Nancey Murphy Profile

PHYSICALISM WITHOUT REDUCTIONISM: TOWARD A SCIENTIFICALLY, PHILOSOPHICALLY, AND THEOLOGICALLY SOUND PORTRAIT OF HUMAN NATURE

by Nancey Murphy

Abstract. This essay[1] addresses three problems facing a physicalist (as opposed to dualist) account of the person. First, how can such an account fail to be reductive if mental events are neurological events and such events are governed by natural laws? Answering this question requires a reexamination of the concept of *supervenience*. Second, what is the epistemological status of nonreductive physicalism? Recent philosophy of science can be used to argue that there is reasonable scientific evidence for physicalism. Third, the soul has traditionally been seen as that which enables human beings to relate to God. What accounts for this capacity in a physicalist theory of the person? This essay argues that the same faculties that enable higher cognitive and emotional experience also account for the capacity for religious experience.

Keywords: dualism; nonreductive physicalism; reductionism; religious experience; supervenience.

A significant point of intersection between theology and science is the debate over the "ontological constituents" of human beings. Although there have been debates on this topic for centuries in philosophy and for more than a century in biblical studies—dualism versus physicalism—

Nancey Murphy is Professor of Christian Philosophy at Fuller Theological Seminary, Pasadena, CA 91182. This article is adapted and reprinted from *Whatever Happened to the Soul?* ed. Warren S. Brown, Nancey Murphy, and H. Newton Malony, © 1998 Augsburg Fortress Press.

[*Zygon*, vol. 34, no. 4 (December 1999).]
 ISSN 0591-2385

recent developments in the neurosciences have brought these debates into public view.

In the ancient and medieval periods it was widely believed that all of the features distinguishing human beings from inanimate objects were to be accounted for as functions of the soul. Latin had two words that translate into English as "soul": *animus*, also translated "mind," and *anima*, also translated "life principle."

With the development of modern biology it is no longer assumed that anything nonphysical needs to be added to inorganic matter to produce a living organism; rather, life is a result of complex organization. Thus, insofar as "mind" is equivalent to *animus*, it may be thought appropriate that current philosophical discussions of the person concern the mind-body problem rather than the soul-body problem. However, the older concept of soul, even apart from its sense as "life principle," was broader than many contemporary concepts of mind. Contemporary views tend to focus on intellectual capacities, to the exclusion of the emotional and appetitive aspects of human life. In addition, in a variety of traditions the soul has been seen as the seat of the capacity for relating to God.

Modern mind-body dualism has been plagued with philosophical problems. These new problems have not been the result so much of changing conceptions of soul/mind as of changing conceptions of matter. For instance, in Aristotle's thought, souls were but instances of the metaphysical concept of *form*. Form was an essential constituent of all material entities, supplying active powers to otherwise passive matter. One might say that the entire metaphysical system was designed to accommodate the notion of soul.

With the rise of modern physics came a new conception of matter itself as no longer a principle correlative with form but as self-sufficient. Now the concept of form has no application, and souls or minds are anomalies in an otherwise purely physical and causally self-sufficient universe. Thus, most philosophers have come to see the problem of mind-body interaction as insuperable.

At the same time, astounding advances in the neurosciences have contributed greatly to physicalist accounts of mental and emotional phenomena. Many, however, are reluctant to accept physicalist accounts of the person, because these often seem to deny the existence, meaning, or value of those aspects of human life that we hold most dear. So the most pressing concern, I believe, is the question of whether a physicalist account of the person can avoid being reductive in a way that calls into question the meaningfulness of human existence. I argue that a proper understanding of *supervenient* levels of description avoids undesirable reductionist consequences.

A second philosophical issue is the epistemological status of nonreductive physicalism. If it is treated simply as a philosophical thesis, then it is diffi-

cult to say how it can be shown superior to any other philosophical view—for instance, no amount of empirical evidence can ever refute mind-body dualism. I propose that it be seen as the central organizing thesis of a broad research program (in the sense defined by philosopher of science Imre Lakatos) involving investigations in a variety of related areas: biology, neuropsychology, cognitive science.

Finally, I raise the question of whether accounts of religious experience will suffer from the replacement of older dualistic theories of the person. While this is not the only point at which accounts of human nature intersect with theology, it is certainly an important one, and one that is of equal interest for all religions, not just Christianity. I argue that religious experience can be understood, nonreductively, as supervenient on configurations of ordinary experiences, subserved by ordinary neural faculties.

Thus, I hope to make three contributions in this essay: (1) to explain how conscious states can be identified with physical states yet without ceding all causal agency to the purely physical level; (2) to shed some light on the sort of argument required to support a nonreductive physicalist account of human nature over its rivals, such as dualism and reductive materialism; and (3) to provide a nonreductive physicalist account of religious experiences.

Hierarchies and Reductionism

Westerners seem always to have attempted to understand the world in terms of hierarchies. The ancient Greeks thought of reality as a hierarchy of *beings* (Lovejoy [1936] 1960). A "generic" Greek view would go something like this:

divinities
(including heavenly bodies)
human beings
animals
plants
inanimate objects

During the modern period (beginning about 1600) a different understanding has gradually supplanted the Greek—that there exists not a hierarchy of beings but a hierarchy of *complex systems*. This hierarchy can be represented by a correlative hierarchy of the sciences that study reality in its varying levels of organization:

biology
chemistry
physics

Here physics is at the bottom because it studies the most basic constituents of reality; chemistry studies these "atoms" as they relate in complex structures (molecules)[2]; biology studies a number of levels of structure, from the biochemical level through the levels of organelles, cells, tissues, organs, and organisms, to colonies of organisms in their environments.

A contentious issue throughout the modern period has been whether psychology and the social sciences could be added in turn to this natural-science hierarchy—psychology being the study of the behavior of whole organisms and the social sciences being the study of human behavior in groups.

Another contentious issue has been reductionism. Here we need to distinguish among various sorts of reductionist theses. *Methodological reductionism* is a research strategy of analyzing the thing to be studied into its parts. *Causal reductionism* is the view that the behavior of the parts of a system (ultimately, the parts studied by subatomic physics) is determinative of the behavior of all higher-level entities. If this thesis—that all causation in the hierarchy is bottom-up—is true, it follows that the laws pertaining to higher sciences in the hierarchy should be reducible to the laws of physics.

Another sort of reductionism is the claim that higher-level entities are nothing but the sum of their parts. This thesis, however, is ambiguous; we need names here for two distinct positions. One is the view that, as one goes up the hierarchy of levels, no new kinds of metaphysical "ingredients" need be added to produce higher-level entities from lower-level ones. No additional "vital force" or "entelechy" is needed to produce living beings from nonliving materials; no immaterial mind or soul is needed to produce consciousness; no *Zeitgeist* is needed to form individuals into a society. Let us use the term *ontological reductionism* for this position. A still stronger claim than the previous one is that the higher-level entities are nothing but the sum of their parts but adds that only the entities at the lowest level are *really* real; higher-level entities—molecules, cells, organisms—are only composites of atoms. This thesis I here designate as *reductive materialism*. It is important to stress that it is possible to hold to ontological reductionism without subscribing to this thesis. Thus, one might say that higher-level entities, such as human beings, are real—as real as the entities that compose them—and at the same time reject all sorts of vitalism and dualism.

A variety of philosophers, biologists, and others have taken care to distinguish between these latter two theses. For example, "organicists" in biology rejected both vitalism and reductive materialism. The American philosopher Roy Wood Sellars developed a view of the entire hierarchy of the sciences that he called, variously, "emergent realism," "emergent naturalism," and "evolutionary naturalism." Sellars argued that organizations and wholes are genuinely significant; they are not mere aggregates of

elementary particles. Reductive materialism, he believed, overemphasizes the "stuff" in contrast to the organization. The levels Sellars countenanced were the inorganic, the organic, the mental or conscious, the social, the ethical, and the religious or spiritual (Sellars [1932] 1966; 1970).

So Sellars and a number of contemporary thinkers as well (Barbour 1990; Peacocke 1993) accept ontological reductionism but vehemently reject reductive materialism. In addition, they say that, while methodological reductionism has been a crucially important strategy in all the sciences, it is a limited strategy and needs to be balanced by studies of how entities at one level relate to higher levels—for example, organisms to their environments. These thinkers reject causal reductionism—one has to take account of causal influences of the whole on the part, as well as of the part on the whole. This process is referred to as "downward causation," "top-down causation," or "whole-part causation."

Let us use the term *nonreductive physicalism* to refer to this constellation of positions: the acceptance of ontological reductionism but the rejection of causal reductionism and reductive materialism. Applied to the specific area of studies of consciousness, nonreductive physicalism denies the existence of a nonmaterial entity, the mind (or soul), but does not deny the existence of consciousness (a position in philosophy of mind called "eliminative materialism") or the significance of conscious states or other ment*al* (note the adjectival form) phenomena. In brief, this is the view that the human nervous system, operating in concert with the rest of the body in its environment, is the seat of consciousness (and also of human spiritual or religious capacities). Consciousness and religious awareness are emergent properties, and they have top-down causal influence on the body. This is the view advocated here. As mentioned earlier, a number of philosophical issues need investigation in order to show that this position is coherent and intelligible. That is, can one consistently say that the neural system performs all of the functions once assigned to mind (and soul) and that this entails no significant loss to our understanding of human life? I believe that this general issue is best considered under the heading of causal reductionism: Is it possible to accept ontological reductionism without causal reductionism? I begin with this issue.

Defeating Causal Reductionism. The central question to be addressed here is, How can a physicalist account of the person *fail* to be reductive? The question of *causal* reduction seems to be the one that matters for retaining our traditional conceptions of personhood. There are several related issues.

First, if mental events can be reduced to brain events, and the brain events are governed by the laws of neurobiology (and ultimately by the laws of physics), in what sense can we say that human beings have free will? Are not their intendings and willings simply a product of blind physical

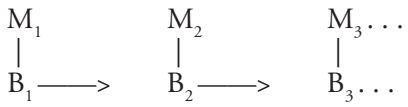
forces, and thus are not their willed actions merely the product of blind forces?

Second, if mental events are simply the product of neurobiological causes, what sense can we make of *reasons*? We give reasons for judgments in all areas of our intellectual lives—moral, aesthetic, scientific, mathematical. It seems utter nonsense to say that these judgments are merely the result of the "blind forces of nature."

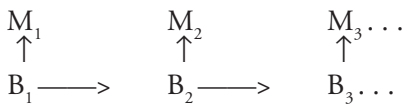
If free will is an illusion and the highest of human intellectual and cultural achievements can (*per impossible*) be counted as the mere outworking of the laws of physics, this is utterly devastating to our ordinary understanding of ourselves. It is equally devastating, of course, to theological accounts, which depend not only on a concept of responsibility before God but also on the justification (not merely the causation) of our theories about God and God's will. So, how is this unacceptable outcome of a physicalist account of the mental realm to be avoided?[3]

Supervenience. I claim that only with the assistance of recent conceptual developments can physicalist accounts of the mental avoid causal reductionism. A variety of strategies have been proposed for understanding the relation of mental events to brain events, and all run into difficulties: for dualists the problem is that of psychophysical interaction; for identity theorists the problem is that mental events become the result of neurobiological causes rather than of conscious reasons.

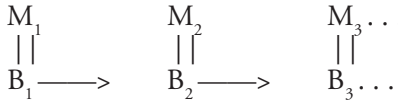
To see where the problem lies, let us begin with the vague thesis that every mental event (state, property) is related to some brain event. Add to this the assumption of causal connections among the neurological events, and we inevitably get a picture like the following, where M_1 through M_3 represent a temporal series of mental events, and the arrows represent causal connections among the brain events:



Until recently only two relations were conceivable between mental events and brain events—identity and causation[4]—so we can make the picture more specific in one of two ways (arrows represent causal relations, | | represents an identity relation):



or



In either case, if we assume causal connections at the physical level, causal reductionism seems inevitable. The mental events appear as mere epiphenomena.

In order to explain how reductionism can be avoided it is advantageous to consider the relation between consciousness and the neural system as but one instance of hierarchical ordering of complex systems, because we can see analogies and borrow concepts from less problematic levels. Recall that Sellars included both the conscious and the ethical as levels in the hierarchy of complex systems. In 1952 R. M. Hare introduced the term *supervenience* as a technical term to relate evaluative judgments (including ethical ones) to descriptive judgments:

> First, let us take that characteristic of "good" which has been called its supervenience. Suppose that we say, "St. Francis was a good man." It is logically impossible to say this and to maintain at the same time that there might have been another man *placed exactly in the same circumstances* as St. Francis, and who behaved in exactly the same way, but who differed from St. Francis in this respect only, that he was not a good man. (Hare 1952, 145; emphasis added)

So the higher-level property or description "good" *supervenes* on a collection of descriptions of Francis' character traits and actions. Or, to say the same thing, these character traits and actions *constitute* Francis' goodness.

In 1970 Donald Davidson introduced the concept of supervenience to describe the relation between mental and physical characteristics. Davidson describes the relation as follows:

> . . . mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition. . . . (Davidson 1980, 214)

The concept of supervenience is now widely used in philosophy of mind, but there is as yet no agreement on its proper definition. Terrence E. Horgan writes:

> The concept of supervenience, as a relation between properties, is essentially this: Properties of type A are supervenient on properties of type B if and only if two objects cannot differ with respect to their A-properties without also differing with respect to their B-properties. Properties that allegedly are supervenient on others are often called consequential properties, especially in ethics; the idea is that if something instantiates a moral property, then it does so in virtue of, i.e., as a (non-causal) *consequence of*, instantiating some lower-level property on which the moral property supervenes. (Horgan 1995, 778–79)

Notice that there are two distinguishable notions of supervenience in this passage. In the first sentence (substituting *S* for *A* for clarity, so that S-properties are supervenient and B-properties are subvenient or *base* properties):

1. Properties of type S are supervenient on properties of type B if and only if two objects cannot differ with respect to their S-properties without also differing with respect to their B-properties.

But from the last sentence we can construct the following definition:

2. Properties of type S are supervenient on properties of type B if and only if something instantiates S-properties in virtue of (as a noncausal consequence of) its instantiating some B-properties.

These two possible definitions are not equivalent; the first statement does not entail the second. The reason can be seen in Hare's original use of the term *supervenience*. Francis' character traits and actions (B-properties) only constitute him (or someone like him) a good person (an S-property) *under certain circumstances*. That is, it is conceivable that identical behavior in different circumstances would *not* constitute goodness. For example, we would evaluate Francis' life much differently if he had been married and the father of children.[5]

The difference between these two accounts of supervenience is absolutely crucial. If mental properties or events are supervenient in the first sense, this ensures the reducibility of the mental to the physical and raises all the problems mentioned above. If mental events or properties are supervenient only in the second sense, then, I claim, reduction is not a necessary consequence. Thus, I offer the following definitions (which I take to be equivalent):

3. Property S is supervenient on property B if and only if something instantiates S in virtue of (as a noncausal consequence of) its instantiating B under circumstance c.
4. Property S is supervenient on property B if and only if something's being B constitutes its being S under circumstance c.

An important feature of the supervenience relation, which has long been recognized, is that supervenient properties are often *multiply realizable*. This is a term from computer science that means that different configurations of hardware (vacuum tubes versus circuits) can realize, or constitute, the same machine considered at the functional level. So if S supervenes on B (given circumstance c), then something's being B entails its being S, but its being S does not entail its being B. For example, goodness is multiply realizable; many life patterns different from Francis' may also constitute one a good person. Thus, from the statement "R. M. Hare was a good man" we cannot infer that he lived as St. Francis did. This is one respect in

which supervenience relations fail to be identity relations (S <—> B): it is not the case that S —> B. (Here arrows represent entailment rather than causation.)

My definition of supervenience recognizes another way in which supervenience relations fall short of identity. The fact that S supervenes on B does not mean that B entails S (B —> S) because of the dependence upon circumstances. Under c_1, B —> S, but under c_2 it may be the case that B —> not-S. For example, under the circumstance of having a family to support, giving away all one's money might not constitute a good act.

For the purposes of clarifying the use of these terms, we need a very simple example, one without the additional complications associated with either the mind-brain issue or moral issues. Suppose that I have a light in my window and that I have arranged with a friend to use it as a signal to let her know whether or not I am at home: on means yes; off means no. I flip the switch; one state of affairs ensues, with two levels of description.

supervenient: The message is "I'm home."
subvenient: The light is on.

It is important to emphasize that there is one state of affairs, two descriptions. Turning the light on *constitutes* my sending the "at home" message under the circumstances of our having made the appropriate prior arrangement.

The "at home" message is multiply realizable. We could have agreed instead that I would leave the light off if I were home, or we could have agreed to use some other device altogether, such as leaving the window shade up or down.

We need a term to call our attention to an opposite sort of failure of the two descriptions to be identical. Not only is it the case that more than one subvenient state can realize the same supervenient state (light on, shade up), but also (again depending on circumstances), the same subvenient state can constitute more than one supervenient state. Suppose, for example, that we have agreed that the light's being on means that I am home only on Mondays, but on any other day the light's being on means that I am out. Depending on the circumstances of the day of the week, the same subvenient state constitutes either one message or the other. I suppose we could refer to this as "multiple supervenience."

It is this latter feature of the supervenience relation that I mean to highlight by emphasizing the role of *circumstances*. The variability in circumstances and their role in such cases is what makes for the difference between a supervenience relation and ordinary identity relations, and thus explains why some supervenient descriptions are not reducible to the lower level. This is the aspect that Horgan's first definition in the quotation above leaves out of account.

Let us now summarize the factors that distinguish between cases in which

reduction is and is not possible. The issues that matter are the following: (a) whether there are multiple circumstances such that B constitutes S in circumstance c but not in circumstance c′; and if so (b) whether or not c, c′, or other circumstances are describable at the subvenient level; (c) whether S is multiply realizable; and if so (d) whether there is a finite disjunctive set of realizands.

Reduction will be possible in the limiting case in which B constitutes S under all circumstances and S is not multiply realizable. (For a justification of these claims see Murphy in Russell, Stoeger, and Ayala 1998.) Reduction will not be possible (1) when there are multiple circumstances that make a difference to the supervenience relation and these circumstances cannot be defined in terms of the subvenient level; or (2) when S is multiply realizable and there is no finite disjunctive set of realizands.

The example above wherein the light's being on has opposite meanings on Monday and Tuesday is an example of the first type of nonreducibility: days of the week cannot be defined in the language of electrical phenomena.

For an example of the second type of nonreducibility, we cannot use agreed-on signals because this will necessarily be a finite list. Instead, consider the variety of natural signs or evidence of someone's being home: lights on, television set on, car in the garage, and so forth. Here no finite list of states of affairs constitutes the supervenient state "evidence of someone's being home"; thus, there can be no laws relating the two levels.

I emphasize that my conclusions here depend on using my more complex definition of supervenience, which gives due attention to circumstances. One might wonder about the disagreement over the definition of supervenience. (John R. Searle, in fact, claims that in philosophy of mind the supervenience relation is a *causal* relation, and he sharply distinguishes its use there from previous use in moral philosophy [1992, 124–26].)

It is important to recognize that many of the theorists working in this area are in favor of reductionism. Thus, the matter cannot be left to mere stipulation; we have to pursue the more difficult task of judging which definition better fits the facts: mine, in which circumstances at a higher level of description need to be taken into account, or the more common definition, in which they do not.

Many cases will fit the simpler definition of supervenience. However, there will also be a number of cases (perhaps most) in which only the more complex definition does justice to the phenomena. A clear case in which nonneural circumstances are widely recognized to make a difference is the role of *mental set* in perception. Two subjects induced to hold different expectations will often have different perceptual experiences resulting from the same physical stimulus. Mental set is a variable easily describable at the mental level, but in most (all?) cases will not be definable in terms of a finite set of neural realizands. For example, consider a well-known experiment in which subjects receive a small electrical shock on the back.

Depending on their mental set, they will experience the sensation either as burning heat or as icy cold. So at the subvenient level there is a series of physical events including the application of the shock, the transmission of a nerve impulse to the brain, and the set of brain events that realize the sensation of either heat or cold. The mental set will, of course, be realized neurobiologically, but it is multiply realizable: it can be the realization of a variety of perceptions of the environment (an ice-cube tray on the counter, burn ointment), or the result of statements by the experimenters, or any one of an unbounded set of other devices resulting in what we can only meaningfully describe at the mental level as the *expectation* of heat or of cold.

Another example: children asked to estimate the size of disks generally estimate coins to be larger than other disks of the same size. The concept of economic value simply does not translate to the neurobiological level of discourse.

This last example is important. If we take the hierarchy of levels to include the moral and the social (with its political, economic, and legal dimensions), we can see that a vast array of concepts will appear that most philosophers would agree is not logically reducible to neurobiological variables. These higher levels of reality are not in danger of causal reduction to the biological level. This, I claim, is exactly what is needed to protect traditional views of the meaningfulness of human intellectual endeavors. One way of putting this is to say that there are *emergent levels* as we go from the neurobiological to the cognitive, to the interpersonal, to the political, economic, and legal, to the moral, and finally to the spiritual. Although all human behavior supervenes on the biological (genetic and neurobiological), little of it is reducible to biology.

Free Will. It would be foolhardy to attempt to solve the problem of free will in one short section of one essay. However, the reflections in the previous subsection are certainly relevant to this issue. Clearly, if I have succeeded in defeating causal reductionism with regard to the mental and the neurobiological, the door is open to treatments of human freedom that do not depend on denying either physicalism or the law-governed character of neurobiological processes. That is, one of our strongest reasons for denying free will in the modern period has been the supposition that causal determinism applies to the human body.[6] Rebut this supposition, and the burden of proof shifts to those who would deny the freedom that seems an obvious fact of human experience.

Discussions of free will often distinguish between incompatibilist and compatibilist accounts. An incompatibilist view maintains that free will is incompatible with a determinist view of the natural world. A compatibilist view maintains that human freedom means being able to act as one chooses. Whether one's choices themselves can be shown to be a product of prior

causes of certain sorts is irrelevant. The important issue, it seems, is whether our choices are determined by the kinds of factors that we believe to be operative, or whether we are self-deceived. For example, is one's choice motivated by the reasons one gives, consistent with one's values, a true reflection of one's character; or is it instead, unbeknownst to the actor, a product of genetic predisposition, unconscious drives, or social manipulation?

The argument of the previous subsection is relevant here. In addition to the above list of suspicions, the physicalist has to answer the question of whether what appear to us to be reasoned choices are not actually the products of the laws of physics (with the laws of neurophysiology being but special cases). It was the intent of the previous argument to show that we can sometimes (and I would actually want to make the stronger claim—*usually*) make causal sense of a series of human actions only by attending to the mental-level description, which includes reasons, judgments, perceptions, and convictions. Yet this is compatible with causal determinism at the neurobiological level.

It is one thing to rebut determinist arguments; it is another to give a positive account of how free will is embodied in neurobiological functioning. My guess is that such an account will come from appreciating the multiple interacting layers of information processing in the brain.

Empirical Support for Nonreductive Physicalism

It has long been recognized that substance dualism cannot be disproved by empirical evidence. For example, no matter how much evidence accumulates suggesting that the brain performs mental operations (in Searle's pithy observation: "Scoop out the brain and the damned thing doesn't work"), it is still possible to claim that there is a substantial mind and that its operations are neatly *correlated* with brain events; Sir John Eccles, one of the most noted of neuroscientists, held exactly this view. It follows, then, that no amount of evidence from neuroscience can *prove* a physicalist view of the mental. This may seem a vexing state of affairs to philosophers who expect conclusive arguments, but most scientists are well aware that adequate evidence can be provided for a thesis without the evidence's ever amounting to proof.

I suggest, then, that we look at the epistemological status of nonreductive physicalism not as a philosophical thesis but as a scientific theory. Imre Lakatos has provided the most illuminating account to date of the structure of science (Lakatos 1978). Reacting to Thomas Kuhn's rather ambiguous account of the history of science as a series of *paradigms* (Kuhn [1962] 1970), Lakatos described it instead as a series of competing *research programs*. A research program is a vast network of theories, logically related to one another and supported by a variety of data. What unifies this network of theory and data is the "hard core"; that is, a thesis, often of a metaphysical nature, about the character of the part or aspect of reality

under investigation. A clear example of the role of metaphysics can be seen in the development of early modern physics, where atomism, one of the competing views of matter in ancient Greek *philosophy*, became the core of a very successful *scientific* research program. So I propose that we think of the metaphysical thesis of nonreductive physicalism as the hard core of a scientific research program.

A scientific research program also has a "positive heuristic," that is, a formal or tacit plan for development of the program that specifies the work to be done—the domain of phenomena that need to be explained according to the basic concepts and principles of the program. The positive heuristic in this case will be the plan to explain physicalistically all of the operations once attributed to the mind or soul. Insofar as researchers (in neurophysiology and anatomy, neuropsychology, psychiatry, cognitive science, and related fields) make progress in explaining "mental" phenomena, the program as a whole is making empirical progress, and its core thesis is thereby corroborated.

Great advances have been made in recent years in giving neurobiological accounts of these faculties. I find the results of brain localization studies to be some of the most impressive pieces of evidence for the physicalist program. Besides simply locating and modeling mental processes as previously understood, these studies sometimes improve our understanding of the mental processes themselves. For example, Antonio Damasio's account of patients with localized brain lesions that cause a combination of anhedonia and deficits in everyday decision making shows that, contrary to what has often been supposed, the emotions contribute positively to practical reasoning (Damasio 1994).

One area of brain research to which I shall attend in some detail is moral reasoning. I select this example because it is often considered one of the highest of human faculties. If progress can be made in explaining it neurobiologically (but without reducing it to mere biology), this will be a dramatic instance of empirical confirmation for nonreductive physicalism.

Here I follow Paul Churchland's work. Churchland and others are engaged in an attempt to supplant an earlier approach to computer modeling of neural processes (Churchland 1995; see also Edelman 1992). The earlier approach was algorithmic; that is, it attempted to model mental processes by writing rules to govern a linear sequence of transformations. This is the way most computers are programmed, but all computer users suspect that it is not the way human thinking works—computers are much better at some things than we (computation) but maddeningly worse at others (recognizing the same intention in the command "\n\ch6" as in "\n\ch-6").

An alternative model employs the notion of prototypes created by trial and error. For example, a computer can learn to distinguish an underwater mine from a rock if appropriate data are fed into it, and then the computer

repeatedly "guesses" and is informed after each trial whether its guess was correct. The hypothesis is that human brains work the same way. One of Churchland's examples is learning to recognize the taste of a peach. We have four types of taste receptors (sour, bitter, sweet, salty). On a four-dimensional graph we could represent a region within the space of all possible taste combinations that contains the exact combination of flavors of a number of peaches. The theory is that through repeated trials one develops a tendency to respond when a combination of signals from the tongue falls into that region. A strong "peach" signal will then be sent if the combination falls in the center of that region; weaker signals will be sent when combinations fall closer to the periphery (Churchland 1995, 21–24).

Now, what is the relevance of all of this to ethics? Churchland argues that much of moral or ethical learning is the development and refinement of prototypes; such learning is a process of discerning how to recognize and categorize a variety of social situations and to respond to them appropriately. For example, this kind of learning involves distinguishing lying from kidding and telling "white lies." Churchland writes: "the intellectual tradition of Western moral philosophy has focused on *rules*, on specific laws or principles. These are supposed to govern one's behavior, to the extent that one's behavior is moral at all" (1995, 144). Human capacities for moral reasoning, however, outpace philosophers' ability to identify the rules such reasoning follows, just as children's ability to speak grammatically precedes any knowledge of the rules of grammar.

> [I]t may be the case that a normal human's capacity for moral perception, cognition, deliberation, and action has rather less to do with rules, whether internal or external, than is commonly supposed.
>
> What is the alternative to a rule-based account of our moral capacity? The alternative is a hierarchy of learned prototypes, for both moral perception and moral behavior, prototypes embodied in the well-tuned configuration of a neural network's synaptic weights. (Churchland 1995, 144)

It is interesting to note that (quite independently of Churchland's work) many moral philosophers and theologians have made a significant turn away from rule-based analyses and toward approaches to morality that focus on virtues (prototypically good human qualities), recognizable only in narrative contexts. The topic of moral *description* has become central as well.[7]

So what does this mean for the topic at hand? One particular subsidiary research strategy within the broad research program of physicalism provides a competing account of the nature of moral reasoning (prototypes rather than rules). This thesis, if true, would explain in biological terms not only what is happening neurophysiologically when one engages in moral reasoning but also why the predominant rule-based strategy in modern Western ethics has turned out to be inadequate. It explains neurobiologically why an approach to moral analysis and moral education

based on narrative accounts of virtuous lives should be more effective than its competitor.

So far, I have not said anything about a nonreductive physicalist program that distinguishes it from a reductive materialist program. I shall not pursue that general question here, but it is clear that an important issue arises from Churchland's work. Is it adequate to say that moral reasoning is *nothing but* developing moral prototypes, and (as Churchland seems to assume) that moral motivation is *nothing but* the recognition that one gets along better in the social world by complying with moral expectations? Churchland writes: "From this perspective, the traditional question posed by the moral skeptic, namely, 'Why should I be moral?', looks peculiar and uncomprehending. As well ask, 'Why should I acquire the skills of swimming?' when one is a fish" (Churchland 1995, 150).

Owen Flanagan argues, rightly, that to reduce ethics to a combination of moral perception and prudence omits the crucial *normative* aspect of ethics. It disallows the question, *Should* it be the case that this society is such that one gets along better in it by conforming to prototype x? (Flanagan 1996, chap. 8). I would argue that to reduce the moral "ought" to a social-prudential "ought," or to biology, or to both, is to fail to understand the meaning of the moral "ought." It is a species of the incoherence into which reductive materialist accounts of the person regularly fall and against which I have attempted to guard in the previous section.[8]

So I conclude that if we take nonreductive physicalism to be not merely a philosophical thesis but also the hard core of a scientific research program, there is ample scientific evidence for it. It can be shown to be consistent with our everyday concepts of the significance of the mental but also confirmed by a burgeoning body of research showing that mental capabilities are realized neurobiologically. I also have hinted, in my critique of Churchland's view of ethics, that a nonreductive program would be more coherent and adequate to experience than a reductive materialist program. However, I cannot argue that here.

A Nonreductive-Physicalist Account of Religious Experience

Some earlier conceptions of the soul in the Christian tradition saw it as the means of contact with God; so the question arises, How are we to explain divine-human interaction using the resources of this new account of the person?

Philosopher of religion Carolyn Franks Davis has provided a useful list of kinds of religious experiences: interpretive, quasi-sensory, revelatory, regenerative, numinous, and mystical. I argue that these experiences supervene on combinations of ordinary experiences; that is, no special faculty is needed in order to experience religious realities. What makes the

experience religious is a meaningful combination of ordinary experiences, under *circumstances* that make it apparent that God is involved in the event in a special way.

My case is easiest to make for Franks Davis's first category. She writes:

> Sometimes a subject sees an experience as religious not because of any unusual features of the experience itself, but because it is viewed in the light of a prior religious interpretive framework. Common examples of such experiences are seeing a misfortune as the result of sins . . . , going through an illness with joy because it is a chance to 'participate in Christ's suffering,' experiencing love for all things of this world because of the belief that they are permeated by the divine, seeing an event as 'God's will,' and taking an event to be the answer to a prayer. (Franks Davis 1989, 33)

Here it is clear that no special faculty is needed to account for the religious experience. In Franks Davis's example the "misfortune" is experienced in the same way as any other event in human life. The higher-level description of "punishment from God" is seen as appropriate because of the circumstances: the prior sin, the belief that God chastises. Her second example is more interesting and more complex. Again, there is a lower-level description of an event experienced in the ordinary way: the person is ill. The sufferer's Christian worldview, however, allows for a higher-level description: participating in Christ's suffering. In addition, this higher-level perception is causally efficacious at the psychological level; in a top-down manner it affects the mood of the subject, producing joy where depression would otherwise be more likely. A more striking top-down effect is the enhancement of immune function that is a likely outcome of the elevated mood.

Franks Davis's second category is quasi-sensory experiences:

> Religious experiences in which the primary element is a physical sensation or whose alleged percept is of a type normally apprehended by one of the five sense modalities are 'quasi-sensory' experiences. These include visions and dreams, voices and other sounds, smells, tastes, the feeling of being touched, heat, pain, and the sensation of rising up (levitation). (Franks Davis 1989, 35–36)

The most common instances of this type discussed in the literature are visions. There are two ways to understand such experiences: (1) spiritual beings are really present in some way, and they are visible to the eye; or (2) the experience is akin to a hallucination but may be a genuine religious experience in that the vision was caused by God for some special purpose or at least in that the person derives some religious value from it. The second account seems to this author the more plausible. Here again no special faculty is required to understand it as a genuine religious phenomenon; presumably the same or similar neural capacities are involved as in hallucinating or dreaming.

This type of religious experience raises the issue of divine action; that is, the description of, say, "a vision of Christ" *justifiably* supervenes on the

description "experience of a man in white" under the circumstance of the experience's truly having been caused by Christ as opposed to, say, the mere wishing of the recipient or the effect of drugs. More on divine action follows.

Franks Davis's third category is revelatory experiences:

> Religious experiences of this category comprise what their subjects may call sudden convictions, inspiration, revelation, enlightenment, 'the mystical vision,' and flashes of insight. They may seem to descend upon the subject out of the blue, unaccompanied by any other feature which would make the experience religious, in which case it is their religious content which makes them 'religious experiences'; or, more frequently, they are the 'revelatory' element in a more complex religious experience, very often a mystical experience. These experiences have distinctive features: (i) they are usually sudden and of short duration, though the aftereffects may last a lifetime (especially in the case of conversion experiences); (ii) the alleged new knowledge seems to the subject to have been acquired immediately rather than through reasoning or sense perception; (iii) the alleged new knowledge usually seems to the subject to have been 'poured into' or 'showered upon' him (metaphors abound) by an external agency; (iv) the 'revelations' carry with them utter conviction, somehow even more than that which attaches to sense perception; and (v) the insights gained are often claimed to be impossible to put into words. (Franks Davis 1989, 39–40)

With the possible exception of "the mystical vision," these are clearly experiences that depend on the same neural functions as ordinary experiences. I want to emphasize the role of narrative context in justifying the description of these experiences in religious terms. Their happening "out of the blue" is often counted as a sign of divine action in the Christian tradition. For example, Ignatius of Loyola describes *consolation*, an experience confirmatory of God's action in a person's life, as an

> interior movement in the soul . . . through which the soul comes to be inflamed with love of its Creator and Lord; and when it can in consequence love no created thing on the face of the earth in itself, but in the Creator of them all.
>
> Likewise, when it sheds tears that move to love of its Lord, whether out of sorrow for one's sins, or for the Passion of Christ our Lord, or because of other things directly connected with His service and praise.
>
> Finally, I call consolation every increase of hope, faith and charity, and all interior joy which calls and attracts to heavenly things and to the salvation of one's soul, quieting it and giving it peace in its Creator and Lord. (Ignatius of Loyola 1978, 206)

Ignatius emphasized that consolation could be distinguished from more ordinary experiences partly by the fact that one had not done anything to induce it. This provides some grounds for believing it to have been effected by God. The long-lasting positive effects in the recipient's life that Franks Davis mentions add powerful confirmation to the judgment that these experiences are indeed revelations *from God*. In short, the church has developed *criteria* for discerning whether an experience is merely a human phenomenon or a true experience of or from God, and this discernment relies heavily on the narrative context of the event.[9]

Franks Davis writes that regenerative experiences are the most frequent type of religious experience among ordinary people. These are experiences that tend to renew the subject's faith and improve his or her spiritual, moral, physical, or psychological well-being. Again, the circumstances are the key: "This category includes a wide range of experiences: experiences of new hope, strength, comfort, peace, security, and joy, seen as 'religious' because they are obtained during religious activity such as prayer . . ." (Franks Davis 1989, 44–45).

Numinous experience has been defined by Rudolf Otto as a combination of awe, dread, or terror with a sense of attraction or fascination (Otto 1923). Here we have, again, ordinary human experiences, although in an unusual combination.

There is much disagreement about the nature of mystical experience. Franks Davis describes it as having the following features: (1) the sense of having apprehended ultimate reality; (2) the sense of freedom from the limitations of time, space, and the individual ego; (3) a sense of "oneness"; and (4) bliss or serenity (Franks Davis 1989, 54).

I shall not attempt an adequate account of these experiences but will only point out that qualitatively similar experiences have been reported by people taking psychoactive drugs. What then distinguishes the true mystical (that is, religious) experience from one phenomenally very similar is, as Franks Davis says, that "mystical experiences are usually . . . the pinnacle of the spiritual journey" (Franks Davis 1989, 55). Thus, it is the setting of the experience in the (often lifelong) quest of the recipient that distinguishes it as a religious experience.

In conclusion, then, I want to suggest that religious experiences do not depend on any special faculties over and above ordinary human emotional and cognitive faculties.[10] Their religiousness consists in (sometimes) their special content, but, more importantly, in their circumstances—circumstances that justify their being interpreted as acts of or encounters with the divine. In brief, religious experience *supervenes* on cognitive and/or affective experience in the *context* of an encounter with God.

Now, in the above account I have been assuming a view of divine action in which God is not only the creator and sustainer of the universe but also an agent in a special way in particular events. This is a common view of divine action in conservative Christian circles, but many liberal theologians would reject or seriously qualify the account of special divine acts. For instance, Maurice Wiles restricts God's action to enacting the whole of history. Revelation is not the result of special action on God's part but is to be explained in terms of special sensitivity of some people to God's general action (Wiles in Thomas 1983, 181–94). The reason for such accounts of divine action is largely that more robust accounts of special divine actions (whether they be of miracles or merely of special providence) have been made problematic by modern science. The same problems arise in at-

tempting to account for the action of a nonmaterial God as in attempting to explain how a nonmaterial mind could have a causal effect on the body.

The nonreductive physicalist account of religious experience is valuable in that it allows believers to accept and make use of research on the biological, psychological, and social realization of religious experience. Without an account of divine action, however, religious experience is reducible to these lower levels in the hierarchy. The nonreductive physicalist account of nature needs to be completed by a theological account in which descriptions of divine action supervene on descriptions of natural and historical events—without being reducible to them. We need to conceive of the hierarchy of the sciences as incomplete without theology and especially to maintain the nonreducibility of theology to other disciplines.[11]

Recognizing the role of Newtonian science in creating problems for an account of divine action, a number of theologians and scientists have called for reconsideration of the problem in light of more recent scientific developments.[12]

Conclusion

Nonreductive physicalism is an important new concept in the philosophical world. While dualism has gradually come to appear untenable in philosophical circles, most philosophers of mind have sought alternatives with a reductive intent. Recent developments in neurobiology and psychology have given aid and comfort to the reductionists. However, radical reductionism (reductive or eliminative materialism) is utterly unacceptable to the Christian. Thus, much needs to be done by scholars in a variety of fields to clarify nonreductive physicalism and to relate it to science, to discussions in fields such as ethics, and finally to the Christian tradition.

I have attempted to take a few steps in this direction, arguing for the coherence of a view that is ontologically, but not causally, reductionist and claiming that this avoids unacceptable consequences such as the denial of human freedom and the meaninglessness of the entire intellectual order. I also have provided a suggestion for how to view the epistemological relations between nonreductive physicalism and the accumulating scientific evidence. Finally, I have suggested a nonreductive physicalist account of religious experience.

Notes

1. This article is an adaptation of a chapter from Brown, Murphy, and Malony (1998), which is the product of a series of conferences on human nature—as seen from the point of view of philosophy, theology, biblical studies, biology, and the neurosciences—held at Fuller Theological Seminary and funded by the Templeton Foundation.

2. Of course this is an oversimplification: physics itself is now many-layered, and atoms as understood by chemists are no longer "atoms" in the philosophical sense of being the most basic constituents of matter.

3. A dualist account does not escape similar problems. Here the problem is also causal: How can a nonmaterial mind have any causal impact on the brain?
4. Mere correlation is also a possibility, but correlation entails all the problems of dualism.
5. A qualification needs to be added here. Someone who wanted to argue for the reducibility of supervenient properties in all cases would point out that anyone whose life was like Francis' in *all* (nonmoral) respects, including his relations to everyone else and everything else in the universe, would necessarily have the same moral properties. That is, even if moral properties do not supervene "locally" (in the first, stronger sense), it must be the case that moral properties supervene "globally" on nonmoral properties. We cannot imagine a possible world like this one in all nonmoral respects but differing only in moral respects. I believe that this claim about global supervenience is true but uninteresting for the issues at hand.
6. This supposition also has been a powerful motivator for dualist accounts of the person; the body may be caught up in the laws of Newtonian mechanics, but the mind is free.
7. In philosophical ethics Alasdair MacIntyre has done more than anyone else to encourage this shift; see *After Virtue: A Study in Moral Theology* ([1981] 1984). In theological ethics, Stanley Hauerwas has been most influential; Hauerwas has a series of books, beginning with one aptly titled *Vision and Virtue: Essays in Christian Ethical Reflection* (1974). See also McClendon (1986) for a "three stranded" analysis of Christian morality, which could nicely be interpreted as an account of three supervenient levels of moral reflection: the bodily and social levels, and the level pertaining to the community's ongoing relation with God.
8. I argue that the reduction of ethics to something else (pleasure calculus, reasonableness, prudence, and now biology, whether neurobiology or genetics) is a confusion, but an understandable confusion in our secular society. Ethics is intrinsically dependent on a higher (theological) level of analysis. See Murphy (1997, chap. 10); and Murphy and Ellis (1996, chaps. 6 and 8).
9. For an account of discernment practices, see Murphy (1990, chap. 5).
10. And perhaps in conjunction with a physiological component. "For example, a 'meditative' religious experience could include a cognitive/mental component, a behavioral component, and a physiological component (or components) all of which could be potentially identified, none of which is uniquely religious but in their clustering become a religious experience" (Dan Blazer, personal correspondence, 28 August 1996).
11. When lecturing on the topic of this essay I have often been surprised to find that some of the audience take the denial of the existence of a substantial soul to imply the denial of the existence of God. This is, emphatically, not my position. Christians need two basic metaphysical categories: God and creation. The claim that God's creation is purely physical does not entail there being no (nonphysical) creator.
12. See Peacocke (1993); Russell, Murphy, and Isham (1993); Russell, Murphy, and Peacocke (1995); and Russell, Stoeger, and Ayala (1998). Relations between nonreductive physicalism and divine action are directly addressed in the newest volume in this series, Russell, Murphy, Meyering, and Arbib (1999).

References

Barbour, Ian. 1990. *Religion in an Age of Science*. The Gifford Lectures, Volume One. San Francisco: Harper and Row.

Brown, Warren S., Nancey Murphy, and H. Newton Malony, eds. 1998. *Whatever Happened to the Soul? Scientific and Theological Portraits of Human Nature*. Minneapolis: Fortress Press.

Churchland, Paul M. 1995. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*. Cambridge: MIT Press.

Damasio, Antonio R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G. P. Putnam's Sons.

Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.

Edelman, Gerald M. 1992. *Bright Air, Brilliant Fire: On the Matter of the Mind*. New York: HarperCollins.

Flanagan, Owen. 1996. *Self Expressions: Mind, Morals, and the Meaning of Life*. New York: Oxford Univ. Press.

Franks Davis, Carolyn. 1989. *The Evidential Force of Religious Experience*. Oxford: Clarendon Press.
Hare, R. M. 1952. *The Language of Morals*. New York: Oxford Univ. Press.
Hauerwas, Stanley. 1974. *Vision and Virtue: Essays in Christian Ethical Reflection*. Notre Dame, Ind.: Fides Publishers.
Horgan, Terrence E. 1995. "Supervenience." In *Cambridge Dictionary of Philosophy*, ed. Robert Audi, 778-79. Cambridge: Cambridge Univ. Press.
Ignatius of Loyola. 1978. *The Spiritual Exercises: A Literal Translation and a Contemporary Reading*. Ed. and trans. David L. Fleming. St. Louis, Mo.: Institute of Jesuit Sources.
Kuhn, Thomas S. [1962] 1970. *The Structure of Scientific Revolutions*. 2d ed. Chicago: Univ. of Chicago Press.
Lakatos, Imre. 1978. "Falsification and the Methodology of Scientific Research Programmes." In *The Methodology of Scientific Research Programmes: Philosophical Papers, Volume I,* ed. John Worrall and Gregory Currie, 8–101. Cambridge: Cambridge Univ. Press.
Lovejoy, Arthur O. [1936] 1960. *The Great Chain of Being: A Study of the History of an Idea*. New York: Harper and Row.
MacIntyre, Alasdair. [1981] 1984. *After Virtue: A Study in Moral Theory*. 2d ed. Notre Dame, Ind.: Univ. of Notre Dame Press.
McClendon, James William Jr. 1986. *Ethics: Systematic Theology. Volume I.* Nashville: Abingdon Press.
Murphy, Nancey. 1990. *Theology in the Age of Scientific Reasoning*. Cornell Studies in the Philosophy of Religion, ed. William P. Alston. Ithaca, N.Y.: Cornell Univ. Press.
———. 1997. *Anglo-American Postmodernity: Philosophical Perspectives on Science, Religion, and Ethics*. Boulder, Colo.: Westview Press.
Murphy, Nancey, and George F. R. Ellis. 1996. *On the Moral Nature of the Universe: Theology, Cosmology, and Ethics*. Minneapolis: Fortress Press.
Otto, Rudolf. 1923. *The Idea of the Holy: An Inquiry into the Non-Rational Factor in the Idea of the Divine and Its Relation to the Rational*. Trans. and pref. by J. W. Harvey. London: Oxford Univ. Press.
Peacocke, Arthur R. 1993. *Theology for a Scientific Age: Being and Becoming—Natural, Divine, and Human*. 2d enlarged ed. Minneapolis: Fortress Press.
Russell, Robert J., Nancey Murphy, Theo C. Meyering, and Michael Arbib, eds. 1999. *Neuroscience and the Person: Theological Perspectives on Divine Action*. Vatican City State: Vatican Observatory, and Berkeley: Center for Theology and the Natural Sciences.
Russell, Robert J., Nancey Murphy, and Arthur R. Peacocke, eds. 1995. *Chaos and Complexity: Scientific Perspectives on Divine Action*. Vatican City State: Vatican Observatory, and Berkeley: Center for Theology and the Natural Sciences.
Russell, Robert J., Nancey Murphy, and C. J. Isham, eds. 1993. *Quantum Cosmology and the Laws of Nature: Scientific Perspectives on Divine Action*. Vatican City State: Vatican Observatory, and Berkeley: Center for Theology and the Natural Sciences.
Russell, Robert J., William R. Stoeger, and Francisco J. Ayala, eds. 1998. *Evolutionary and Molecular Biology: Scientific Perspectives on Divine Action*. Vatican City State: Vatican Observatory, and Berkeley: Center for Theology and the Natural Sciences.
Searle, John R. 1992. *The Rediscovery of the Mind*. Cambridge: MIT Press.
Sellars, Roy Wood. [1932] 1966. *The Philosophy of Physical Realism*. New York: Russell and Russell.
———. 1970. *Principles of Emergent Realism: The Philosophical Essays of Roy Wood Sellars*. Ed. W. Preston Warren. St. Louis, Mo.: Warren H. Green.
Wiles, Maurice. 1983. "Religious Authority and Divine Action." In *God's Activity in the World*, ed. Owen C. Thomas, 181–94. Chico, Calif.: Scholars Press.