5 An Old Ghost in a New Body

C. WADE SAVAGE

Descartes regarded a living human being as a machine inhabited by a person—a ghost, critics say—and he called this inner person "the soul." The operator of a crane is often used to illustrate his conception, but mistakenly. For the crane can do nothing without the help of its operator. The human machine, on the other hand, can do many things without the help of the inner person: It can digest food, breathe air, adjust its own temperature, and so on. Descartes believed that animals are mere machines, and therefore held that the human machine can do anything an animal (even an ape) can do without the help of the soul.

It is more accurate, therefore, to use the analogy of the pilot of a sophisticated robot, such as a spacecraft landing on the moon or a lunar exploration vehicle possessing legs, grasping claws, and sensors. The robot can do a great many things without the aid of its pilot: detect a rock of an unusual color, grasp the rock and place it in a storage compartment, and so on. But some things it cannot do unaided; for example, avoid a soft spot on the moon surface that looks solid, detect the unusual rock in deep shade, and so on. So the robot is designed, we will suppose, to transmit live pictures of its environment to the pilot's compartment, and with a set of controls that permit the pilot to take

C WADE SAVAGE Minnesota Center for Philosophy of Science, University of Minnesota

G. G. Globus et al. (eds.), *Consciousness and the Brain*

control of all or most of the craft's functions. The pilot cannot supply the robot with better sensory capacities, since he, like the robot, must rely on the signals that produce the television image; nor can he supply the robot with better motor operation, since with his controls he activates the same machinery the robot controls activate. But he can supply the robot with reasoning, or thought, which provides for a better analysis of the sensory data (for perception), better control of the motor operations (for decisions, and will); and—perhaps unfortunately for the robot—with the desires, purposes, and aversions that he possesses. The total system—robot plus pilot—thus becomes more intelligent and more subservient to the needs of persons. If we supply the robot with a loudspeaker connected to a microphone in the pilot's cabin, then the system acquires language—the physical expression of thought.

On Descartes' view, a human being is analogous to a robot-plus-pilot. But analogous only: for when we examine the interior of a human body we fail to find any pilot, any inner person. The reason, according to Descartes, is that the inner person is an immaterial, nonspatial agency—called the soul, or mind—which somehow acts upon and is acted upon by the body's nervous system. And here the view encountered its first difficulty. For Descartes was unable to explain to the satisfaction of his critics, or even to many of his disciples, how an immaterial, nonspatial system acts on a material, spatial system. So acute was the difficulty that Descartes' successors, while retaining the immaterial soul, abandoned the hypothesis that it acts on or is acted on by the body. These philosophers—the parallelists—held that the soul and body behave in parallel, *as if* the one acted on the other and the other on the one, when in fact there is no interaction. (When asked for an explanation of this parallel behavior, they often replied that it was arranged by God.) Thus the soul became useless in the explanation of behavior. If, as the parallelists would have it, the human body can behave as it does without being acted upon by or acting upon the soul, then all human behavior can be explained and understood without the hypothesis of the soul. It is, of course, a short step to materialism. If the hypothesis of the soul is unnecessary in the explanation of human behavior, why not abandon the troublesome hypothesis altogether?

There are difficulties in Descartes' hypothesis of the soul in addition to those arising from its immateriality and its alleged interaction with the body. As Descartes saw it, the human robot (body) can do much of what it does without the aid of its pilot (soul)—anything, in fact, that an ape can do, which is considerable. What he believed it not capable of, unaided by the soul, is conscious, rational, reflective (think-

ing), genuinely purposive, voluntary behavior—behavior that requires conscious perception, reasoning, thinking, formulating plans or goals, and consciously executing plans (willing). These are functions of the soul. Since an ape does not have a soul, it is not capable of the sort of behavior that requires these function. Now, this Cartesian argument can be reversed, as follows. Since an ape *is* capable of conscious, rational, reflective, purposive, voluntary behavior, the ape has a soul. The Cartesian may insist that apes are not capable of the behavior in question. But on what grounds? As every animal psychologist knows (or should know), apes exhibit the sort of behavior that in human beings is called conscious, rational, reflective, purposive, and voluntary. If the Cartesian replies that only behavior that requires a soul merits these labels, he begs the question. If he continues to insist that apes do not have souls, it can be pointed out to him that since conscious, rational, reflective, purposive, or voluntary behavior in an ape does not require a soul, such behavior in human beings does not require a soul either. Again, the hypothesis of the soul is seen to be unnecessary in explaining human behavior. Why not abandon it?

That is to say, why not abandon the hypothesis of an *immaterial* soul. Perhaps some part of the nervous system of both apes and humans generates conscious, rational, reflective, purposive, voluntary behavior. If so, we might call that part of the nervous system the *material* soul, since it performs the functions Descartes assigned to the immaterial soul. Whether there is any such part of the nervous system is, of course, a question to be decided by neurophysiologists, psychologists, and cyberneticians. Whatever they decide, we are safe in assuming, as materialists, that either some part of the body, or the body as a whole, performs the functions that Descartes assigned to the soul. We may call that part, or the body itself, the material soul. That humans have souls in this sense is the germ of truth in Descartes' view, and a minute germ it is.

In light of these familiar difficulties, it is a bit surprising to discover in our midst the ghost of Descartes, embodied in that intelligent machine known as Sir John Eccles. I would have supposed that three centuries of criticism, together with developments in animal psychology and computer science, had laid that spirit to rest. But perhaps my desire for progress in philosophy has blinded me to residual advantages in the Cartesian conception. In any case, Eccles' papers afford an opportunity to review the old issues with up-to-the-minute neurological flesh on their bones, and to see if we can learn anything new from them. Perhaps we shall learn how it is possible for

a contemporary neuroscientist to remain a Cartesian. And it should be noted that many contemporary scientists are Cartesians, though few as classically so as Eccles. Perhaps we shall also be able to test an intriguing hypothesis. Some philosophers believe that Descartes was at heart a mechanist; and that had he been nourished by contemporary scientific culture instead of 17th-century Christian culture, he would not have posited the soul, but would have held that humans as well as animals are simply machines. Professor Eccles may be arguing and theorizing as Descartes would were he among us.

INTERACTION

As we have seen, one major difficulty in the Cartesian view is that of explaining how an immaterial soul acts on and is acted on by a material body. Eccles believes that contemporary neurophysiology can contribute to the solution of this difficulty. Descartes' animal spirits and the nerves as conduits for these were abandoned long ago. The neuron with its axon and dendrites has been the standard building block of central nervous tissue for some time. More recently, nerve impulses have been considered to be neurochemical waves of disturbance. None of these developments has had the slightest bearing on the adequacy of Descartes' conception of the human being. The most significant recent development in neurophysiology, as Eccles sees the matter, is the hypothesis of the neural "column" as the functional unit of the cerebral cortex. These columns are clusters of interlaced neurons—pyramidal cells and others—positioned at right angles to the surface of the cortex; and it appears that, not individual neurons, but columns of these must discharge in order to activate the organism. Professor Eccles appear to believe that somehow such discoveries bring us closer to understanding how the soul acts on the body. As the sequel shows, this is simply wrong. The last 350 years of neurophysiological science have not brought us one bit closer to understanding the mystery.

Eccles diagrams the relation of the soul to the brain in Figure 4 (this volume, p. 114). His usual term for the soul is "the conscious self." (In the diagram it is called "World 2" for reasons that need not concern us here.) The downward arrows from the circle representing the conscious self represent volitions, or acts of will—the actions of the soul on the body. The upward arrows represent perceptions or consciousness of sensations—the actions of the body on the soul. We are given the following explanatory comment.

> It must be recognized that Figure 4 is an information-flow diagram and that the superior location adopted for the conscious self is for diagrammatic convenience. It is of course not meant to imply that the conscious self is hovering in space above the dominant hemisphere! It is postulated that in normal subjects activities in the minor hemisphere reach consciousness only after transmission to the dominant hemisphere, which very effectively occurs via the immense impulse traffic in the corpus callosum, as is illustrated in Figure 4 by the numerous arrows. Complementarily, . . it is postulated that the neural activities responsible for voluntary actions mediated by the pyramidal tracts normally are generated in the dominant hemisphere by some willed action of the conscious self (see downward arrows in Figure 4) (this volume, p. 113).

Here we have a contemporary neurophysiologist telling us where the soul acts on the body: not on the pineal gland, as Descartes sometimes suggested, but on a specific area of the cerebral cortex of the dominant hemisphere.

In an attempt to become more specific, he ventures the following hypothesis:

> . . . we have to assume that for "will" to be operative, large populations of cortical neurones are subjected to strong synaptic bombardments and are stimulated thereby to discharge impulses which bombard other neurones . . . at any instant the postulated action of the "will" on any one neurone would be effectively detected by the "critically poised neurones" on which it acts synaptically . . . in the active cerebral cortex the patterns of discharge of large numbers of neurones would rapidly be modified as a result of an "influence" that initially caused the discharge of merely one neurone. But further, if we assume that this "influence" is exerted not only at one node of the active network, but also over the whole field of nodes in some sort of spatiotemporal patterning, then it will be evident that potentially the network is capable of integrating the whole aggregate of "influences" to bring about some modification of its patterned activity, that otherwise would be determined by the pattern of afferent input and its own inherent structural and functional properties (1974, pp. 100–101).

The basic features of this theory can be briefly presented as follows. The primary activity of the brain is the action of *individual neurons* on other neurons through synaptic connections on the axons and dendrites of each. The secondary activity of the brain (at least the cortical areas) is the action of *columns of neurons* on other columns. The tertiary activity of the brain—and that which controls the organism by the producing motor impulses—is the action of *groups of columns of neurons* on other groups (in the manner depicted in this volume, p. 110 Figure 3). The resulting patterns, or waves, of neural activity lead

Eccles (following Sherrington) to liken the brain to an "enchanted loom, weaving a dissolving pattern, . . . a shifting harmony of sub-patterns" (this volume, p. 110), and to say that "it is absurd to state that the brain is as 'mechanical as clockwork' " (this volume, p. 103).

These remarks are misleading, for they can easily be taken to imply that, in some cases, the action of groups of neurons—the "patterned activity of the brain"—cannot be explained by means of the actions of the individual neurons comprising the group; and that the actions of the individual neurons must be explained, at least in part, by means of the action of the groups they are in. This is an emergentist position, whose leading exponent among neurophysiologists is R. W. Sperry.

> . . . consciousness is conceived to have a directive role in determining the flow pattern of cerebral excitation . . . conscious awareness . . . is interpreted to be a dynamic emergent property of cerebral excitation . . . the more molar conscious properties are seen to supersede the more elemental physio-chemical forces, just as the properties of the molecule supersede nuclear forces in chemical interactions . . . Individual nerve impulses . . . are simply carried along or shunted this way and that by the prevailing overall dynamics of the whole active process (in principle—just as drops of water are carried along by a local eddy in a stream or the way the molecules and atoms of a wheel are carried along when it rolls down a hill, regardless of whether the individual molecules and atoms happen to like it or not) (Sperry, 1969, pp. 533–534).

Sperry calls this position "emergent interactionism": "emergent" for reasons made clear in the passage above, "interactionism" because, while the emergent conscious properties direct individual nerve impulses, these properties are at the same time "directly dependent on the action of the component neural elements. Thus a mutual interdependence is recognized between the sustaining physico-chemical processes and the enveloping conscious qualities" (Sperry, 1969, p. 534).

Sperry may have developed this view in an attempt to solve the problem of mind–body interaction. If so, the attempt seems a failure. It is clear enough how the conscious, emergent properties depend on individual nerve impulses, for these properties are patterns of individual nerve impulses (such dependence is, incidentally, not *causal* dependence). But it is completely unclear how the conscious, emergent properties "direct" individual nerve impulses. Drops of water are not "carried along" by the local eddy, They and their actions on one another *constitute* the local eddy. However we assess Sperry's position, it is, apparently, not the position of Eccles. The emergentist

does not posit a soul, or conscious self, as an entity or system separate from the brain. For him consciousness is certain (emergent) brain properties.

Now the problem of how the will influences the behavior of the organism—how for instance, I will my finger to move—arises for Eccles in the following way. Behavioral activity is assumed to be caused by motor nerve activity and motor nerve activity by the activity of groups of neurons in the cortex. So if will is to influence behavioral activity, it must influence neuron-group activity. But it can only do this by influencing the activity of individual neurons in the group. To solve this problem, Eccles speculates that when the cortex is at a high general level of arousal, the firing of a single neuron can cause the discharge of several critically poised neurons, and each of these causes several more to discharge, and so on until a large group of neurons are discharging. In this way the will can, by acting on a single neuron, or on each of a few critically located neurons, influence neuron-group activity. The production of an avalanche provides an analogy for the mechanism posited here. A person on a mountaintop can, by releasing a single rock, or a few critically located rocks, produce an avalanche involving thousands of rocks. The person could in this manner cause (what is analogous to a behavioral act) the destruction of a building at the foot of a mountain. It should now be clear that Eccles is not an emergentist. In his view, the avalanche of falling rocks—or of neuronal discharges—is determined by its components, and not the other way round.

The problem of how the will influences behavior is thus reduced to the problem of how the will acts on individual neurons. To this latter problem Eccles offers not even the hint of a solution, and cannot offer a solution as long as he holds that the soul is an immaterial, nonspatial system. Does the soul act on a neuron in the way that one neuron acts on another? If so, the soul is something like a neuron—a material, spatial agency. Does the soul act on a neuron in the way that a cosmic ray might act on a neuron to cause it to discharge? If so, the soul is something like a cosmic ray (or the sun that emits the ray) and is a material, spatial entity. Does the soul act on a neuron in a manner quite unlike those mentioned? If so, the action of the soul is utterly mysterious and inexplicable. This is precisely the unsatisfactory state in which Descartes left the problem of interaction. For all his brandishing of contemporary neurophysiological theory, Eccles has made no contribution to its solution. He could make such a contribution by identifying the soul with some material system; for example, with the sun and its cosmic rays bombarding the brain. This would be a strange theory (my soul is the sun? and yours also?), and is probably objectionable

for a number of reasons, but it has no interaction problem. Or Eccles could identify the soul with some part or system of the brain, a theory that has no obvious defect in the present writer's opinion. Eccles' reasons for avoiding this theory will become clear in the sequel.

As long as the action of the soul on the body is mysterious and inexplicable—as it must be if the soul is an immaterial, nonspatial system—there is no need for Eccles to go to such trouble in speculating how the discharge of a few neurons (caused by the soul) can cause large collections of neurons to fire. He can simply say that the soul sometimes causes large collections of neurons to fire. The ability of the soul to cause one neuron to fire is no less mysterious than its ability to cause a collection of neurons to fire. This point shows quite clearly that Eccles "explanations" of interaction are pseudoexplanations. Since interaction between an immaterial soul and a material brain is mysterious and inexplicable, one may presume it to occur in whatever manner one wishes (consistent with the observable brain effects); which is equivalent to saying that one has no idea how it occurs.

The dualist may be tempted to argue that, since the immaterial soul is nonspatial, the problem of how it interacts with the brain is specious. The immaterial soul is not at any distance from the brain. It is just as correct to say that the immaterial soul is "in" individual neurons as it is to say that it is "outside" individual neurons, and just as correct to say that the soul is a "force internal" to the individual neuron as it is to say that it is a "force external" to the individual neuron. Suppose we say, then, that the soul is internal to the individual neuron on which it acts; then the problem of interaction disappears, does it not? It does not, because the solution is specious. The solution erroneously assumes that we can meaningfully say that the soul is inside, or internal to, something. If the soul is nonspatial, it makes as little sense to say that the soul is inside or internal to something as it does to say that soul is outside or external to something. But what then *is* the problem of interaction? How is the problem to be stated?

On the assumption that the soul is an immaterial, nonspatial entity, there is indeed no interaction problem of the kind we find in physics or physiology. If the neurophysiologist theorizes that nervous impulses are generated by neurons acting on neurons, he is required to say how this interaction occurs. The current explanation is that interaction consists in ion flows across membranes at the synapses of the interacting neurons. This is a physical explanation of physical interaction. There cannot be a physical explanation of soul–body interaction, since it is not a physical interaction; so it seems unfair to require the

dualist to provide a physical—hence inappropriate—explanation of interaction. However, it is entirely fair to require the dualist to provide an *appropriate* explanation—a nonphysical explanation—of the nonphysical interaction between soul and body. And we have no idea what such an explanation might consist in, for our concept of an explanation is that of a physical explanation. We cannot even *formulate* an explanation of soul–body interaction, much less *test* our formulation by experiments. The dualist can always say that this demonstrates the limitations of our conceptual system, the limitations of our concept of explanation. But as scientists we must surely regard this as obfuscation. If the hypothesis of an immaterial soul is, by definition, the hypothesis of an entity whose operations are inexplicable to human scientists, then human scientists should reject the hypothesis.

INTELLIGENT BEHAVIOR

A second major difficulty in the Cartesian view emerges if we ask why human behavior is supposed to require a soul when animal behavior does not. Eccles says:

> . . I think that there is no meaning or scientific value in ascribing mental properties to systems that exhibit order or apparent purpose or memory, or even intelligent action . . . Even when we come to the apparently intelligent actions of higher animals with their remarkable abilities to learn and remember, I have not found any reason to go beyond the purely mechanistic neurophysiology in explaining their brain performances, which of course was the position of Descartes (1974, pp. 87–88).

Then why go beyond the purely mechanistic neurophysiology in explaining the performances of *humans*? Eccles' answer can be divided into three parts: the hypothesis of the soul is held to be required to explain (a) intelligent behavior, (b) consciousness, and (c) free agency. We will discuss these parts of the answer in order listed.

Eccles holds that the hypothesis that the human body has a soul is required to explain certain intelligent human behavior; behavior that, however similar it may be to that of higher animals, is different enough to require a different explanation. Human linguistic behavior (language, we will say) is perhaps the best exemplar of the sort of behavior Eccles has in mind. (Human language has long been a principal refuge of the soul. For a retrospective of Descartes' treatment of language, see Gunderson, 1971.) The difficulty for Eccles is, of

course, that some animals exhibit linguistic behavior. He takes note of this fact and discounts it.

> Undoubtedly, the experimental investigations on chimpanzees, with respect to their developing both a sign language (Gardner and Gardner, 1969) and a symbol language (Premack, 1970), show that the chimpanzee brain exhibits considerable levels of intelligent and learned performance, but this chimpanzee communication is at quite a different level from human speech (1974, pp. 95–96).

Eccles will not find it difficult to discover differences between human and chimpanzee speech. Human speech is more creative (capable of greater novelty), communicative, comprehensive, etc.; and it contains symbols for numbers, quantities, abstractions, etc. However, he will find it difficult—impossible in the author's opinion—to discover differences that require us to posit a soul for the human but not for the chimpanzee. (Some linguists believe that human language requires the hypothesis of innate generating mechanisms—genetic wiring of the brain, so to speak. This, of course, is not the hypothesis of a soul, at least not an immaterial soul.)

The existence of animal language is no longer the major challenge to Cartesianism. The current focus is on computer language, about which Eccles says nothing in the papers under review. For several decades, computers have been capable of emitting language in the form of printed-out "answers" in a human script to "questions" fed into the computer in a special computer language into which the human operator translates his questions. In the early models, such responses exhibited no creativity, and could be made only to the stimuli of a computer language. But new models are being designed and built that can "talk" to human beings in a humanlike manner; i.e., that can respond directly to human language in language that exhibits the same sort of creativity exhibited by human language (see, e.g., Winograd, 1973). There is no good reason (none known to the author, at least) to doubt that one day we shall have a computer whose linguistic performance is indistinguishable from that of a human. The arrival of that day may not disabuse the Cartesian of his belief in the existence of the human soul. But it will establish that the soul is not required to explain such intelligent behavior as human language, and thus will deprive the Cartesian of a principal argument for the existence of the soul.

Observing the remarkable computer imagined above, Eccles would probably say: "But it does not think or reason, since it has no soul. Its language is mere imitation of human language, and, unlike the latter, is meaningless, nonreflective, and nonrational." One can predict

this response from statements such as: "One does not conduct a rational argument with a being who makes the claim that all its responses are reflexes, no matter how complex and subtle the conditioning" (this volume, p. 101). Now, we do not determine whether another human being can think or reason by inspection of its mental processes. Rather, we observe its behavior, including the linguistic; and if it behaves in a reflective, rational manner, we conclude that it thinks and reasons. If the computer behaves in a reflective, rational manner in its linguistic or other behavior, we will, if consistent, conclude that it thinks and reasons. As for meaning, it is not some "essence" or "life" breathed into behavior by a soul or other agency; it is a feature of the behavior. A meaningful utterance is one made in an appropriate way under appropriate circumstances. If the utterance of a human has these features, we pronounce it meaningful. If the utterance of a computer, or an animal, possesses these features, then, to be consistent, we must pronounce it meaningful.

Pressed by these arguments, the Cartesian may say of our talking computer: "Well, since it is capable of uniquely human behavior, behavior that requires a soul, the computer has a soul." (Apparently Eccles would *not* make this response.) One cannot help seeing this response as an *ad hoc,* desperate attempt to maintain the hypothesis of the soul. If the talking computer is endowed with a soul, who so endowed it? God, presumably, since human manufacturers apparently cannot create souls. But then we must suppose that God waits while the manufacturers build the computer and, at the moment of its completion, installs in it a soul. And why does He wait until the moment of completion? The computer is being built according to a certain design, and it is the design, rather than its realization in a particular collection of hardware, that seems to require soul. (Some thinkers have suggested that the soul of either a computer or a human being is the *program* of the machine. This conception seems too insubstantial and abstract to satisfy the Cartesian.) Given these theological problems, it seems the better part of wisdom to suppose that the computer is completely made by human manufacturers and is not equipped with a soul, especially since none is required to explain its behavior. And if none is required for the talking computer, none is required for the human being.

Descartes defined the soul as a thinking thing, and he believed that the soul is the *only* thing that thinks. But consider. Human linguistic behavior is generated by thought. This proposition follows from our ordinary definition of thought. For, as we ordinarily define it, thought is whatever generates such behavior as human language. Now

if computers can generate humanlike language, then computers can think; and if computers do not have a soul, computers can think without a soul. It follows that the soul is not the only thing that thinks. If computers can think without a soul, then so can the human body think without a soul. It can think by means of its brain alone, which can be regarded as a biological computer.

This conclusion seems forced upon us by biological and computer science. Eccles cannot remain a Cartesian and accept it. Yet he comes within a hair's breadth of doing just that.

> Since neural events in the minor hemisphere do not directly give the subject conscious experiences, we have to postulate that the neuronal machinery concerned in these specific operational tasks works at an unconscious level, which would be in good accord with the psychiatric concept of the unconscious mind. For example, in listening to music it can be envisaged that initially immense and complex operational tasks such as decoding, synthesizing and patterning are carried out in the temporal lobe of the minor hemisphere. Communication via the corpus callosum to the liaison areas of the dominant hemisphere with the consequent conscious experiences presumably is delayed until these most sophisticated neural operations have been carried out in the special musical centers. In their operational function these centers can be regarded as being analogous to the speech centers . . . (1974, p. 97).

The passage above illustrates a pervasive feature of Eccles' view of the conscious self, the tendency to suppose that the brain does all the work—analyzing, integrating, inferring, in short, thinking—and the conscious self sits back and "experiences" the results. This makes the soul a freeloader, something that does no work, and something that can be dispensed with.

If such tasks as the decoding, synthesizing, and patterning of music can be carried out by the cortex unconsciously—i.e., without the activity of the conscious self—then it is easy to suppose that thinking can be carried out by the cortex without the activity of the conscious self. For understanding language can be compared with understanding music, and understanding language is thinking. Why doesn't Eccles face it?The brain can think, and can do so without the soul. That the brain can think without the soul does not quite entail that the soul does not exist. But if the soul does exist, what is its function?

Does the soul think, as well as the brain? If not, then there is one less reason for positing its existence, since there is one less function it participates in or performs. If so, if both brain and soul are capable of thinking independently, then why the duplication of

function? On this view the soul is like a person equipped with a computer capable of performing operations of thought. Why would the person need the computer? Well, the computer can continue to work when the person becomes fatigued; and the computer may be able to think faster, or better than the person. The admission that the brain might think better than the soul does not make the dualist inconsistent, but it should make him uncomfortable. The more critical question for the dualist is, "Why does the computer need the person?" Well, the computer might think less well than the person in some areas (theoretical, or abstract areas, for instance); and the computer might not be as skillful or imaginative in programming computers as the person. But if, as seems likely from developments in computer science and neurophysiology, the brain can do everything in the way of cognitive processing a soul can do and do it just as well, then why retain the troublesome hypothesis of the soul?

CONSCIOUSNESS

Eccles second reason for positing the soul (conscious self) is that only the conscious self is capable of consciousness. He says: "By consciousness I mean conscious experience, which each of us has privately for himself" (1974, p. 87). What Eccles means by "consciousness" and "conscious" is the major difficulty in understanding and assessing this part of his argument for the conscious self, and the above definition does little to ameliorate the problem. What are "experiences"? Perceptions, feelings, volitions, beliefs, thoughts, decisions, purposes, and so on, presumably. Now if it makes sense to speak of conscious perceptions, feelings, etc., it must also make sense to speak of unconscious perceptions, feelings, etc. But does Eccles want to say that there can be unconscious thoughts, unconscious purposes, unconscious decisions? If so, then thinking, purposing, and deciding do not require a conscious self; and, therefore, there is no reason to deny that animals think, purpose, and decide. If, on the other hand, Eccles says that thoughts, purposes, and actions are necessarily conscious—cannot be unconscious—then his use of the term "conscious" is obscure. If I can be conscious of something (perception, thought, etc.), then I can fail to be conscious of it; i.e., it can be unconscious.

Most of Eccles' pronouncements about consciousness are made in the course of examining the "split-brain" experiments of Sperry and his associates, experiments with human subjects in whom the corpus callosum—the neural connecting system between the two

brain hemispheres—has been surgically interrupted. Such a subject is typically unaware of all perceptual inputs to the minor (right, usually) hemisphere—i.e., is not conscious of seeing objects in his left visual field; and he is not conscious of controlling his left hand. Of his left hand he says such things as, "I cannot work with that hand," "That hand is numb," and "I just can't feel anything or do anything with it." Nonetheless, the subject can carry out with his left hand intelligent, goal-directed, creative actions that require sensation and memory for their performance. For example, he may pick up a quarter from the table—no dollar notes being made available to him—when a dollar sign is flashed in his left visual field (a stimulus he is unaware of seeing). And he may be able to produce geometric drawings and mosaic constructions with his left hand that are superior to those he can produce with his right, of which latter he is conscious and does have conscious control (1974, p. 91).

Eccles concludes that the conscious self is directly connected, both in normal and in split-brain subjects, to the dominant hemisphere; and that the reason the split-brain subject is unconscious of actions and perceptions on his left is that the minor hemisphere has been disconnected from the conscious self.

> We can regard the minor hemisphere as having the status of a very superior animal brain. It displays intelligent reactions and primitive learning responses and it has a great many skills, particularly in the spatial and auditory domains, but it gives no conscious experience to the subject. Moreover, there is no evidence that this brain has some residual consciousness of its own (1974, p. 92).

It is tempting to clarify the terms "conscious" and "consciousness" as follows. That the split-brain subject is doing something with his left hand (drawing, for instance) is undeniable; but he is not conscious of what he is doing. Similarly, that he sees things in his left visual field—his left hand, the paper on which he is drawing—is undeniable; but he is not conscious of what he sees. He is conscious in the sense that he sees: we may call this consciousness$_1$. He is not conscious in the sense that he is not conscious of what he sees: we may call this consciousness$_2$. Consciousness$_2$ is a sort of *self*-consciousness: consciousness of the actions and perceptions of oneself. Better still, it is *meta*consciousness, at least as regards perception, since it is consciousness of consciousness. With this terminology, we can describe Eccles' position precisely. Consciousness$_1$ is possible without a conscious self, or soul. The minor hemisphere of the split-brain subject, and the brain

of an animal, are capable of consciousness in this sense. Such brains are not capable, however, of consciousness$_2$. Consciousness in this sense is a function of a conscious self. Only when the brain is connected to a conscious self, as it is in the normal human being, is it capable of consciousness$_2$.

The above clarification helps us understand why Eccles insists that consciousness cannot be a function of the brain, but must be a function of the conscious self. It is because he means by "consciousness" consciousness$_2$. He holds that consciousness$_1$ (seeing, for instance) is a function of the brain. But since consciousness$_2$ is consciousness of consciousness$_1$, it cannot be a function of the brain. For no material system, such as the brain, can be conscious of itself. Eccles does not explicitly make this argument in the papers under review, but it seems to be suggested by several things he does say, for example, "How can brain states describe themselves?" (1974, p. 88).

This is an ancient argument, and a fallacious one. In the first place, it is no more difficult for a material system to be conscious of itself that it is for an immaterial system (such as the conscious self) to be conscious of itself. Self-consciousness has seemed puzzling to many philosophers. Some have likened it to an eye seeing itself, and have argued that, just as an eye seeing itself is logically impossible, self-consciousness is logically impossible. But if self-consciousness is possible, as Eccles surely assumes, it is just as possible in a material system as it is in an immaterial one. To illustrate this point, consider the following, completely speculative, hypothesis. When a person sees a drawing, certain waves of neuronal discharge are generated in his cortex. Perhaps these waves have relatively low intensity when the subject is not conscious of seeing the drawing, and relatively high intensity when he is conscious of seeing the drawing. This is mere speculation and may be incompatible with current neurophysiological theory. But the general point is clear: some modification of the same brain processes that constitute consciousness$_1$ can be identified with consciousness$_2$. Which is to say that a material system can be conscious of itself.

In the second place, when something is conscious$_2$ of consciousness$_1$, it is not conscious of itself, for the two consciousnesses are different. Consequently, we can suppose that consciousness$_1$ is one brain function, and consciousness$_2$ another brain function that impinges on the first. Thus, consciousness$_2$ may be a system in the dominant hemisphere, and consciousness$_1$ another system in the minor hemisphere, on which consciousness$_2$ impinges in normal

subjects. In the split-brain subjects, the consciousness$_2$ system in the dominant hemisphere cannot impinge on the consciousness$_1$ system in the minor hemisphere, owing to interruption of the corpus callosum. There is no need to suppose that consciousness$_2$ is an immaterial system (the conscious self) separate from the brain.

The analysis above has been based on our distinction between consciousness$_1$ and consciousness$_2$, and the assumption that the split-brain subject is not conscious$_2$ of his perceptions and actions on his left. It has to be pointed out that this assumption seems false. If the subject is not conscious$_2$ of the movements of his hand and of seeing the marks on paper he makes as he draws, how does he accomplish this highly difficult task? Drawing requires continual monitoring of the marks made and the movements that make them, and this monitoring seems to be nothing other than consciousness$_2$. If this is correct, and if, as Eccles claims, consciousness$_2$ is a function of a conscious self, then a conscious self must be connected to the minor hemisphere of the split-brain subject, and to the brains of chimpanzees, who are capable of tasks as sophisticated as drawing. If, on the contrary, a conscious self is not required for consciousness$_2$ in these cases, then it is not required for consciousness$_2$ in the normal human subject. But it seems undeniable that there is a sense in which the split-brain subject is not conscious of his perceptions and actions. So it seems we must introduce another sense of consciousness, consciousness$_3$, with which to express this fact.

To understand this third sense of consciousness, we need to consider additional cases. We say of sleepwalkers, persons dazed from a blow to the head, and hypnotized persons that they are not conscious (or, not fully conscious). And we say of normal persons absorbed in some activity that they are not conscious (or, not fully conscious). What we mean by this is that they do not know what they are doing, or what their situation is. Consciousness$_3$ is consciousness of what one is doing and what one's situation is. It is self-consciousness, in the ordinary sense of that term. Our major test for whether a person knows what he is doing and what his situation is consists in getting him to *describe* what he is doing and what his situation is, either at some later time (from memory) or at the moment. Thus, to discover whether a person is walking in his sleep, we question him about what he is doing, where he is going, what his name is, where he lives, and so on. We use similar tests to discover whether a dazed boxer has emerged from his daze. (Note that the answers do not all have to be true for the subject to be conscious$_3$.)

We tend to regard a subject's description of what his situation

is and what he is doing as the *only* test of whether he is conscious$_3$. And since animals cannot provide such descriptions we conclude that they are not conscious$_3$, not self-conscious. (Eccles at least, seems to reach the conclusion in this manner.) But consider. If a dog is brought home from a long stay in the hospital, and immediately proceeds to search for familiar objects and places, then he knows what his situation is, and is conscious$_3$. If the dog is surprised in the act of eating a steak waiting to be broiled, and slinks away with ears down and tail between his legs, then he knows what he is doing, and is conscious$_3$. So if consciousness$_3$ requires a soul, dogs (some of them, at least) have souls.

Discussing what we have here called consciousness$_3$, Eccles says:

> . . . by forging linguistic communication of ever increasing precision and subtlety, man must gradually have become a self-conscious being aware of his own identity or selfhood. As a consequence he also became aware of death, as witnessed so frequently and vividly in other members of the tribal troup that he recognized as beings like himself (1974, pp. 102–103).

The penultimate developments are said to be those "in religion, in philosophy and in science that are associated with his attempts to understand the manner of being he was, his origin, and his destiny" (1974, p. 103). Most dogs have no awareness of death, at least, not their own death. And pretty obviously they make no attempts to understand the manner of being they are, their origin, and destiny. But it does not follow that they are not self-conscious. All that follows is that they do not have as high a degree of self-consciousness as do humans. They have a less complicated conscious self, or soul (and may be the better for it). Or, if Eccles denies that dogs have souls, to be consistent he must also deny that humans do. The relative complexity of human self-consciousness is not an adequate reason for positing souls in humans and none in dogs.

Finally, it should be noted that even if the only proper test for self-consciousness (consciousness$_3$) were the subject's *description* of what he is doing and what his situation is, it would follow that some animals are not self-conscious; but it would not follow that only humans have souls, and that a soul is required for self-consciousness. The talking computer imagined in the previous section is able to describe its situation ("I am a model X-001 computer. I was programmed in 1990.") and what it is doing ("I am now checking the last computation."). Either this self-conscious (conscious$_3$) computer has a soul, or the hypothesis of the soul is not required to explain self-consciousness. The latter is surely the more reasonable alternative.

FREE AGENCY

Eccles' third, and major, reason for positing the soul (conscious self) is that without it free will, or free agency, is inexplicable.

> . . . it is postulated that the neural activities responsible for voluntary actions mediated by the pyramidal tracts normally are generated in the dominant hemisphere by some willed action of the conscious self (see downward arrows in Figure 4). When destined for the left side, there is transmission to the minor hemisphere by the corpus callosum and so to the motor cortex of that hemisphere (this volume, p. 113).

In the split-brain subject, the corpus callosum has been interrupted, and this explains, according to Eccles, why the behavior of the subject's left hand is not free; the willed action of the conscious self cannot be transmitted to the minor hemisphere which controls the movement of the left hand.

It is worth noting, for its application to a later point, the controversial character of Eccles' interpretation. Eccles assumes that the subject's left-hand movements are unfree (involuntary), and explains this fact by the disconnection of the left hand from the conscious self. It is possible to assume that the subject's left-hand movements are free (voluntary) movements, even though the subject is not conscious of these movements, and then to explain this fact in one of two ways. The first is to hypothesize that there is a second conscious self connected to the minor hemisphere, but suppressed by the dominant conscious self and unable to communicate with the experimenter because of its lack of connection to any linguistic area of the brain. (Sperry proposes a hypothesis similar to this [1969, p. 532].) On the second explanation, there is only the one conscious self connected not only to the dominant hemisphere but also the minor hemisphere, and it wills the movements of the left hand without being conscious that it does so. Eccles may reply that one cannot will an action without being conscious that he does so. But this is doubtful. Both laymen and psychologists are accustomed to distinguishing between "consciously willed" and "unconsciously willed" actions, and Eccles himself sometimes employs the first of these two phrases.

Some dualists would argue that even if *every* action of a subject is like the actions of the left hand in the split-brain subject, still the human has free will and requires a conscious self. For as long as the conscious self of the human performs acts of will (represented by the downward arrows in Eccles' Figure 4), the human has free will, even if these acts of will are ineffective and do not cause neurons to discharge. This is a parallelist position, and not the position of the interactionist

Eccles, who says:

> If in willing an action one does not *effectively* influence the patterns of neuronal activity in the cerebral cortex and so bring about the desired discharge of motor pyramidal cells, then free will is an illusion, however subtle the philosophical arguments (this volume, p. 103).

It is, therefore, most accurate to say that Eccles' reason for positing the conscious self is to explain free *agency*, not free will.

This precision helps us to uncover fallacies in some of Eccles' arguments. He says: "That we have free will is a fact of experience" (this volume, p. 101), meaning, one would suppose, that since each of us has the experience of free will, each of us has free will, and therefore a conscious self that does the willing. This is a familiar and well-criticized argument. Unless freedom is a feeling, "I feel free" does not entail "I am free," any more than does, "I feel imprisoned" entail "I am imprisoned." And if freedom is just a feeling, it hardly seems worth having. Free *agency*, on the other hand—the ability to act freely—is worth having. But its existence cannot be established by introspecting our feelings, or experiences. "I feel that I am acting freely" does not entail "I am acting freely." Free agency is not a "fact of experience."

Because he fails to distinguish between free will and free agency, Eccles does not see that we can infer nothing about the existence or nature of free agency from "experience."

> It has been known for many years that electrical stimulation of the motor cortex of conscious subjects evokes actions which are disowned by the subject. As Penfield reports: "When a subject observes such an action, he remarks, 'That is due to something done to me and is not done by me.' " Evidently a motor action emanating from the motor cortex in response to voluntary command has some concomitants that are not present when a similar action is artificially evoked from the motor cortex (this volume, p. 118).

There is a sense in which Eccles' conclusion—stated in the last sentence—is trivially true. The concomitant of the voluntary action that is not present when a similar action is artificially evoked is Penfield's electrode! Of course, Eccles does not intend this trivial interpretation. Plainly he thinks that some important inference can be based on the subject's experience of unfreedom. This is an error. Suppose Penfield had been clever enough to evoke the subject's motor response together with the remark, "that is something done by me" (there is no reason to think this impossible). Should we then infer that the subject's action was free? Obviously not. That the subject *feels* free (or unfree) is poor evidence that he *is* free (or unfree).

This point can also be illustrated by the split-brain experiments of which Eccles makes so much. The split-brain subject says of his left hand, "I cannot work with that hand," "I just can't feel anything or do anything with it." And all the while the subject is drawing geometrical figures! Should we infer that the drawing is unfree? Plainly, we should not. The subject's actions (drawing) are remarkably similar to actions that in normal subjects we call "free," or, at least, "voluntary." It is therefore just as reasonable to infer that free or voluntary actions do not require an experience of freedom, and therefore do not require the operation of Eccles' conscious self.

If we should not infer that a subject's actions are free from the fact that he feels (experiences) them to be free, it is plainer still that we should not infer that they are free in Eccles' sense. Free action, according to Eccles, is action ultimately caused by acts of will on the part of an immaterial conscious self. Why an *immaterial* conscious self? To develop this query, consider the following simple modification of Eccles' diagram (Figure 4). The circle "hovering" above the dominant hemisphere is taken to represent another part (or system) of the *brain*, separately drawn for illustrative purposes. The downward arrows (as well as the upward) are taken to represent the action of neurons on other neurons; hence, acts of will are discharges of special neurons. These discharges are the "experiences of freedom" subjects often have. And the reason such experiences are sometimes deceptive is that the neuronal discharges are sometimes ineffective. In this model, the conscious self has become a material system. Why doesn't Eccles adopt this materialist model? Won't it explain free action as well as his immaterialist model?

The answer requires distinguishing two versions of the materialist model: the determinist and the indeterminist versions. The determinist holds that every event has a cause. Consequently, on the determinist version, volitional discharges of neurons are caused by the discharges of other neurons. Eccles is an indeterminist; he holds that some events do not have causes, namely, those acts of will that cause free action. So, he would reject the determinist version. But there is an indeterminist version of the materialist model, on which the volitional discharges of neurons are uncaused. A volition is, in this view, the *spontaneous* discharge of a neuron. Eccles would reject this view because it allows that some physical events—"volitional" neuronal discharges—are uncaused. Eccles wishes to remain a determinist in the physical–physiological sphere while holding that volitions are uncaused. To do so consistently, he must hold that volitions are nonphysical events and locate them in an immaterial soul (conscious self). It thus seems that he can have his cake and eat it.

But can he? The major motivation for determinism in the

physical sphere is our hope that physical events are predictable, hence manipulable, hence subject to rational control. If physical events are caused, they are predictable; if uncaused, then unpredictable. The spontaneous discharge of a neuron is uncaused and *completely* unpredictable. The discharge of a neuron which is caused by an uncaused mental volition is not completely unpredictable, since it can be predicted from the volition. But the discharge is *highly* unpredictable (as highly unpredictable as it can be without being completely unpredictable), since it cannot be predicted until the causing volition has occurred. On Eccles' indeterminist immaterialist model, neuronal events are only slightly less unpredictable than they are on the indeterminist materialist model; not enough, it would seem, to compensate for the disadvantages in the immaterialist model.

Immaterialism is therefore not required to explain (or allow for) free agency: one can hold that free actions are caused by uncaused physical events. But there is a deeper point to be made. Many philosophers believe that indeterminism is not required to explain (or allow for) free agency. This tradition was begun by Hume and reaffirmed by Schlick and the logical positivists. Such philosophers are called soft determinists, or compatibilists, since they hold that there is no incompatibility between the view that every event—every action included—has a cause and the view that some actions are free. Free action is defined by these philosophers as action with a special kind of cause. For one thing, the cause must be internal to the brain, so an action caused by an electrode touched to the cortex is unfree. Eccles makes vague objections to this view, but he fails to cite the most important one. How does the compatibilist distinguish between free and unfree actions where both are internally caused? The actions of a normal person are caused by neuronal discharges in his cerebral cortex; but so are the actions of an epileptic during a seizure. What justifies our calling the one type of action free, and the other type unfree? The compatibilists have never answered this objection. And their failure to do so is the most important argument in favor of indeterminism.

But the indeterminist view is no less objectionable. In contrast to compatibilism, it provides a clear definition of a free action; namely, a free action is one caused by uncaused volition. But it forces us to hold that if any action is free, then some events are uncaused. And it is an article of scientific faith that every event has a cause. To believe otherwise is to open the door to superstition and magic. This objection will not impress those who believe that science has limits and cannot enmesh everything in its causal web. For them, we provide a second, much more important objection.

On the indeterminist definition, a free action is one caused by an uncaused volition; whether the volition be identified as a spontaneous neuronal discharge, or a spontaneous mental act of will. The volition is not only not caused by any physical event, such as a neuronal discharge, it is not caused by any mental event either. Hence, it is not caused by a wish, desire, thought, purpose, plan, decision, or intention. The volition is therefore either an unpredictable, random event, or one that can be predicted only by statistical methods without certainty. Consider the flexing of my finger. If this action is free on the indeterminist definition, then I do not know when I will flex my finger. If the finger flexing is free, then I do not think about flexing it, nor do I wish, desire, plan, decide, or intend to flex it; or, if I do any of these, they have nothing to do with my flexing it. Free agency in this sense is of no use to me. Indeed, it is a liability, since free actions are, apparently, among the class of actions over which I have no control.

Compatibilism (soft determinism) seems to imply, by contrast, that I do have control over my free actions. For these are a subclass of the class of actions that are affected by my wishes, desires, thoughts, purposes, plans, decisions, and intentions. Since compatibilism also adheres to the scientific article of faith that every event has a cause, it seems much the more preferable view.

The above treatment of the advantages and disadvantages of determinism and indeterminism is standard among contemporary philosophers, but it is not conclusive. How is the issue finally to be resolved? Is neurophysiology of any help at all? Many contemporary philosophers believe that developments in neurophysiological science (and perhaps science in general) have no bearing on such large and enduring metaphysical issues as dualism versus materialism and freedom versus determinism. Eccles believes otherwise. There is at least one point where he seems to be right and the philosophers wrong.

Neurophysiological data do seem to have bearing on the truth or falsity of compatibilism. If the view that some actions are free and some unfree is compatible with the view that every event has a cause, then there must be some difference between the causal neurophysiological mechanism of actions we call free and that of the actions we call unfree. Neurophysiological evidence that there is no such difference is evidence that compatibilism is false. Neurophysiological evidence that there is such a difference is partial evidence that compatibilism is true. Partial, because the difference must be of the right sort, i.e., must correspond to the presence and absence of wishes, thought, plans, decisions, etc. Neurophysiologists may be developing evidence of the

latter sort—partial evidence that compatibilism is true. Curiously, this evidence is cited by the indeterminist Eccles in support of his own position.

Kornhuber and his associates recorded brain potentials in a subject who was instructed to flex his finger "at will." They discovered, over a wide area of the cerebral surface, "a slowly rising negative potential, called the *readiness potential*" which usually "began almost as long as 0.8s before the onset of the movement" (Figure 5, this volume p. 116). Eccles says:

> These experiments at least provide a partial answer to the question: What is happening in my brain at a time when a willed action is in process of being carried out? It can be presumed that during the readiness potential there is a developing specificity of the patterned impulse discharges in neurones so that eventually there are activated the correct motor cortical areas for bringing about the required movement. It can be regarded as the neuronal counterpart of the voluntary command. The surprising feature of the readiness potential is its wide extent and gradual build-up. Apparently, at the stage of willing a movement, there is a very wide influence on the patterns of neuronal operation, or, as we will consider below, on the patterns of module operation. Eventually this immense neuronal activity concentrates onto the pyramidal cells in the proper zones of the motor cortex for carrying out the required movement (this volume, pp. 115–116).

These experiments provide some evidence for the soft determinist, or compatibilist, view, on which free (voluntary) action is action with a special sort of cause. The special cause may be the readiness potential, the developing specificity of patterned impulse discharges, the concentration of neuronal activity onto the cells that produce the free movement.

The Kornhuber results are regarded by Eccles, however, as evidence for his own indeterminist view.

> My hypothesis would be that the highly specialized modules in the regions of the brain in liaison with the conscious self . . . can function as extremely sensitive detectors of consciously willed influences, at least when they are poised at special levels of activity . . . As a consequence, the willing of a movement produces the gradual evolution of neuronal responses over a wide area of frontal and parietal cortices of both sides, so giving the readiness potential. Furthermore, the mental act that we call willing must guide or mold this unimaginably complex neuronal performance of the liaison cortex so that eventually it "homes in" onto the appropriate modules of the motor cortex and brings about discharges of their motor pyramidal cells . . . (this volume, pp. 116–117).

It is just as reasonable to construe the Kornhuber results as evidence for

a materialist, determinist view as it is to construe them as evidence for Eccles' immaterialist, indeterminist view. There is no compelling reason to suppose, as Eccles does, that the readiness potential is caused by a "mental act of willing." Willing on the materialist view is some neural process, and we can suppose that this neural process generates the readiness potential. Perhaps the neural event of willing is the discharge of one, or a few, neurons. If so, willing may generate the readiness potential in the manner hypothesized by Eccles. The point is that willing need not be regarded as an immaterial, uncaused event. Eccles would undoubtedly respond that if willing is a neural event then it is caused by other neural events; which is contradictory, since willing is by definition an uncaused event. The reply of the compatibilist (soft determinist) is that this argument merely begs the question of whether willing can be identified with a neural process having a certain kind of cause.

When I freely flex my finger, I am the agent: "I do it." We often describe such actions by saying they are done "at will." Such ordinary descriptions lead Eccles to say that when a subject flexes his finger freely, his mental willing "guides" the neuronal discharges in his cortex so that they "home in" on those modules of the motor cortex whose discharge is required to actuate the muscles of the finger. Eccles is misled, according to the compatibilist. For the compatibilist, "doing it myself," or doing it "at will," is the effecting of the finger movement by a certain kind of cause—call it a "volition" or "act of will" if you must—which itself has a cause of a certain kind. The cause of my finger flexion is *my* neural activity, it is neural activity in *me*, and in that sense it is *I* who move my finger. Now, since some of my actions are caused by my neural activity and are nonetheless *un*free, the compatibilist must distinguish between those of my neural causes that effect free action and those that do not. The Kornhuber results may help to make this distinction. In any case, they provide just as much evidence for the compatibilist (soft determinist) view as they do for Eccles' indeterminist view.

To summarize this section, free action does not require the hypothesis of an immaterial, uncaused soul, and the experimental data do not support that hypothesis.

DISPROVING CARTESIANISM

In this paper it has been argued that all human behavior can be emitted by a body unaided by an immaterial soul, that the explana-

tion of human behavior does not require the hypothesis of an immaterial soul. It may be of value to see just how unpersuasive this argument can be to a committed Cartesian. The dualist points to some feature (creativity, for example) of a certain human behavior (speech, for example) and says: "There, you see? Only a being with a soul is capable of that sort of behavior." And the materialist replies: "No, such behavior can be generated by a sufficiently advanced mechanical brain." How can we resolve the dispute? By philosophical argumentation? By theoretical scientific considerations? By experiment? Suppose a friend and I encountered an unfamiliar machine picking up steel girders and riveting them into place as the superstructure of a building. My friend says: "It must have an operator; only a machine with a person controlling it is capable of doing things like that." And I reply: "No, a sufficiently advanced automaton could do just what that contraption is doing." How could the dispute be resolved?

Surely not by philosophical—i.e., *a priori*—argument. *A priori* argument can no more determine whether a crane-riveter automaton is possible than it could have determined whether the vacuum cleaner or washing machine is possible. Similarly, it would seem, *a priori* argument cannot determine whether a speaking automaton is possible.

Theoretical scientific considerations, on the other hand, are relevant. Suppose I could show my friend automata in operation that are rather like, in what they do, the crane-riveter contraption we encountered, and could produce a theoretical description (blueprint?) of an automaton that can do exactly what the crane-riveter does. This would afford considerable evidence that such automata are possible. Similarly, it would seem, if we could show the dualist a working automaton that emits language rather like human speech, and could produce a theoretical description (a computer program?) of an automaton that speaks exactly as humans do, then we would have produced considerable evidence that a speaking automaton is possible, and that speech does not require a soul. It would not be conclusive evidence, because the theory in whose terms the description of the automaton was given might be incorrect; it could be that on a correct theory the automaton is theoretically impossible.

A more expensive way of settling the dispute is by trying to construct an automaton that does what the crane-riveter does. If we succeed, then, by definition, we have shown that a crane-riveter automaton is possible. Note that we have not shown that the particular crane-riveter we encountered is an automaton: *that* one may have an operator. If we fail to construct a crane-riveter automaton, we have not thereby shown that such automata are impossible. There were, one would suppose, many unsuccessful attempts to construct an automatic

washing machine. Fortunately, the would-be inventors did not conclude that such automata were impossible. Similarly, if computer scientists succeed in constructing a speaking automaton (in "simulating human speech") they will, of course, have shown that speaking automata are possible, that speech can be emitted by devices without souls. Note that they will not have shown that human beings do not have souls. But they will have shown that it is not *necessary* to hypothesize the soul in explaining human speech. If computer scientists fail to construct speaking automata, they will not thereby have shown that such automata are impossible.

The direct method of settling the dispute is to discover whether the crane-riveter contraption has an operator, or is rather an automaton. If we discover that the contraption has no operator, then we have of course discovered that it is possible for such contraptions to have no operator, and the dispute is settled. But if we discover that the contraption has an operator, the dispute remains unsettled. For it may still be possible to build machines that do without an operator what this one does. Note that it may be quite difficult to determine whether the machine has an operator. That there is a person in the cab is not definitive; he may be a featherbedder who has absolutely nothing to do with the operation of the machine. That there is no person in the cab is also not definitive; the machine may be operated by remote control or radio. Similarly, if we discover that the human body has no soul, we will have discovered that human behavior can be emitted by an automaton. And if we discover that the human body has a soul, the question of whether human behavior can be emitted by an automaton will remain unsettled.

But how can we discover whether the human body has a soul? When we examine the interior of the human body, we find no operator. This in itself shows nothing, since the operator may be influencing the body from a distance by some radio-like process. Eccles makes precisely this suggestion when he says:

> I would postulate that in the liaison areas these neuronal patterns of module activity are the receiving stations or antennae for the ongoing operations in the consciousness of World 2 [the conscious self] as illustrated in Figure 4 (this volume, p. 117).

Therefore, to establish that the brain is not being influenced by a soul from a distance, we must show either that the body has no device for receiving such distant influences, or that such distant influences have no way of propagating themselves to the brain. Establishing either point is enormously difficult.

Proving that the brain is not influenced by a soul is rather like trying to prove that God does not intervene in the physical universe. To prove the latter conclusively we would have to show that every physical event is caused by some other physical event rather than by God. And how will we do that? It is practically impossible (logically impossible if the physical universe is infinitely large) simultaneously to examine every event in the physical universe. Similarly, to prove conclusively that the brain is not influenced by a soul, we would have to show that every brain event is caused by some other brain or physical event. We would have to show that the discharge of every neuron is caused either by the discharge of some other neuron or by some other physical event. And how will we do that? It is practically impossible—and probably always will be so—simultaneously to monitor the activity of every neuron.

Nevertheless, there comes a point at which the hypothesis that God intervenes in the physical universe, even though it cannot conclusively be disproved, becomes unreasonable. The hypothesis that the soul intervenes in the physical system called the brain becomes unreasonable at a much earlier stage of scientific investigation, since the brain is a finite, relatively small system. In the opinion of the author, that stage has been achieved. Chimpanzees are capable of language that is very like a primitive human language. Computers are being built that can respond to human language in humanlike language. Neurophysiological investigations have so far failed to uncover any neural events not caused by physical events. There is, in short, a mountain of evidence that all human and animal behavior is caused by neural events, and that every neural event is caused by some physical event. The evidence is not conclusive, but it rarely is even in the best of science. In the face of what evidence we do have, it is unscientific to believe that the soul acts on the body.

The same sort of evidence cannot be used to disprove the epiphenomenalist theory that the body acts on the soul, since the effects of this action are immaterial. Such effects leave the physical universe just as it is. Nor can this sort of evidence be used to disprove the parallelist theory that each human being has a soul that neither acts upon nor is acted upon by the body. It is not clear that the parallelist theory can be disproved by empirical evidence of *any* kind. (Whether it can is an unsettled question in philosophy.) But this issue is irrelevant to present concerns. Parallelism and epiphenomenalism are of no use to a physical scientist. Since, on these theories, the soul does not act on physical and biological systems, the soul cannot be employed as a hypothesis to explain the behavior or internal opera-

tions of any of these systems. If my soul does not act on my body, then it does not generate my intelligent behavior, nor my conscious behavior, nor my free (voluntary) behavior.

Interactionist dualism (the theory that the body and an immaterial soul act on one another) would be useful, but, as we have seen, it has no scientific basis. Eccles should confess to this, and cease practicing on us the delusion that neurophysiological investigations can confirm the existence of the soul.

Eccles occasionally betrays religious and political motivations for his dualism (not altogether surprising in light of the history of the position). For example, commenting on the views of B. F. Skinner, as contained in the latter's widely excoriated book, *Beyond Freedom and Dignity*, Eccles says:

> This type of behaviorism leads to a caricature of man—beyond freedom and dignity—that ignores the personal experiences that for each of us is the primary reality. It can appeal only to the philosophically naive and to those seeking the power that devolves from the absolute control of man (1974, pp. 89–90).

In these same pages we find, what we would expect, that Eccles rejects, not merely Skinnerian behaviorism, but all forms of behaviorism and materialism, including the recent psychoneural identity hypothesis.

Eccles seems to be echoing the old complaint that materialism degrades us by placing us on a level with animals and machines and plant and rocks. It degrades us only if we begin with a degraded conception of animals, machines, plants, and rocks. It is possible to be a materialist while holding that all life is sacred. It is even possible to be a materialist while holding the more profound view that all *being* is sacred (although this view is usually associated with monism). Given the political uses to which dualism has sometimes been put (justification of cruel treatment of animals, for instance) one might even suppose that materialism is the more humane view, in the broad sense of the term (as in "the humane society"). But this would be a mistake. Dualism is humane in this sense as long as it concedes that animals, too, have souls: feelings, desires, purposes, thoughts, consciousness, rights—the same rights to life and to absence of pain that we accord humans. Its refusal to make this concession seems, to this author, to be the product of human vanity.

REFERENCES

Eccles, J. C. (1974): Cerebral activity and consciousness. In: *Studies in the Philosophy of Biology: Reduction and Related Problems.* Ed. by F. J. Ayala and T. Dobzhansky, Berkeley: University of California Press.

Descartes, R (1967). *The Philosophical Works of Descartes* Vol 1 Ed by E S. Haldane and G R T Ross, New York: Dover Publications
Gunderson, K (1971): *Mentality and Machines.* Garden City: Doubleday
Sperry, R W (1969)· A modified concept of consciousness. *Psychol. Rev.* **76,** 532–536.
Winograd, T (1973) Research in natural language *Computer* **6,** 25–30