failed to explain the unity of the self in a way that allowed for distinct selves to be individuated.

These accounts also relied on one aspect of the mental states in terms of which the account was being constructed. In the case of thoughts, the unifying factor was the mental attitude held toward some propositional content. Since a single self might be seen as holding a large number of such mental attitudes, the attitudes would thereby serve to relate the self to the relevant propositional contents. Similarly, perceptual fields were seen as a possible link between the self and mental states, this time between the self and sensations. These fields function to contain sensations, and therefore serve, like mental attitudes, as a kind of mental frame. Centers of consciousness play a similar role as a mental frame, for intuitively they provide the consciousness of conscious mental states; a mental state is conscious
because it is present to some center of consciousness. None of the foregoing accounts relied, save incidentally, on the distinctive content of the mental states under consideration.

Given the kind of content that thoughts, sensations, and desires have, this reliance on mental frame was justified. Except for the special propositional content of higher-order thoughts considered in connection with the second model of consciousness, nothing about the propositional content of thoughts helps explain how one’s thoughts are united into distinct selves. One might counter that the propositional content of one’s thoughts typically serves to unify those thoughts by virtue of the rational interrelation of thoughts with each other and with other mental states. But the rational connections among one’s thoughts can hardly serve to unite more than a small fraction of them, and the resultant unity is too abstract to be a good candidate for the self. Similar considerations apply to the qualitative character of sensations, which supplies their content. Interpreted perceptually, the phenomenal qualities of one’s sensation presumably form a coherent whole, but except for details of perspective, little if anything about those qualities, as opposed to the field that contains them, can serve to unite the sensations that have the qualities. Moreover, neither type of content permits any ready distinction between different selves. The rational connections among one’s thoughts, for example, provides no basis for distinguishing one person’s thoughts from another’s, and many or most of one person’s sensations could be exchanged with those of somebody similarly situated without significant disruption of whatever coherence obtains among them.

Though the coherence that holds in the mental content of one’s thoughts and sensations is insufficient to explain the diversity of selves, desires fare somewhat better on this score. For one’s individuality is doubtless better reflected in the particular desires one has than it is in one’s thoughts or sensations. But the opposite difficulty arises for the content of desires. Most desires are relatively transitory, and they are often inconsistent. So the distinctive way one’s desires fit together, while it may help capture the way individual selves differ, cannot explain how various desires are unified into particular selves.

Not all emotions are long lived, of course, nor are all desires transitory. Emotions, such as fear of a momentary event, can pass quickly, and desires pertaining to long-range goals can outlast most emotions. Still, desires are keyed to specific desiderata. Ordinarily, when these are obtained, or become impossible, the desire lapses. By contrast, the life span of emotions is far more independent of external circumstance. Fear, anger, and jealousy sometimes subside when the events that provoked them are over, but sometimes these emotions linger. Elation and sadness typically last beyond the events that occasion them. Moreover, when long-range
desires do occur, they must ordinarily be fueled by fairly strong emotions. These considerations lend additional support to the idea that emotions, and the ways emotions fit together, reveal far more about the individuality and unity of particular selves than do desires.

Emotions, such as anger, jealousy, joy, fear, love, sadness, and affection do better than desires, thoughts and sensations in reflecting the distinctive unity of the self, and better than either thoughts or desires in reflecting the individuality of diverse selves. Like desires, emotions have an intimate connection with action, both types of mental state enter into our motives for doing things. And also like desires, emotions are not always rationally consistent. We have mixed feelings about things, and we often feel as we do despite what we rationally think. But emotions, unlike desires, combine in complex patterns that contribute substantially to the makeup of one’s character and personality. The ways that having emotions and being disposed to have them go to produce aspects of character and personality is notoriously difficult to articulate. In part, this is due to the vast variety of character and personality types; the variability of character and personality seems almost limitless, and the ways emotions contribute to the different traits people can exhibit requires detail and subtlety to describe. But it is also difficult to capture the different types of personality and character, and the way various emotions produce them, because the ways emotions combine do not generally reflect ordinary standards of rationality. In this respect, emotions are largely unlike sensations and thoughts, whose content generally fits together in some rationally coherent manner. Emotions cohere, instead, in patterns that reflect various standards of appropriateness that are largely independent of issues of rational coherence.  

The immense multiplicity of ways that the emotions fit together, sometimes satisfying standards of appropriateness and sometimes not, is crystallized in our ordinary talk of personality and character traits. This is vividly illustrated in ethical discussions of the virtues and vices, which rely heavily on connections between the virtues and vices and the emotions. In the *Nicomachean Ethics*, for example, Aristotle contrasts the vices of cowardice and foolhardiness with the virtue of bravery (III, 7). Clearly, cowardice is intimately connected with fear, and while it may be more difficult to specify the particular emotions that underlie bravery and rashness, there can be little doubt that distinctive emotions do correspond to them. These considerations may help explain the prominence in Spinoza’s *Ethics* and Hume’s *Treatise* of detailed and extended discussions of the emotions, in both cases occurring in between general discussions of the mind and explicit treatment of issues of morality.

Because of the great variety in the way the emotions fit together to form traits of personality and character, the content of the emotions is a
promising tool for understanding the individuality of distinct selves. Personality and character are, of course, manifested in overt behavior. But no reference to overt behavior need be made to ascribe such traits, and the underlying connections among the emotions, to an individual, just as no reference to overt action need be made to ascribe to an individual various desires. Unlike the content of thoughts and desires, which may be largely interchangeable between two individuals without affecting the distinctive individuality of the two people, the content of one’s emotions are ordinarily highly distinctive, enough so to characterize distinct selves with an exceptional degree of specificity. And, in those cases in which thoughts and desires seem to be highly characteristic of a person’s individuality, an emotional component is commonly involved. Religious or political beliefs, and desires pertaining to career or way of life, can all be central to a person’s individuality. But to be so, such beliefs and desires must be, to a very high degree, emotionally charged.

In a series of exceptionally lucid and penetrating articles, Robert M. Gordon has argued that a large number of emotions imply knowledge, and that all the rest exclude knowledge. If I am angry that \( p \), I must know that \( p \); if I am worried that \( p \), I must fail to know whether or not \( p \). This analysis implies that one cannot correctly attribute emotions to somebody without also very often becoming committed to various nonmental matters of fact. Gordon’s elegant argument is compelling, but his conclusion shows no more than what is commonly accepted for verbs of perception, such as ‘see’ or ‘hear’: one cannot see that \( p \), or see an \( a \), without its being true that \( p \) or an \( a \) being present. Gordon suggests a disanalogy when he writes that

\[
\text{there seems to be no form of words that entails everything that is entailed by sentences of the form “S is angry about the fact that } p \text{” except that it does not entail (or presuppose) ‘} p \text{’ (“The Aboutness of Emotions,” p. 34).}
\]

Gordon suggests that the best we can do is the cumbersome circumlocution “S is angry because S believes that \( p \)” (ibid.). But a far more natural and concise form of words is readily available: ‘S feels angry that \( p \)’ does not imply or presuppose that \( p \). This move exactly parallels the familiar Cartesian shift from seeing and hearing to seeming to see and hear, and in both cases the move has the desired effect of abstracting the purely mental component of a state, leaving to one side all the nonmental implications. Gordon adds locutions such as ‘feels ashamed’ to his list of verbs of emotion that imply knowledge (“Emotions and Knowledge,” p. 410), but he does not justify this inclusion or even discuss the implications of descriptions of emotion governed by the verb ‘feel’. Perhaps he neglects the peculiarities of such locutions because of his less convincingly argued claim that, since emotions are not “entirely” a matter of how one feels, how one feels is not important to one’s emotions (“The Aboutness of Emotions,” p. 30–31).
The great variety of ways that the contents of emotions fit together to form personality and character traits helps capture not only how distinct selves are individuated, insofar as this is done in mental terms, but also helps explain how the self is unified. Since the contents of one's emotions combine in so many distinct ways, the resultant unity is richer and can therefore make more concrete the unity that constitutes individual selves. By contrast, the sort of unity that results from the juxtaposition of thoughts and sensations, however varied they may be in content, is a relatively abstract affair, devoid of features that could be seen as distinctive of different selves. Moreover, emotions are more often long-range states than desires are, and long-range desires ordinarily require accompanying long-lasting emotions to sustain them. And, since the life span of desires is more dependent on external circumstance than that of emotions, one's emotions will more often be governed by long-range dispositions than will one's desires. So emotions and dispositions to have emotions reflect a long-term continuity that is central to the unity of the self.

There is another respect in which the content of emotions holds out promise for an account of the self cast solely in mental terms. Thoughts and desires can be about anything whatever, and sensations can represent any perceptible nonmental object as well as many particular states of the body. But emotions are typically concerned with other selves. Anger, jealousy, love, indignation, envy, affection, hatred, devotion, compassion, and contempt all are normally concerned with other people. I cannot envy or feel compassion for inanimate objects, and if I am angry or hate a situation, for example, I typically if not always think of the situation as though it had been brought about by some animate agency. Joy, fear, pride, disappointment, disdain, and other emotions may or may not be about other selves, but these emotions have an application to other selves as a large part of their ordinary function. Because of the privileged place of other selves in the content of our emotions, this content captures an especially central feature of the self. For the contrast between self and other is not simply a contrast between oneself and everything else, of whatever sort. Rather, it is a contrast between oneself and other beings of the same sort as oneself, that is, other selves. The idea of a self surrounded by other objects none of which are also selves is a degenerate notion of the self. A full-blooded notion of the self requires an environment that contains other selves that interact with one, and with respect to which one forms an idea of what it is to be a self. The content of the emotions reflects this feature of the self to a far greater extent than any other sort of mental state.

Many emotions are sufficiently complex in structure to make it doubtful that one could have them without a fairly clear, operative conception of the difference between oneself and other selves. Among these are the emotions that Hume especially focuses on in Book II of the Treatise, for
example, pride, humility, malice, esteem, respect, benevolence, envy, and pity. Other emotions, such as fear, anger, astonishment, sorrow, disgust, delight, and disappointment, seem to presuppose such a contrast to a lesser extent. But our understanding of even these emotions, and our ability to distinguish among them both conceptually and in practice, still relies on the contrast between oneself and others. For our emotions result from the peculiar effect other things have on us, and these other things must be conceived as distinct and isolable parts of the environment. One may perhaps discern readily enough that one is in a pleasant or an unpleasant emotional state without needing thereby to make any reference to oneself as distinct from other things. But without actively drawing this distinction, it is difficult to see how one could go on to discriminate anger from disappointment, fear from distress, and sorrow from frustration.

Other selves include more than people or members of the human species. Descartes’s extraordinary doctrine about nonhuman animals notwithstanding, there is every reason to think that all mammals have relatively impressive mental lives, and the same is very likely true of many other forms of life as well. Presumably, these mental lives include thoughts, desires, and sensations, as well as emotions. But the close tie between emotions and the self holds with the nonhuman animals as well as with ourselves. We relate to other animals in terms of the emotions we take them to have far more vividly than we do in terms of their thoughts and sensations, which we have a difficult time understanding. And while we may discern the desires of nonhuman animals with no more difficulty than we note their emotions, emotions cover a broader spectrum of the animals’ view of themselves in relation to the environment than is covered by their discernible desires. Doubtless, the vast multiplicity of character and personality traits in humans is not matched in other, less elaborate species by any comparable variability. Still the connections among emotions in nonhuman animals are still far richer and variable than those which we can readily discern as holding among other sorts of mental states.

Not only do the emotions combine to yield distinctive patterns of personality and character; they also serve to unify many of our other mental states as well. Many thoughts and perceptual sensations evoke rich emotional content in us; joy and sorrow, anger and gratitude, pleasure and grief, are all associated with occurrences of thoughts and sensations that have distinctive propositional or sensory content. These associations result in particular patterns among thoughts and sensations that result in unified groups of mental states. In forming these unified groups, the emotions act somewhat in the manner of the higher-order thoughts considered above in connection with the consciousness of mental states. By being about other mental states, these thoughts succeeded in unifying distinct mental states, at least to a limited extent.
But higher-order thoughts accomplish their unifying by virtue of their propositional content, whereas emotions unify various mental states simply by being associated with them. Repeated association presumably results in the relevant thoughts and sensations taking on a kind of emotional tone, and this tone comes to seem to us as a constituent part of the content of the unified mental states. And there is little reason to think that anything other than repeated and increasingly familiar associations occur between emotions and other mental states, which could be responsible for the emotional tone of various memories and other unified thoughts and sensations. Since the unifying that higher-order thoughts accomplish is due to their propositional content, the number of mental states these thoughts can unify is quite small. Because emotions unify mental states by mere association, very large numbers of such states can be brought together. Indeed, the kind of emotional tone that induces a kind of order in many patterns of thoughts and sensations operates sufficiently widely to encompass a large proportion of our conscious mental lives.

The content of mental states is not the only respect in which emotions are superior to thoughts, desires, and sensations in reflecting the unity and individuality of selves. Like desires, emotions motivate actions and other behavior. Because of this, there are close causal ties between emotions and desires. Thus the notion of agency pertains to emotions as well as to desires. Moreover, because emotions have a component that we identify as a kind of bodily feeling, emotions resemble sensations in being bound together because they occur in a single experiential field. In both respects, emotions are unified in ways that resemble the distinctive sorts of unity peculiar to other mental states.

Emotions are not simply bodily sensations; nor is the experiential field in which they occur the same as that in which bodily sensations occur, even though we describe them by using much of the same sort of bodily vocabulary for the two cases. A generalized ache, like a feeling of fear, can be suffused over much of the visceral area; a pang of jealousy can be located in much the way that a stomach ache is. But it is odd to think of the two, in either case, as actually occurring, phenomenologically, in the same space. A pang of jealousy, even if experienced in one’s stomach, for example, in no way intrudes phenomenologically on any bodily sensation one may have. Nor does such a pang seem even to overlap any bodily sensation. It is as though we have two bodily fields of experience, which we have somehow not learned to merge in the way we have merged our auditory and visual fields into a single perceptual field. But in the case of visual and auditory fields, there is overwhelming reason to identify locations in the one with locations in the other; particular auditory and visual sensations co-occur because of common external causes, and this co-occurrence suggests the association of the relevant phenomenal locations across sensory modal-
ties. By contrast, no such reason exists to correlate our ordinary bodily sensations with that component of our emotions which involves some bodily feel. So there is little reason to correlate or identify the phenomenal locations of one kind of bodily feel with the phenomenal locations of the other and, hence, little reason to merge the two fields. Still, it is clear that the bodily-feel component of emotions does occur in a unified experiential field. If I locate a feeling of anger and also a feeling of fear as both occurring around the heart, for example, I experience both as occurring in the same phenomenal location.

Emotions are not only unified in ways that mirror the unity of desires and sensations; they are also unified in ways distinctive of thoughts. For emotions resemble thoughts and desires in having propositional content. It may be an exaggeration to claim, as some do, that emotions are simply a special sort of judgment. But the conceptual content of emotions is clearly manifest in their being about various objects and states of affairs. As in the case of thoughts, this conceptual content involves the holding of various mental attitudes; part of what distinguishes, for example, being furious, horrified, pleased, ashamed, delighted, upset, and excited is the attitude held toward the relevant conceptual content. And, as in the case of thoughts, the occurrence of these mental attitudes seems to presuppose a unified self that holds the attitudes; the attitudes intuitively link the content of the emotions to the self. These various ways in which the emotions resemble thoughts, desires, and sensations suggest that emotions share somehow in the sort of mental frames that are peculiar to each of the other three sorts of mental state. And, since the mental frames of each of those three sorts of mental state helped explain the unity of selves, the emotions once again appear to have the closest tie to the self of any sort of mental state.

Accounts of the nature of the emotions are very frequently reductive in spirit and intention. One reads that the emotions are special sorts of judgments, that they are bodily sensations, that they are elaborate forms of desire and aversion, or that they are suitable compounds of these various elements. If some such reductive account were correct, the argument of this paper would be jeopardized. For, if emotions are special sorts of judgment, desire, or bodily sensation, it will be wrong to urge, without suitable explanation, that emotions contribute more to our concept of the self than these other sorts of mental states do. It would still be possible, in that case, that, even though thoughts, desires, and sensations contribute little in general to our understanding of the self, those special sorts of thought, desire, or sensation which are the emotions contribute much. One would then need to explain, however, how the emotions accomplish this despite the arguments advanced above against thinking that any sorts of thought, sensation, or desire can be of significant help. Similar considerations hold
for the reductivist view that emotions are suitably compounded mixtures of thoughts, desires, and sensations.

There is another kind of consideration that might be invoked in favor of a reductivist account of the emotions. The idea of what it is to be a mental state can be captured in two different ways, by appeal to immediate awareness of some sort, and by appeal to families of characteristics that are distinctive of mental states. But these two marks do not give coextensive results, and it is arguable that the first is inadequate. Mental states occur of which we are unaware. Such states cannot be distinguished by their having consciousness, but it is arbitrary to deny them mental status simply on that account. Immediate awareness is therefore not what determines states as mental, though it is the way we fix the extension of the term ‘mental’, as well as the various terms for the kinds of mental states. Rather, there are characteristics shared by conscious and nonconscious mental states, which are the basis of a rough and ready mark of the mental.

The obvious candidates for these characteristics are propositional content and sensory quality. States of both kinds are picked out by immediate awareness, so it is no objection that propositional and sensory character simply yield two unrelated groups of states. The possibility of being the subject of introspective awareness is enough to count otherwise diverse states as all being mental states, just as reproductive and self-nourishing capacities suffice to bring diverse phenomena all under the heading of life forms. One might object that one cannot tell what it would be like for phenomenal qualities to occur without consciousness, or what it would be like to think something but not consciously. But this only shows that one cannot know what it is like to be in these states without being in conscious versions of them, not that qualitative and propositional mental states must be conscious. There is also no way to know what a table, person, or mathematical theorem is like without having been aware of the table, person, or theorem.

Nonmental items do have propositional content and qualitative character; tables have visible color and tapestries have presentational content. But such color and content cannot, according to common sense, be the subject of introspective awareness. Moreover, the mental frame of mental states is absent in these other cases. Though an analogue of the mental state of thoughts occurs in speech acts, it is only an analogue. One might object that these considerations make any such account of mentality circular; what is mental is distinguished by having a mental frame, or by being able to be the subject of that peculiarly mental form of awareness we call introspection. But there is compelling reason to doubt that an ultimately noncircular mark of the mental is possible. Nor is it obvious that such a mark should be missed. Noncircular marks of important categories can seldom be found without extensive scientific investigation and sophisticated theoretical de-
velopment. In advance of such input, common-sense marks tend either to be question begging or extensionally inadequate. But informative, extensionally adequate marks of specific subject matters are not needed to fix those subject matters, as sciences such as biology and chemistry testify. • The quest for traditional essences is frequently a confused search for a way to fix the extension of central terms in a way that also expresses a nontrivial, intensionally adequate definition.

The idea that sensory quality and propositional content are the distinctive characteristics of the mental suggests a reason for giving a reductive account of the emotions. Thoughts have propositional content, but lack sensory quality; desires are perhaps the same, if we abstract from imagistic representations of the desired object. Sensations, by contrast, have sensory qualities without propositional content, at least if one abstracts from the thoughts that invariably accompany experiences of actual perception and, presumably, many hallucinations. Emotions seem, therefore, to be the only kind of mental state having both sorts of distinctive mental characteristic. It may be tempting to conclude that emotions are merely amalgams of sensory mental states and propositional mental states, rather than distinctive types of mental states themselves.

These considerations, however, are far from compelling. The number of distinctive categories of mental states need not correspond to the number of basic kinds of properties in terms of which we characterize mental states. The broad types of mental states will in part be distinguished by virtue of these properties, but other factors will enter as well, such as kinds of mental frame and how the states causally interact with other states and with behavior and stimulations. To some extent, how many kinds of mental state we say there are is arbitrary. Moreover, it seems idle whether we say, for example, that a thought that \( p \) is a state of an organism that has, in turn, the property of being that \( p \), or instead that it is directly a propositional property of the organism. Our choice about these issues will hinge largely on terminological convenience and indirect theoretical considerations. But, even if mental states are seen as immediate propositional and qualitative properties of organisms, emotions would not need to be regarded as compounds of such properties. Rather, they could be single properties that demand description in both qualitative and propositional terms.

If, moreover, emotions are compounds of propositional and sensory mental states, rather than mental states with both sensory and propositional character, two possibilities are open. Either they are simply compounds of thoughts and sensation, though perhaps of special sorts, or they are compounds of other kinds of purely propositional and purely sensory mental states. The second alternative, however, is uninviting. No conception presents itself of pure propositional mental states other than thoughts, nor of undiluted sensory mental states other than sensations. To hypothesize that
emotions are compounds of such otherwise unknown states is therefore unmotivated. Unless the components of emotions are at least as well understood as the emotions themselves, there is no reason to suppose that they are built up of such components at all, even if those components are pure in particular ways. If the emotions are amalgams of other mental states, therefore, it is reasonable to think they are made up of thoughts, bodily sensations, and desires.

There are independent reasons, however, for being skeptical of any such reductivist account of the emotions. Emotions resemble thoughts in having propositional content; to be angry or proud is to be angry or proud that something is the case. But the mental frame of thoughts is a mental attitude that appraises the truth of the relevant propositional content. In the case of the emotions, by contrast, the truth of the propositional content is either taken for granted, as with anger and delight, or is not at issue, as with worry and terror.

Emotions also resemble bodily sensations; both are felt mental states, and we describe both in terminology that is heavily laden with bodily metaphor. But, if emotions were special sorts of bodily sensations, or at least contained them as parts, it would be difficult to understand why they occur in distinct, nonoverlapping phenomenological fields. Finally, emotions resemble desires in that both involve evaluations. But, whereas desires pertain directly to one’s doing or not doing something, emotions involve vastly more complex evaluative attitudes towards one’s behavior, towards oneself, and towards others. These considerations constitute a prima facie case that reductive accounts of the emotions will fail to do justice to their distinctive character. Moreover, even if one could imagine a reductive account complicated enough to capture this distinctive character, there could be little hope of arriving at an adequate understanding of the emotions without first having constructed an acceptable nonreductive account. For the internal complexity of the emotions, and the remarkable heterogeneity of the various kinds of emotions there are make it unlikely that a reductive account would capture the peculiar nature of the emotions. And only if we had an account that treated emotions as sui generis among types of mental state would we be able to form a warranted judgment about the adequacy of a proposed reductive account.

The problem of judging the success of a proposed reductive account is more pressing here than in many other cases. Some emotions, such as regret, resentment, indignation, and pride, have distinctive kinds of propositional content that help us pick them out. Others, such as disgust, excitement, and perhaps terror, have particular bodily feels that aid in distinguishing the kind of emotion. But those which have distinctive bodily feels seem to lack propositional content that is peculiar to just those emotions, and conversely. And many or most emotions, such as embarrass-
ment, disappointment, surprise, anger, shame, hopefulness, and worry, may or may not have any accompanying conscious propositional content or sensory character that distinguishes them as kinds of emotions. Disappointment and surprise, for example, or anger and dislike, may occur with identical propositional content and with qualitative feels that do not suffice to distinguish them. Other such pairs can be found. And the mechanical character of attempts to define the varieties of emotions by reference solely to qualitative feel and propositional content also testifies to this difficulty. Doubtless some of our ability to distinguish emotional states is due to nonconscious propositional and qualitative properties. But much of our ability to discriminate kinds of emotions may result instead from some sense we have, whether conscious or not, of how these states causally connect with other mental states, with actual and possible behavior, and with stimulations. But the difficulties in giving any account of the kinds of emotional state which proceeds solely in terms of propositional and qualitative properties tells heavily against the possibility that a reductive account of the emotions might succeed.

The idea that there is an especially intimate tie between the emotions and our conception of the self is not new. This claim is central to Solomon’s discussion in The Passions, for example, and Roger Scruton remarks on this connection in his illuminating and suggestive article, “Emotion, Practical Knowledge and Common Culture.”21 But the foregoing arguments both substantiate and help explain the idea that such an intimate connection holds. For, if no sort of mental state except the emotions contributes significantly to our understanding of the unity and individuality of selves, it will be clear that such a connection must obtain and, to a certain extent, what sort of connection it must be.

Emotions have received relatively little systematic attention since the time of Descartes, compared with other sorts of mental phenomena. Spinoza’s and Hume’s discussions are the most conspicuous exceptions, followed perhaps by that of Descartes himself. And all three of these treatments are strongly reductionist. This tendency to give short shrift to the emotions is doubtless due in part to the difficulty of giving an account that does justice to phenomenological complexity. But another factor that has led to the emotions being somewhat shortchanged is that, of all the principal types of mental state, the emotions are farthest from the Cartesian paradigm of mentality, within which we tend to operate. Mental states, on Descartes’s view, are wholly independent of bodily factors; thus thoughts are paradigms of Cartesian mental states. For sensations to figure as mental states on this conception, we must abstract from the contribution of the sense organs and of the objects perceived. For desires to qualify, we must abstract from overt actions. It is far harder to know what sort of abstraction
would be needed, in the case of the emotions, in order to have pure, nonbodily mental states.

The Cartesian paradigm of mentality not only implies complete independence from bodily factors; it also involves an especially close connection to consciousness. Consciousness is a definitional characteristic of mental states; one might not take note of a particular mental state one is in at a particular time. But in Descartes's words, "there can exist in us no thought of which, at the very moment that it is present in us, we are not conscious." There is nothing needed to be aware of one's mental states, on this conception, other than one's simply casting one's attention that way. Again, thoughts fit the model best; phenomenologically, it seems to take nothing more than one's simply having a thought for one to be aware that one has it. Desires and sensations follow closely. Here, it seems, we may have to focus our attention more carefully, and perhaps also ask ourselves what it is that we are sensing, feeling, or wanting.

But the emotions exhibit no such close connection with consciousness. This is partly because one can be unclear about just what emotion one has. Emotions are not felt in distinctive locations, and the content of one's emotions may well be confused and poorly focused. If one can be confused about which emotion one is feeling, moreover, then knowing which cannot merely be a matter of casting one's attention in the right direction. In addition, we sometimes learn of our emotions not by introspection, but by noting our own behavior, and hearing others comment on what we must be feeling. Even if these occurrences are perhaps not paradigmatic, they are also not rare. A perceptive person can often catch onto what somebody feels, in a particular situation, a bit ahead of that person. Though this also happens to some extent with what we perceive, want, and think, it is the most noticeable with the emotions. The distinctive Cartesian tie between mental states and consciousness is therefore weakest in the case of the emotions.

Because Cartesian models dominate our thinking about the mental, the emotions generally get treated, among mental states, somewhat as second-class citizens. The close bodily connection and the loose connection with consciousness both fly in the face of the Cartesian paradigm. Given the additional tendency, since Descartes, to try to understand the self in mental terms that obey the Cartesian paradigm, it is natural that the close ties between emotions and the self would be overlooked or ignored.

That the emotions fit least well with the Cartesian model of mental states has implications about the program of giving an account of the self solely in mental terms. That those mental states with which the self is most closely connected is the least paradigmatically mental and the most closely connected with the body suggests that a strictly mental account of the self...
may be radically misguided. This conclusion should occasion no surprise, at least when a Cartesian conception of mind is under consideration. For a concept of the mental which requires that the mental be entirely abstracted from everything bodily is an unlikely candidate to serve as the basis of a unified and individuating self.

Descartes's notion of a nonphysical, self-transparent mind echoes remarks Aristotle makes about *nous*: "in the case of those things which have no matter," Aristotle writes in the *De Anima*, "that which thinks and that which is thought are the same." But, as Aristotle also notes, *nous* can do nothing to individuate selves (*DA* III 5, 430*ab*18–25; 429*ab*22–28). The idea of a thinking thing as self may be intelligible in the context of Descartes's search for foundations of knowledge. If one rejects all knowledge except knowledge of oneself, knowing and being a self will thus far be indistinguishable. And it may seem sufficient for something to be a self that it functions reflexively, by being transparent to itself and by being the sole proper object of its own knowledge. But the Cartesian conception is a conception of the self solely by default. Only because nothing exists with which we can contrast the self in any operational way can we succeed in thinking of it as a self at all. It is surprising that an idea of the mind that owes so much to Aristotle's *nous* could ever have gained currency as a conception of the self.25

**FOOTNOTES**


9. For details about such difficulties, see my “Thinking That One Thinks,” forthcoming.


11. I develop such an account in “Two Concepts of Consciousness,” forthcoming.


14. On the relation between emotions and character, see Adam Morton, “Character and the Emotions,” in EE.


16. This claim is advanced, for example, by Robert Solomon, The Passions (Garden City: Anchor Press, 1977), especially pp. 188-190.

17. A useful and detailed defense of a cognitive view of the emotions that does not treat them as kinds of judgment can be found in William Lyons, Emotion (Cambridge: Cambridge University Press, 1980).

18. I have argued against this possibility for the case of propositional mental states in “Talking about Thinking,” Philosophical Studies, XXIV, 5 (September 1973): 283-313. Considerations such as those raised by Sydney Shoemaker (in “Functionalism and Qualia,” Philosophical Studies, 27, 5 [May 1975]: 291-315 and in “Absent Qualia are Impossible—A Reply to Block,” The Philosophical Review, XC, 4 (October 1981): 581-599 suggest that no noncircular account can be given for qualitative mental states either.


21. In EE.


23. It does not follow, of course, from the introspective appearance that thoughts are inseparable from consciousness that thoughts cannot be nonconscious; see “Two Concepts of Consciousness.”


25. I am indebted to Margaret Atherton and Robert Schwartz for helpful discussion and critical comments.
EMOTION: PHILOSOPHICAL STUDIES

EDITED BY K. D. IRANI AND G. E. MYERS