Faith and Philosophy: Journal of the Society of Christian Philosophers

Volume 12 | Issue 4

Article 6

10-1-1995

Concerning the Unity of Consciousness

William Hasker

Follow this and additional works at: <https://place.asburyseminary.edu/faithandphilosophy>

Recommended Citation

Hasker, William (1995) "Concerning the Unity of Consciousness," *Faith and Philosophy: Journal of the Society of Christian Philosophers*: Vol. 12 : Iss. 4 , Article 6.

DOI: 10.5840/faithphil199512429

Available at: <https://place.asburyseminary.edu/faithandphilosophy/vol12/iss4/6>

This Article is brought to you for free and open access by the Journals at ePLACE: preserving, learning, and creative exchange. It has been accepted for inclusion in Faith and Philosophy: Journal of the Society of Christian Philosophers by an authorized editor of ePLACE: preserving, learning, and creative exchange.

CONCERNING THE UNITY OF CONSCIOUSNESS

William Hasker

Ever since Descartes there have been philosophers who have claimed that the unity of conscious experience argues strongly against the possibility that the mind or self is a material thing. My contention is that the recent neglect of this argument is a mistake, and that it places a serious and perhaps insuperable obstacle in the way of materialist theories of the mind.

Ever since Descartes¹ there have been philosophers who have claimed that the unity of conscious experience argues strongly against the possibility that the mind or self is a material thing. This argument surfaced from time to time during the seventeenth, eighteenth, and nineteenth centuries, but recently it has been neglected. My contention is that this neglect is a mistake, and that the argument places a serious and perhaps insuperable obstacle in the way of materialist theories of the mind.

I

The argument against materialism from the unity of consciousness is clearly present in Leibniz, as Margaret Wilson has shown.² Indeed, the classic form given the argument in the Second Paralogism may well be viewed as Kant's reflective development of the argument taken from Leibniz. Kant's version goes as follows:

Every *composite* substance is an aggregate of several substances, and the actions of a composite, or whatever inheres in it as thus composite, is an aggregate of several actions or accidents, distributed among the plurality of the substances. Now an effect which arises from the concurrence of many acting substances is indeed possible, namely, when this effect is external only (as, for instance, the motion of a body is the combined motion of all its parts). But with thoughts, as internal accidents belonging to a thinking being, it is different. For suppose it be the composite that thinks: then every part of it would be a part of the thought,³ and only all of them taken together would be the whole thought. But this cannot consistently be maintained. For representations (for instance, the

single words of a verse), distributed among different beings, never make up a whole thought (a verse), and it is therefore impossible that a thought should inhere in what is essentially composite. It is therefore possible only in a *single* substance, which, not being an aggregate of many, is absolutely simple.⁴

In dubbing this a “paralogism,” Kant seemingly invites us to regard it as wholly unsound—yet his own attitude towards the argument was far more complex than that.⁵ We must remember, to be sure, that the chief objective of Kant’s “transcendental psychologist” is to establish a speculative proof of immortality. And it clearly is Kant’s view that neither this argument nor any other can succeed in doing this. But in certain other respects, his assessment of the unity-of-consciousness argument is far from negative. On the contrary, in the second edition of the *Critique* he wrote, “From this [argument] follows the impossibility of any explanation in *materialist* terms of the constitution of the self as a merely thinking subject.”⁶ And the same claim is repeated in *On the Progress of Metaphysics*: “...therefore materialism can never be used as a principle for explaining the nature of the soul.”⁷ Allison summarizes Kant’s attitude towards the argument as follows:

Thus, Kant’s position seems to be that the unity of consciousness, which the rational psychologist (presumably Leibniz) uses erroneously to establish the positive metaphysical doctrine of the simplicity and hence incorruptibility of the soul, can be used legitimately to establish the weaker thesis of the impossibility of a materialist explanation of the conceptual activities of the mind.⁸

It has to be said that Kant’s partially favorable attitude towards the argument has not been shared by his recent commentators and critics.⁹ But rather than pursue this debate in the Kantian context, I think it will be most helpful for us to move directly to a contemporary version of the argument.¹⁰

II

First, we begin with some stipulations about the materialism against which the argument is directed. This materialism will be assumed to view the brain as a computer-like network of electro-chemical interconnections, processing sensory (and other) data much in the way a computer processes information. Each item of information held within the brain is modelled in the physical state of some portion of the brain. There is not conscious awareness of every item of information the brain contains, but awareness when it occurs consists of the brain, or some part thereof, being in the appropriate physical state. An important question here, still in dispute among brain scientists, is whether items of information are stored in discrete sub-units of the brain, or whether information is stored “holographically”¹¹ so that any given item is spread over a large region of brain-tissue. In what follows, this question will be left open.

As an example of the unity of consciousness, I cite my awareness of my

present visual field.¹² This field includes the impressions from a set of shelves in my apartment's living room, with books on the lower shelves and a number of plants, mostly miniature cacti, above them; behind there is the horizontal pattern of the mini-blinds. All this I observe without scanning or refocusing my eyes: momentarily, as it were. This visual field is not unified in any interesting aesthetic sense, but it is a fact that I experience it as a unity, all at once and not as a succession of discrete experiences.¹³

Before proceeding further, it will be helpful to see what our datum looks like as viewed from the perspectives of two contemporary schools of materialism. According to functionalism, my being aware of my conscious field is my being in a state with certain causal and functional properties. And according to eliminativism, what I describe as my being aware of my visual field is a matter of certain parts of my brain being in certain physical states. Now, it may well be that my being aware of my visual field has exactly the causal and functional properties attributed to it by functionalism. And it might also be true that when I take myself to be aware of my visual field my brain is in such-and-such a physical state. But what I *mean* when I say that I am aware of my visual field is *not* that I am in a state with certain causal and functional properties, or that my brain is in such-and-such a physical state. The meaning of the assertion that I am aware of my visual field is, I think, neither obscure nor difficult to grasp.¹⁴ Furthermore, I believe that, on many occasions, what is thus asserted by me is *true*. The reader, then, is invited to take this datum at face value and to investigate it in its own right, without first translating it into the argot either of functionalism or of eliminativism.¹⁵

Now, my procedure is to take a specific conscious state—the state I am in when I am aware of my visual field, as described above—and ask what physical entity it is that is in that state. The question is surprisingly difficult for the materialist to answer. Initially, to be sure, the materialist may reply that the physical object which is aware of my visual field is I *myself*—that is, my body. But not all of my body is equally relevant to the state in question; my brain, we generally believe, is essential to it in a way that other parts of my body are not. (The brain could be in a similar state even if it were in a vat.) So is it my brain, rather than I, that is aware of the visual field? This raises a question of terminology. When a person performs some action in virtue of some part of that person's body doing something, we sometimes say that the *person* has done something *by means of* the bodily part, and sometimes that the *part itself* has done it. (I point *with my finger*, but also *my finger points*.) It is not clear, in the case of mental states such as thinking and awareness, that one of these ways of speaking is more correct than the other: Do I "use my brain" to figure things out, or does my brain figure things out?¹⁶ In what follows, the latter form of speech will be employed, but all the points made will apply in either case.

Let us say, then, that it is my brain that is aware of my visual field, and I am aware of it in virtue of my brain's being aware of it. But not all of my brain need be involved—some of it may be performing other functions which contribute indirectly or not at all to my awareness of my visual field. So let us narrow the focus even further. Let *V* be the smallest part of my

brain which contains the modelling of all of the information from my visual field. The existence and functioning of *V* may or may not be sufficient for my having the awareness of the field. But it clearly is necessary for the awareness, since if part of *V* is not functioning some of the information-content of the field will be lost.

Should we say, then, that it is *V* which is aware of the visual field? Perhaps so, but we need to consider the composition of *V*. In the light of the computer analogy sketched above, we can say that *V* is a whole composed of physical parts. Many of these parts model information from various parts of the visual field. But no proper part of *V* models all of this information, so it is not possible for any of these parts to be aware of the entire visual field. But if *V* is a whole composed of parts each of which is *not* aware of the visual field, how can *V* itself be aware of it? If we assume that each item of information is modelled in a discrete sub-unit of the brain, we might suppose that each sub-unit is aware of the information it contains, and that in virtue of this *V* is aware of the entire field. (On the holographic model, on the other hand, *all* of *V* is needed to carry the full information for *any part* of the visual field.) But even if the sub-units are assumed to be aware of the items of information they contain, this does *not* enable us to explain the awareness of the entire field; that is as if one were to say that each student in a class knows the answer to one question on an examination, and that in virtue of this the entire class knows the material perfectly! It is true that the members of the class are able, working together, to *reproduce* all of the information, but there may in fact be no one at all who *knows* or *is aware of* all of it. The point is simply that the kind of awareness we are discussing is essentially unitary, and it makes no sense to suggest that it may be "parcelled out" to entities each of which does *not* have the awareness. *A person's being aware of a complex fact cannot consist of parts of the person being aware of parts of the fact.* A conjunction of partial awarenesses does not add up to a total awareness.¹⁷

But may we not suppose that, in addition to the various sub-units which model items of information from the field, *V* also contains a scanning mechanism, *S*, which collects and integrates the items of information and thus makes possible the unified awareness? We may of course suppose this. But note that for this mechanism to do what is required of it, *S* must somehow represent within itself *simultaneously* all of the information of the visual field. Now if *S* does contain all of this information (even if only momentarily), we must still reflect that *S* is itself a whole consisting of parts—and thus the argument begins all over again, with the same result as before.

I think what one naturally wants to say at this point is something like the following: It may be, indeed must be, that the information from my visual field is distributed among a number of distinct brain-units. But this fact obviously does not make a unified consciousness impossible—it would, after all, be absurd to require that all of the information should somehow be concentrated into a mathematical point! Rather, what happens is that the various items of information are as it were drawn together and co-presented in a unified consciousness, and this process is not frustrated by the spatial "spread" of information within the brain.

It seems to me that something like this must be true. But note that "consciousness," as used in the preceding paragraph, does not refer to a brain or to any part of a brain. So I repeat my question: what *physical entity* is it that is aware of my visual field? If materialism cannot answer this question, it is in serious difficulty.¹⁸

For many materialists, the most appealing solution may lie in rejecting the demand, which is implicit in the discussion to this point,¹⁹ that the properties attributed to the brain be accounted for in terms of properties of, and relations between, its parts. For an especially clear formulation of this demand, we turn to Wilfrid Sellars and his "principle of reducibility," which he states as follows:

If an object is *in a strict sense* a system of objects, then every property of the object must consist in the fact that its constituents have such and such qualities and stand in such and such relations, or roughly, every property of a system of objects consists of properties of, and relations between, its constituents.²⁰

Commenting on this principle, Sellars states:

Telling us, as it does, that if an object *is* (as contrasted with *is correlated with*) a whole of parts, its having *P* consists in its parts having properties and standing in relations, it also tells us that if an object has a property which violates the principle, then *in that context* it is *correlated with* rather than *consists of* the "parts."²¹

Three points need to be noted with regard to Sellars' principle of reducibility. First: Sellars' name for his principle may suggest the notion of reduction as employed in the philosophy of science, but it would be a mistake to take it that way. This is a *metaphysical* principle, belonging, as Sellars says, "to logic or general ontology rather than to the philosophy of science."²² Second: Since the intention of the principle is metaphysical rather than epistemological, what is required is that the properties of the system shall consist in properties *actually possessed* by the parts, not necessarily in those *we know* (or believe on independent evidence) to be possessed by them. (Thus, in some cases the requirements of the principle might be satisfied by postulating new, previously unknown, properties for the parts.) Third: The principle deals with the relations between the properties of a system and properties of its parts *at a given time*; it is, one might say, synchronic rather than diachronic.

While I have little desire to plunge into a maze of Sellarsian exegesis, his expression "consists of" (or "in") requires some comment. I take it that Sellars is describing a situation in which a whole is *nothing but* a system of objects—where there is no such thing as a whole "over and above" the sum of the parts—in distinction from a situation in which there is such a whole which has properties of its own which cannot be accounted for in terms of the parts and their properties. The phrase can be helpfully illustrated by Sellars' own examples. The redness of a brick wall *consists in* the redness of its constituent bricks—the redness of the wall amounts to noth-

ing over and above the fact that the visible surface of the wall is made up of the surfaces of the bricks (disregarding mortar lines), and those surfaces are red. And for a certain object to be a ladder "is for its parts to be of such and such shapes and sizes and to be related to one another in certain ways."²³ In contrast, the pinkness of a pink ice cube does *not* consist in its constituent elementary particles having certain properties and standing in certain relations, for none of the particles have properties of which "being pink" could consist. (This helps to motivate Sellars' well-known contrast between the "manifest image" and the "scientific image" of man-in-the-world, as well as the doctrine that the scientific image stands to the manifest image as reality to appearance.)

But perhaps "consists in" is still less than transparently clear; at the very least, we should like to have some criterion for determining when one property consists in other properties. My proposal for understanding this notion is as follows: A property of a whole consists in the properties of the parts if it follows from those properties by logical or conceptual necessity.²⁴ This exploits the familiar idea that, in a valid deductive argument, there is nothing in the conclusion that is not already "contained" in the premises. So we can say that

If an object *O* is a system made up of elements $e_1, e_2, e_3, \dots e_n$, then all the properties of *O* are logical consequences of the properties of, and relations between, the e_i .²⁵

Thus, the fact that the visible surface of a wall is made up of red bricks, implies that the wall itself is red. (The color of an opaque object is the color of its visible surface; the visible surface of the wall is identical with the visible surfaces of the bricks that make it up; all the visible surfaces of the bricks are red; therefore, the wall is red.) And from the fact that an object consists of such-and-such bits of wood in such-and-such relations, it follows that the object is a ladder; one might say that it follows from the description that the object "fits the definition" of a ladder.²⁶ On the other hand, it does *not* follow from the microphysical properties of an ice cube, that the cube is pink,²⁷ and so pinkness is *not* a property of an ice cube considered as a system of subatomic particles.²⁸

I think it is clear that a computer is, in Sellars' phrase, "in a strict sense, a system of objects." It is the very fact that the computer's computational capabilities are deducible from, and thus explicable in terms of, the properties and relations of the computer's parts, that enables us to design computers and to understand their functioning—and this, in turn, is just what makes the computer attractive as a model for the understanding of mind. But the unity-of-consciousness argument places an important barrier in the way of this project, by pointing to a property of the mind—my awareness of my present visual field—which is *not* a logical consequence of the properties of and relations between the brain's physical parts.²⁹

We may conclude, then, that the mind is *logically irreducible* to the brain, in that it has properties which are not logically implied by the properties of, and relations between, the physical parts of the brain. But from this it follows, in view of the principle of reducibility, that the mind is also *onto-*

logically irreducible to the brain; in Sellars' phrase, the mind is "correlated with" rather than "consists of" the brain's parts. This does not, to be sure, go very far towards telling us, in a positive sense, what the mind *is* or how its existence is to be explained. But it does tell us something about what the mind is *not* and about how its existence *cannot* be explained, and for a topic as contentious as this one that is a major gain.

It would perhaps be generally conceded that, *as things now stand*, we do not see how such properties as "being aware of one's visual field" could be deduced from properties of, and relations between, the neurons in the brain. But is there any reason to believe this state of affairs is permanent? Is this not merely one more case in which we can confidently wait for the advance of science to relieve difficulties which at a given juncture seem insurmountable?³⁰

It is always difficult to resist such appeals to the future progress of science, in view of the astonishing progress which has already been made. Nevertheless, resistance may sometimes be called for. The difficulty in the present case is not merely that the particles of physics lack certain properties (e.g., color) that they apparently would need in order for the properties we find in the world to be deduced from them. What is at issue is that, whatever properties we may encounter in the world, these properties are co-presented for us *in a unified experience*—or, we may also say, *to a single subject*. The notion that such an experience may in the final analysis *not* be the experience of a single subject—that in fact the experience inheres in a number of different entities, each of which does *not* have that experience as a whole—is I think simply unintelligible.³¹

The materialist who has followed the discussion to this point has in effect three options. She may take the heroic path laid out for her by eliminativism, by simply denying that it is ever true that a person is aware of a visual field (or of any other complex fact). Or, she may become a logical reductionist, asserting that such complex mental properties are after all logically derivable from the physical properties of the brain's parts. For many, however, the most appealing course may be to reject the principle of reducibility, enabling us to attribute to the brain as a whole (or to the person as a whole) properties which are not implied by the properties of and relations between its constituents. In order to explore this possibility, we will next consider an extremely interesting discussion of panpsychism initiated by Thomas Nagel.

III

In his paper entitled "Panpsychism,"³² Nagel argues for that position on the basis of four premises, "each of which is more plausible than its denial, though perhaps not more plausible than the denial of panpsychism."³³ The four premises are:

1. *Material composition:* Living organisms are complex material systems, made exclusively of ordinary matter.
2. *Nonreductionism:* Mental properties are not physical properties of the organism³⁴ and are not implied by physical properties alone.
3. *Realism:* Mental properties really are properties of the organism (i.e., eliminativism is false).

4. *Nonemergence*: "There are no truly emergent properties of complex systems. All properties of a complex system that are not relations between it and something else derive from the properties of its constituents and their effects on each other when so combined."³⁵

Nagel then goes on to say:

Panpsychism seems to follow from these four premises. If the mental properties of an organism are not implied by any physical properties but must derive from properties of the organism's constituents, then those constituents must have nonphysical properties from which the appearance of mental properties follows when the combination is of the right kind. Since any matter can compose an organism, all matter must have these properties. And since the same matter can be made into different types of organisms with different types of mental life (of which we have encountered only a tiny sample), it must have properties that imply the appearance of different mental phenomena when the matter is combined in different ways. This would amount to a kind of mental chemistry.³⁶

For our purposes the crucial premises of Nagel's argument are Nonreductionism and Nonemergence. Unfortunately, Nagel's explanations of these premises (especially Nonemergence) are not as clear as one might wish, so we shall need to sharpen the focus. Furthermore, the interpretations of the two premises must be related in a certain way if the argument is to succeed. Nonemergence states that the properties of wholes *must* be related to the properties of parts in a certain way, and Nonreductionism states that the mental properties of persons *cannot* be derived from the physical properties of matter. So for the argument to work, the sense in which mental properties *can't* be derived from physical properties has to be the same as the sense in which mental properties *have to be* accounted for by *some* properties of matter, thus forcing the ascription of mental (or proto-mental) properties to the ultimate constituents of matter.

According to Nonreductionism, mental properties are not physical properties and are not implied by physical properties alone. It seems most natural and plausible to take "implied by" here as representing logical implication, and we shall so understand it. But if so, then Nonemergence must be understood as requiring that the properties of a whole are logically implied by the properties of its parts. According to James Van Cleve, however, this

is setting the standards for explainability extremely high. If a property of a whole that follows only with nomological necessity from the properties of its parts is held on that score to be *inexplicable*, won't we have to say the same about any effect that does not follow with logical necessity from its cause? If so, the momentum of Nagel's argument will be such that it can't be stopped short of *complete causal rationalism*—Spinoza's view that causes and effects are related by logical necessity. This is a view that few philosophers nowadays would embrace.³⁷

Van Cleve concedes, however, that there is a way to defend Nagel's premise short of this. If we distinguish between determinative relations which hold *at* a time and those which hold *over* time, then it may be more plausible to insist on the relation of logical implication between the properties of a whole and those of its parts for synchronic than for diachronic relations. "The argument against causal rationalism is at its strongest in the case of effects that succeed their causes in time; for how can what happens at one moment have any logically necessary connection with what happens at any other?"³⁸ Nevertheless, Van Cleve finds even this weaker requirement (which he terms "mereological rationalism") to be implausibly strong; he holds that "there is no reason why emergent properties can't follow with *nomological* necessity from properties of the parts, and in that case they would not be inexplicable for anyone but a causal rationalist."³⁹

Now mereological rationalism, as defined by Van Cleve, is equivalent to the principle of reducibility discussed earlier in this paper. So it becomes necessary for me to explain why the requirement of logical reducibility, as stated by that principle, is the correct one, and should not be replaced by Van Cleve's weaker requirement of nomological reducibility.

Causes may in some sense necessitate their effects, but the connection does not seem to be one of logical or conceptual necessity; thus causal rationalism is implausible. But the principle of reducibility does not require causal rationalism. Suppose that in a complex whole there are two properties, occurring simultaneously, which are linked by nomological necessity. Note that the principle of reducibility does not, so far, require a logical connection between these properties. The connection between the two properties can be as opaque, conceptually speaking, as you please. However, if one of the two properties is said to be a property of the *system as a whole*, the principle poses a question: what exactly is it that *has* the property? By hypothesis, the whole simply *is* "in the strict sense, a system of objects"; there *is no* whole "over and above" the parts of which it is composed. So whatever nonrelational properties the whole has must consist of properties of, and relations between, the parts; there simply is nothing else of which they *could* consist. If a property of the whole is not logically grounded in the properties of the parts, then it is "floating in mid-air," unattached to any real individual—but this is unintelligible.

To see this more clearly, we briefly recapitulate the main argument. A person is aware of some complex object—say, a complicated visual field. By hypothesis, it is the person's body—specifically, her brain—that is aware of this object. Now, take the smallest region of the brain that contains all the information for the visual field. This brain-region is a whole consisting of parts, but none of its proper parts can be aware of the visual field as a whole. And the properties of, and relations between the parts, do not logically imply that there is an awareness of the visual field.

So far, Van Cleve would agree, but he urges that "there is no reason why emergent properties can't follow with *nomological* necessity from the properties of the parts." Now if one state of affairs follows from another with nomological necessity, it is because there is a *causal process* linking those two states of affairs. Causal processes, furthermore, occur in and/or between concrete individual objects. So what concrete objects are available

for the causal processes which result in awareness of the visual field? There are the parts of the brain-region in question, of course—but the properties of, and relations between, the parts are already accounted for; by hypothesis they do *not* imply that there is an awareness of the visual field. And the brain as a whole is not an *additional* concrete object over and above its parts, any more than, in Gilbert Ryle's example, Oxford University is an additional object over and above the colleges, libraries, and so on of which it consists. So Van Cleve's causal processes either occur in and/or between the parts of the brain, in which case they do *not* imply awareness of the visual field, or they do not occur at all.

Since this argument is crucial for the paper as a whole, it may be helpful to state it here in a semi-formalized version, as follows:

Let 'p' designate a human being, by hypothesis a physical object consisting of parts $e_1, e_2, \dots e_n$. 'Ap' says that p is aware of her visual field. 'R' is a complex relational predicate such that ' $Re_1, e_2, \dots e_n$ ' describes all the properties of and relations between the parts of p. Ap is the state of affairs such that 'Ap' is true. 'Imply' stands for conceptual implication.

1.	$Re_1, e_2, \dots e_n$ and Ap, and $Re_1, e_2, \dots e_n$ does not imply Ap.	
2.	$Re_1, e_2, \dots e_n$ causally necessitates Ap.	Provisional Assumption
3.	There is a causal process (CP) connecting $Re_1, e_2, \dots e_n$ with Ap.	2
4.	CP consists of events involving concrete individuals.	
5.	The concrete individuals involved in CP include at most $e_1, e_2, \dots e_n$ and p.	
6.	p is not a concrete individual distinct from $e_1, e_2, \dots e_n$	
7.	The concrete individuals involved in CP include at most $e_1, e_2, \dots e_n$.	5,6
8.	All the events involving $e_1, e_2, \dots e_n$ are already included in $Re_1, e_2, \dots e_n$.	
9.	CP is not a causal process that connects $Re_1, e_2, \dots e_n$ with Ap.	1,6,7,8
10.	It is not the case that $Re_1, e_2, \dots e_n$ causally necessitates Ap.	2,9 indirect proof

I conclude, in view of this, that the principle of reducibility (a.k.a. mereological rationalism) is correct, and under this interpretation Nagel's argument appears to be successful. Ironically, however, the very principle that undergirds the argument also defeats the conclusion he seems to want to draw from it. Recall the reference to "mental chemistry," whereby matter has "properties that imply the appearance of different mental phenomena when the matter is combined in different ways." Nagel seems to assume that, once we attribute mental (or proto-mental) properties to the ultimate constituents of matter, it may indeed be possible to explain the mental properties of human beings (such as being aware of one's visual field) in terms of their parts in a way that satisfies the principle of reducibility (=

Nonemergence = mereological rationalism). But it's clear that this simply cannot be done; to repeat it once more, a person's being aware of a complex object cannot consist of parts of the person being aware of parts of the object. So the result of Nagel's argument is indeed ironic: it defeats not only the materialist views he rejects, but also the "atomistic panpsychism" he intends it to support.

IV

The argument began with an appeal to the unity of consciousness, as exemplified by the phenomenon of one's awareness of one's visual field. In the light of the principle of reducibility it was argued that this awareness must be the experience of a unitary subject, one which is not decomposable into parts none of which possesses this awareness. This conclusion, in turn, rules out as unacceptable standard varieties of materialism, as well as panpsychism of the sort adumbrated by Nagel. Assuming all this is correct, what possibilities are left?

Pretty obviously, standard varieties of substance dualism are not threatened by the unity-of-consciousness argument, and adherents of such views may welcome the argument as support. Suppose, however, that for one reason or another one finds those options unappealing?⁴⁰ There is, I want to suggest, another possibility, one which on first inspection is not all that distant from Nagel's panpsychism. Suppose we attribute to the fundamental constituents of matter (i.e., to matter-energy as such), not indeed conscious experience, but rather a certain *potentiality for consciousness*, a potentiality which is realized given a certain kind (or, one of a range of appropriate kinds) of arrangement of matter in a living organism. However, this potentiality will *not* take the form of "mental properties" of, say, quarks which, when appropriately combined, can yield our everyday garden variety mental properties in accordance with the principle of reducibility. This we have seen to be impossible. Rather, it is the potentiality to produce, under appropriate physical conditions, a *new individual*, a psychic unity which, in Sellars' terminology, is *correlated with* but does not *consist of* the particles of microphysics. Here a comparison can be drawn with such a familiar physical phenomenon as a magnetic field, which is *correlated with* but does not *consist of* the "system of objects" which is the generating magnet. Such an "emergent self" can, I would argue, play the role not only of the consciousness which is aware of the visual field, but of the acting subject of free moral choice. Such a view, which may be termed "emergent dualism," certainly runs the risk of becoming a philosophical orphan, rejected by dualists and materialists alike. On the other hand, I would argue that it ought to be more acceptable to dualists than any of the standard versions of materialism, and more acceptable to materialists than the usual kinds of dualism. But these are matters which require a great deal more discussion than is possible here.

Nowhere in this discussion has appeal been made to considerations which are unique to Christian or theistic philosophy. Nevertheless, it will be evident that these are issues in which Christian philosophy has a major stake. And I believe that the self as conceived in emergent dualism is a

good candidate for the core of identity of a human person before God, and a plausible recipient for the promise of eternal life. If so, then there is much about this view that should interest us.⁴¹

Huntington College

NOTES

1. Meditation 6.
2. Margaret D. Wilson, "Leibniz and Materialism," *Canadian Journal of Philosophy* 3 (June, 1974), pp. 495-513.
3. It is difficult to make sense of the notion that each part of the composite would be a part of the thought. Roderick Chisholm translates the passage in question as follows: "Suppose a compound thing were to think. Then every part of that compound would have a part of that thought. The thought that the compound would then have would be composed of the thoughts of the parts of that compound...." ("On the Simplicity of the Soul," *Philosophical Perspectives* 5, 1991, p. 175).
4. Immanuel Kant, *Critique of Pure Reason*, tr. N. K. Smith (New York: St. Martin's, 1965), p. 335 (A 352).
5. See Karl Ameriks, *Kant's Theory of Mind* (Oxford: Clarendon Press, 1982), Ch. 2.
6. *Op. cit.*, p. 376 (B 420).
7. *Kants gesammelte Schriften*, IV, p. 308; quoted by Henry E. Allison, "Kant's Refutation of Materialism," *The Monist* 79 (April, 1989), p. 195.
8. *Op. cit.*, p. 195. Even this may be overly simple, however. For a detailed discussion of Kant's view of the argument, see Ameriks.
9. See Wilson, "Leibniz and Materialism," pp. 509-512; Ameriks, *Kant's Theory of Mind*, pp. 55-64; Allison, "Kant's Refutation of Materialism," p. 18. Roderick Chisholm, who is not unsympathetic to the argument, nevertheless rejects it because, in his opinion, we lack a sense for the expression, "a part of a thought," and thus "have no reason to accept the statement" in Kant's argument which employs this expression (Chisholm, p. 175). But this is far from conclusive. If we think of the "thoughts" in question as propositions, then it seems rather natural (though it is not uncontroversial) to think of the concepts involved in the proposition as "parts" of the thought. But more importantly, the argument does not actually need to make use of the notion of a "part of a thought." What is needed is merely that a person's being aware of a complex cannot consist of the actions of parts of the person each of which is not aware of that complex.
10. Some of the material in the next few pages has been adapted from my article, "Emergentism," *Religious Studies* 18 (1982), pp. 480-484.
11. On a holographic film information from any point in the field of view is distributed all over the film, rather than being concentrated on a single point as in an ordinary photographic negative. If a part of a holographic film is destroyed no part of the field of view is lost; rather, the quality of the entire image is degraded to some extent.
12. The term "visual field," as here employed is perhaps not quite everyday usage, so some clarification may be in order. We certainly are familiar with the notion that each of us has at any given time a "field of vision" such that, for example, unobstructed objects in front of one are in one's field of vision but may leave the field of vision by moving behind one's back. The "visual field," as the term is here employed, is simply one's

field of vision insofar as one is aware of it at a given time, assuming one is not asleep, one's eyes are open and functioning, and so on.

13. I wish to emphasize that the "unity" in view here is the mere fact of being experienced simultaneously by a single subject; it does not involve any of the more sophisticated (and possibly debatable) ways in which some philosophers (e.g., Kant) have alleged that experience is and must be unified. One might also consider as a datum the "unified" experience of simultaneously feeling a tickle, hearing a twig snap, and smelling gasoline.

14. But might not the mental property I thus take myself to have turn out to be identical with some (perhaps rather complex) physical property? I am not at this point concerned to decide this question; I merely assert that I do in fact sometimes have the property of being aware of my visual field.

15. Does this mean that eliminativism and functionalism are rejected as theories of the mind? As noted, what has been said so far does not entail the denial that my being aware of my visual field has exactly the causal and functional properties attributed to it by functionalism, or that when I take myself to be aware of my visual field my brain is in exactly the physical state postulated by eliminativism. According to eliminativism, however, the statement "I am aware of my visual field," like other statements in what they (misleadingly) call "folk psychology," is not true. Since I believe this and similar statements are often true, I do reject eliminativism. It is not so clear that functionalism rejects commonplace assertions about the mental as false. What is clear is that on the view taken here the causal-functional story told by functionalism cannot be the whole truth about the mind. As Searle says, mental states have an "irreducibly subjective ontology."

16. A dualist, of course, would have a metaphysical reason for preferring the former locution. But in the present context we are assuming the perspective of the materialist, who would not have a reason of this kind.

17. But suppose the brain includes a number of parts, each of which contains all the information from the visual field? In this case we have, not a single V , but many V_i , many brain-units, each of the which might conceivably (according to the argument thus far) be aware of the entire visual field. Now, take any one these—say, V_n —and consider its composition. In the light of the computer analogy sketched above, we can say that V_n is a whole composed of physical parts. Many of these parts model information from various parts of the visual field. But no proper part of V_n models all of this information, so it is not possible for any of these parts to be aware of the entire visual field. But if V_n is a whole composed of parts each of which is not aware of the visual field, how can V_n itself be aware of it? And with this question, we return to the main argument; the postulation of multiple V_i doesn't change the situation at all.

18. If at this point we revert to saying that the person as a whole has this awareness, an awareness which is not possessed by any part of the person, we confront once again the truth that a person's being aware of a complex fact cannot consist in the actions of parts of the person, each of which does not possess this awareness.

19. And in Kant's argument: "the actions of a composite, or whatever inheres in it as thus composite, is an aggregate of several actions or accidents, distributed among the plurality of the substances."

20. "Philosophy and the Scientific Image of Man," in *Science, Perception and Reality* (London: Routledge & Kegan Paul, 1963), p. 27.

21. "Science, Sense Impressions, and Sensa," *Review of Metaphysics* 24 (1971), p. 411. Concerning the justification for this principle, Sellars wrote, "I have as yet published no...defence, but have simply stated that I find the

principle in accordance with my logical intuitions" (ibid., p. 393).

22. Ibid., p. 411.

23. "Philosophy and the Scientific Image of Man," p. 26. Two points require mention here. First, this example shows it is not a consequence of Sellars' principle that, in order for an object to be *F*, it must consist of parts which are *F*. This requirement would be satisfied in the case of the redness of the wall, but not in the case of the ladder's being a ladder. Second, it is clear that Sellars thinks of the object's identity as a ladder as determined by its physical structure. One might argue, however, that in order to be a ladder it must have a certain history (it must have been made in order to be a ladder, as opposed to originating by chance or for some other purpose) or that it must be used as a ladder. If these requirements do indeed obtain, then being a ladder would be a relational property involving the ladder's maker and/or user. For purposes of the present discussion, we shall ignore such historical and functional properties.

24. In saying this I commit myself to a notion of conceptual necessity (or analyticity) which is narrower than metaphysical necessity (and unlike the latter, determinable a priori) but does not require deducibility in a particular formalism, such as first-order logic. For a good recent defense of this notion, see Alan Sidelle, *Necessity, Essence, and Individuation* (Ithaca: Cornell University Press, 1989), especially chapter 5. (Sidelle's rebuttal of the arguments against analyticity in this chapter is independent of, though certainly consonant with, his endorsement of a conventionalist account of modality.)

25. This applies only to the monadic properties of *O*. For relational properties of *O*, one must consider the "system" made up of *O* and the other relata.

26. Admittedly, the task of giving a definition of the physical structure of a ladder—i.e., of stating the logically necessary conditions for an object to be a ladder—is not a trivial one. And any such definition will be infected with vagueness; there is not, for instance, a sharp line of demarcation separating ladders from stairways. But these problems do not invalidate the point being made here.

27. One can, to be sure, give "physical definition" of pink and other colors in terms of wave-lengths of light. But such definitions would not capture the experiential quality associated with "pink"; thus they are not definitions of color-terms as we ordinarily use them.

An interesting application of the principle of reducibility has been suggested to me by Philip Quinn. Water is, pretty clearly, a "system of objects," namely hydrogen and oxygen atoms. And water has the property of being liquid under standard conditions of temperature and pressure. But the predicate "*x* is a liquid" belongs to thermodynamics and does not occur in atomic theory; thus one cannot deduce, from the properties of hydrogen and oxygen atoms as described in atomic theory, that water is a liquid. So the principle of reducibility seems to be too strong.

What is at issue, however, is not what is logically derivable within a particular theory, but what is implied by the properties hydrogen and oxygen atoms actually possess. There is no reason whatever to doubt that the characteristics in virtue of which water is a liquid "consist of" various properties and relations of the water molecules (e.g., the lack of rigid bonding between molecules), and that a material whose molecules are characterized by these properties and relations necessarily behaves in such a way that it will be classified as a liquid.

28. But could we not replace the principle of reducibility, as here pro-

posed, with the weaker requirement that properties of systems should be supervenient upon the properties of their parts? Nothing that has been said so far is inconsistent with the possibility that mental properties may supervene upon physical properties. But if the supervenient property is "being aware of the visual field," the question to be asked is, What is this a property of? If it is a property of the biological organism, then exactly the same questions arise as have already been discussed—so property supervenience has no effect on the argument.

29. It might be suggested that the property of being aware of one's visual field is in fact identical with some complex physical property, and that this complex property is logically implied by the properties of the parts of the brain. This raises the question of criteria for property identity; if (as seems plausible) property identity requires logical equivalence, then mental properties in general will not be equivalent to physical properties. But suppose the requirement is weaker—say, metaphysical equivalence, coextensiveness in all possible worlds? There are of course objections (e.g., "multiple realizability") to supposing that even this weaker relationship holds in general between mental and physical properties. For present purposes, however, I make no use of such objections. But I do have specific objection in the case of "being aware of one's visual field" and similar properties, as follows: Being aware of one's visual field is essentially a unitary property, one which necessarily does not consist of and is not implied by properties of parts of a person each of which does not have the property in question. Any complex, on the other hand, would be essentially such that it necessarily consists of and is implied by the properties of the parts of the object which possesses it. So the mental property necessarily lacks a (second-order) property which the physical property necessarily possesses, and so they cannot be the same property.

30. Thus, Allison alleges that Kant's is "an argument from ignorance, in the sense that it appeals to an excessively narrow, essentially eighteenth century conception of scientific explanation" (Allison, p. 18; cf. also Wilson, p. 512).

31. For a discussion of Sellars' own way of dealing with this problem, see "Emergentism," pp. 484-86. Additional light is thrown on the unity-of-consciousness argument by the following quotations from Brentano and Chisholm. First, Brentano:

When someone thinks of and desires something, or when he thinks of several objects at the same time, he is conscious not only of different activities, but also of their simultaneity.... Now if we find the perception of seeing in one thing and the perception of hearing in another, in which of these things do we find the perception of their simultaneity? Obviously, in neither of them. It is clear, rather, that the inner cognition of one and the inner cognition of the other must belong to the same real unity (Franz Brentano, *Psychology from an Empirical Standpoint* [London: Routledge and Kegan Paul, 1973], p. 160, quoted in Roderick Chisholm, *The First Person: An Essay on Reference and Intentionality* [Minneapolis: University of Minnesota Press, 1981], p. 87).

Chisholm then comments as follows:

Brentano is telling us simply that, when a person is aware that he is seeing something and also aware that he is hearing something, then he is also aware that he is both hearing something and seeing something. Could we

perhaps settle for less—say that the person who sees something and hears something is aware that the seeing and the hearing are ‘parts of the same consciousness’ or that they are ‘co-present in consciousness’? I think not. What could it mean to say they are ‘parts of the same consciousness’ or ‘co-present in consciousness’ other than that the same person is aware of both? (Ibid., p. 88)

32. In Thomas Nagel, *Mortal Questions* (New York: Cambridge, 1979), pp. 181-95.

33. Ibid., p. 181.

34. Here I do not think Nagel is in contradiction with Searle’s contention that mental properties are physical properties. In Searle’s usage, mental properties are physical properties merely in virtue of being properties of physical organism. Nagel, on the other hand, is denying that they are physical properties in a sense “roughly equivalent to Feigl’s ‘physical2’” (ibid., p. 183n.).

35. Ibid., pp. 181-82. I have deviated from Nagel’s wording in some cases; he gives a short paragraph to explain each of the premises. But I am confident that these formulations capture his meaning.

36. Ibid., p. 182.

37. James Van Cleve, “Mind-Dust or Magic? Panpsychism Versus Emergence,” in *Philosophical Perspectives* 4 (1990), p. 217.

38. Ibid., p. 218.

39. Ibid.

40. For my own objections to dualism, see “Emergentism,” pp. 476-79; also, “The Souls of Beasts and Men,” *Religious Studies* 10 (1974), pp. 265-67.

41. I am indebted to a great many people who have offered comments and criticisms on earlier versions of this material, including Karl Ameriks, Robert Audi, David Burrell, Robin Collins, Scott Davison, Thomas Flint, Fred Freddoso, Stewart Goetz, Robert Kane, Bruce Langtry, Timothy O’Connor, Brad Petrie, Alvin Plantinga, Victor Reppert, William Rowe, Richard Swinburne, and Dean Zimmerman.