

A Psychiatric Dialogue on the Mind-Body Problem

Kenneth S. Kendler, M.D.

Of all the human professions, psychiatry is most centrally concerned with the relationship of mind and brain. In many clinical interactions, psychiatrists need to consider both subjective mental experiences and objective aspects of brain function. This article attempts to summarize, in the form of a dialogue between a philosophically informed attending psychiatrist and three residents, the major

philosophical positions on the mind-body problem. The positions reviewed include the following: substance dualism, property dualism, type identity, token identity, functionalism, eliminative materialism, and explanatory dualism. This essay seeks to provide a brief user-friendly introduction, from a psychiatric perspective, to current thinking about the mind-body problem.

(*Am J Psychiatry* 2001; 158:989–1000)

Of all the human professions, psychiatry, in its day-to-day work, is most concerned with the relationship of mind and brain. In a typical clinical interaction, psychiatrists are centrally concerned with both subjective, mental, first-person constructs and objective, third-person brain states. In such clinical interventions, the working psychiatrist traverses many times the “mind-brain” divide. We have tended to view etiologic theories of psychiatric disorders as either brain based (organic or biological) or mind based (functional or psychological). Our therapies are divided into those that impact largely on the mind (“psycho” therapies) and on the brain (“somatic” therapies). The division of the United States government that funds most research in psychiatry is termed the National Institute of “Mental” Health. The manual of the American Psychiatric Association that is widely used for the diagnosis of psychiatric disorders is called the Diagnostic and Statistical Manual of “Mental” Disorders.

Therefore, as a discipline, psychiatry should be deeply interested in the mind-body problem. However, although this is an active area of concern within philosophy and some parts of the neuroscientific community, it has been years since a review of these issues has appeared in a major Anglophonic psychiatric journal (although relatively recent articles by Kandel [1, 2] certainly touched on these issues). Almost certainly, part of the problem is terminology. Neither medical nor psychiatric training provides a good background for the conceptual and terminologic approach most frequently taken by those who write on the subject. In fact, training in biomedicine is likely to produce impatience with the philosophical discourse in this area.

The goal of this essay is to provide a selective primer for past and current perspectives on the mind-body problem. No attempt is made to be complete. Indeed, this article reflects several years of reading and musing by an active psy-

chiatric researcher and clinician without formal philosophical training.

Important progress has been recently made in our understanding of the phenomenon of consciousness (3). Some investigators have proposed general theories (e.g., Edelman and Tononi [4] and Damasio [5]), while others have explored the implications of specific neurologic conditions (e.g., “blindsight” [6] and “split-brain” [7]). Although this work is of substantial relevance to the mind-body problem, space constraints make it impossible to review all of this material here.

I take a time-honored approach in philosophy and review these issues as a dialogue on rounds between Teacher, a philosophically informed attending psychiatrist, and three residents: Doug, Mary, and Francine. These three have sympathies for the three major theories we will examine: dualism, materialism, and functionalism. Doug has just finished a detailed presentation about a patient, Mr. A, whom he admitted the previous night with major depression.

The Dialogue

Teacher: That was a nice case presentation, Doug. Can you summarize for us how you understand the causes of Mr. A's depression?

Doug: Sure. I think that both psychoanalytic and cognitive theories can be usefully applied. Mr. A has a lot of unresolved anger and competitiveness toward his father and this resulted—

Mary: Doug, come on! That is so old-fashioned. Psychiatry is applied neuroscience now. We shouldn't be talking about parent-child relationships or cognitive schemata but serotonergic dysfunction resulting in deficits in functional transmission at key mood centers in the limbic system.

Teacher: Mary, I'm glad you raised that point. Let's pursue this discussion further. Could Doug's view and your view of Mr. A's depression both be correct? Could his unresolved anger at his father or his self-derogatory cognitive schemata be expressed through dysfunction in his serotonin system?

Doug: I am not sure, Teacher. My approach to psychiatry has always been to try to understand how patients feel, to try to make sense of their problems from their own perspectives. People don't feel a dysfunctional serotonin receptor. They have conflicts, wishes, and fears. How can molecules and receptors have wishes or conflicts?

Mary: Wait a minute, Doug! Are you seriously claiming that there are aspects of mental functioning that cannot be due to brain processes? How else do you think we have thoughts or wishes or conflicts? These are all the result of synaptic firings in different parts of our brains.

Teacher: Let me push you a bit on that, Mary. What precisely do you think is the causal relationship between mind and brain?

Mary: I haven't thought much about this since college! I guess I have always thought that mind and brain were just different words for the same thing, one experienced from the inside—mind—and the other experienced from the outside—brain.

Teacher: Mary, you are not being very precise with your use of language. A moment ago you said that mind is the result of brain, that is, that synaptic firings are the cause of thoughts and feelings. Just now, you said that mind and brain are the same thing. Which is it?

Mary: I'm not sure. Can you help me understand the distinction?

Teacher: I can try. It's probably easiest to give examples of what philosophers would call identity relationships. Simply speaking, identity is self-sameness. The most straightforward and, some might say, trivial form of identity occurs when multiple names exist for the same entity. For example, "Samuel Clemens" is Mark Twain. Of more relevance to the mind-body problem is what have been called theoretical identities, identities revealed by scientists as they discover the way the world works. Theoretical identities take folk concepts and provide for them scientific explanations. Examples include the discoveries that temperature is the mean kinetic energy of molecules, that water is H₂O, and that lightning is a cloud-to-earth electrical discharge.

Mary: I think I get it. It wouldn't make any sense to say that molecular motion causes temperature or that cloud-to-earth electrical discharge causes lightning. Molecular energy just is temperature, and earth-to-cloud discharge of electricity just is lightning.

Teacher: Exactly! Now, let's get back to the question. Do mind and brain have a causal or an identity relationship? Having looked at identity relationships, let's ex-

plore the causal model. Am I correct, Mary, that you said that you think that abnormal serotonin function can cause symptoms of depression? Put more broadly, then, does brain cause mind? Does it only go in that direction?

Mary: Do you mean, can mind cause brain as well as brain cause mind?

Teacher: Precisely.

Doug: Wait. I'm lost. I'm still back on the problem of how what we call mind can possibly be the same as brain or even caused by brain. The mind and physical things just seem to be too different.

Francine: Me, too. I can't help but feel that Mary is barking up the wrong tree trying to see mind and brain as the same kind of thing.

Teacher: I can pick up on both strands of this conversation, as they both lead us back to Descartes (1596–1650), the great scientist and polymath who started modern discourse on the mind-body problem. Descartes agreed with you, Doug. He said that the universe could be divided into two completely different kinds of "stuff," material and mental. They differed in three key ways (8). Material things are spatial; they have a location and dimensions. Mental things do not. Material things have properties, like shape and color, that mental things do not. Finally, material things are public; they can be observed by anyone, whereas mental events are inherently private. They can be directly observed only by the individual in whose mind they occur.

Doug: Yes. That is exactly what I mean. The physical world has an up and down. Things have mass. But do thoughts have a direction? Can you weigh them?

Mary: Doug, do you realize how antiscientific you are sounding? How can you expect psychiatry to be accepted by the rest of medicine if you talk about psychiatric disorders as due to some ethereal nonphysical thing called mind or spirit?

Doug: However, maybe that is exactly what psychiatry should do, stand as a bastion of humanism against the overwhelming attacks of biological reductionism. Science is a wonderful and powerful tool, but it is not the answer to everything. Is science going to tell me why I find Mozart's music so lovely or the poetry of Wordsworth so moving?

Teacher: Hold on, you two. How about if we agree for now to ignore the problem of how psychiatry should relate to the rest of medicine? Let's get back to Descartes. He postulated what we would now call substance dualism, the theory that the universe contains two fundamentally different kinds of stuff: the mental and the material (or physical).

Mary: So, he rejected the idea of an identity relationship?

Teacher: Absolutely. But he had a big problem, and that is the problem of the apparent bidirectional causal relationship between mind and brain. Even in the 17th cen-

tury, they knew that damage to the brain could produce changes in mental functioning, so it appeared that brain influences mind. Furthermore, we all know what would happen if a mother were told that her young child had died. One would see trembling, weeping, and agitation—all very physical events. So it would also appear that mind influences brain. Descartes never successfully solved this problem. He came up with the rather unsatisfactory idea that somehow these two fundamentally different kinds of stuff met up in the pineal gland and there influenced each other.

Doug: But if mind and brain are completely different sorts of things, how could one ever affect the other?

Teacher: Precisely, Doug. That is one of the main reasons why Descartes's kind of dualism has not been very popular in recent years. It just seems too incredible.

Mary: So, are you saying that the identity relationship makes more sense because causal relationships are hard to imagine between things that are so different as mind and brain?

Teacher: That is only part of the problem, Mary. There is another that some consider even more serious. It is easy for us to understand what you might call brain-to-brain causation, that is, that different aspects of brain function influence other aspects. We are learning about these all the time, for example, our increasing insights into the biological basis of memory (9). Many of us can also begin to see how brain events can cause mind events. This is easiest to understand in the perceptual field—let's say vision. We know that applying small electrical currents to the primary and secondary sensory cortices produces the mental experience of perception.

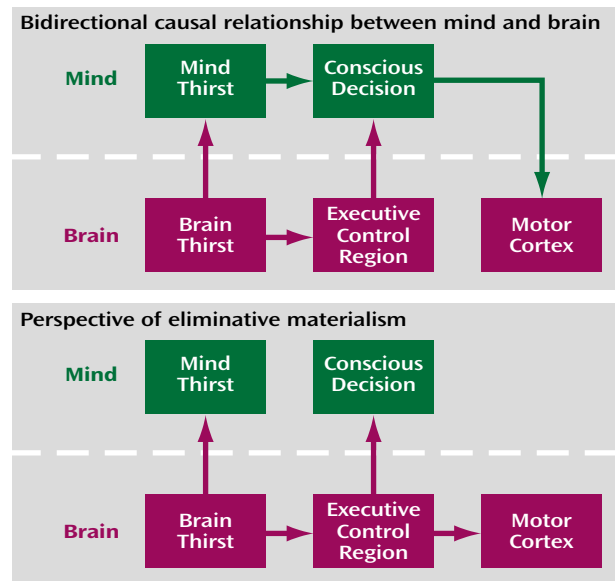
Mary: Those kinds of experiments sure sound like brain-to-mind causation to me.

Teacher: Yes. I agree that seems like the easiest way to understand them. But I am trying to get at another point. Let's set up a very simple thought experiment. Bill is sitting in a chair eating salty peanuts and reading the newspaper. He has poured himself a beer but is engrossed in an interesting story. In the middle of the story, he feels thirsty, stops reading, and reaches for the glass of beer.

Now, the key question is, what role did the subjective sense of thirst play in this little story? Was it in the causal pathway of events? I am going to have to go to the blackboard here and draw out two versions of this (Figure 1, upper part).

Both versions begin with a set of neurons in Bill's hypothalamus noting that the sodium concentration is rising owing to all those salty peanuts! We will call that brain thirst. The interactionist model assumes that the hypothalamus sends signals to some executive control system (probably involving a network of several structures, including the frontal cortex). Somewhere in that process, brain thirst becomes mind thirst; that is, Bill

FIGURE 1. Two Views of the Causal Relationship Between Mind and Brain in the Experience of Thirst and Act of Reaching for a Beer^a



^a In the view shown in the top part of the figure, which depicts a bidirectional causal relationship between mind and brain, the critical decision to reach for beer because of thirst is made consciously in the mind; the decision is conveyed to the motor cortex for implementation. From the perspective of eliminative materialism, all causal arrows flow within the brain. The mind is informed of brain processes but has no causal efficacy. Conscious decisions apparently made by the mind have, in fact, been previously made by the brain.

has the subjective experience of thirst. In his executive control region, then, based on this mind thirst and his memory about the nearby glass of beer, Bill makes the mind-based decision, "I'm thirsty and that beer would sure taste good." The decision (in mind) now being made, the executive control region, under the control of mind, sends signals to the motor strip and cerebellum saying, "Reach for that glass of beer."

The main advantage of this little story is that it maps well onto our subjective experience. If you asked Bill what happened, he would say, "I was thirsty and wanted the beer." He would be clear that it was a volitional act in his mental sphere that made him reach for the beer. But notice in Figure 1 that we are right back into the problem confronted by Descartes. This little story has the causal arrows going from brain to mind and back to brain. A lot of people have trouble with this.

An alternative approach to this problem is offered by the theory of eliminative materialism or, as it is sometimes called, epiphenomenalism.

Doug: Ugh. It is these kinds of big words that always turn me off to philosophy.

Teacher: Bear with me, Doug. It might be worth it. Philosophy certainly does have its terminology, and it can get pretty thick at times. But then again, so do medicine and psychiatry. The concept of eliminative materialism is

very simple: that *the sufficient cause for all material events is other material events*. If we were to retell this story from the perspective of eliminative materialism, it would look a lot simpler. All the causal arrows flow between brain states, from the hypothalamus to the frontal cortex to the motor strip (Figure 1, lower part).

Doug: But what about Bill's subjective experience of thirst and of deciding to reach for the glass of beer?

Teacher: The theory of eliminative materialism wouldn't deny that Bill experienced that but would maintain that none of those mental states was in the causal loop. The hypothalamic signal might enter consciousness as the feeling of thirst, and the work of the frontal control center might enter consciousness as the sense of having made a decision to reach for the glass, but in fact the mental experiences are all epiphenomenal or, as some say, inert. Let me state this clearly. The eliminative materialism theory assumes a one-way casual relation between brain and mind states (brain → mind) and no causal efficacy for mind; that is, there is no mind → brain causality. According to this theory, mind is just froth on the wave or the steam coming out of a steam engine. Mind is a shadow theater that keeps us amused and thinking (incorrectly) that our consciousness really controls things.

Doug: That is a pretty grim view of the human condition. Why should we believe anything that radical?

Teacher: Doug, have you ever had the experience of touching something hot like a stove and withdrawing your hand and *then* experiencing the heat and pain?

Doug: Yes.

Teacher: If you think about it, that experience is exactly what is predicted by the eliminative materialist theory. Your nervous system sensed the heat, sent a signal to move your hand, and then, by the way, decided to inform your consciousness of what had happened after the fact.

Doug: OK. I'll accept that. But that is just a reflex arc, probably working in my spinal cord. Are you arguing that that is a general model of brain action?

Teacher: I'm only suggesting that it needs to be taken seriously as a theory. In famous experiments conducted in the early 1980s (10), the neurophysiologist Ben Libet asked students to perform spontaneously a simple motor task: to lift a finger. He found that although the students became aware of the impulse around 200 msec before performing the motor act, EEG recordings showed that the brain was planning the task 500 msec before it occurred. Now, there are a lot of questions about the interpretation of this study, but one way to see it, as predicted by eliminative materialism, is that the brain makes up its mind to do something, and then the decision enters consciousness. The mind thinks that it

really made the decision, but it was actually dictated by the brain 300 msec before. My point with this story and the reference to Libet's work is to raise the question of whether consciousness is as central in brain processes as many of us like to think.

Let me try one more approach to advocating the eliminative materialist perspective. Before the development of modern science, people had many folk conceptions that have since been proven incorrect, such as that thunder is caused by an angry god and that certain diseases are caused by witches. Perhaps the concepts of what has been called folk psychology are the same sort of thing, superstitious beliefs that arose when we didn't know anything about how the brain worked.

Doug: You mean that believing that our actions are governed by our beliefs and desires is like believing in witches or tribal gods?

Mary: To put it in another way, magnetic resonance imaging (MRI) machines that show the brain basis of perception and cognition are like Ben Franklin's discovery that lightning wasn't caused by Zeus but could be explained as a form of electricity.

Teacher: Exactly. So, we have now examined some of the problems you get into when you want to try to work out a causal relationship between mind and brain.

Mary: Yeah. That identity theory is looking better and better. After all, it is so simple and elegant that it must be true. Mind is brain and brain is mind.

Teacher: Not so fast, Mary. There are some problems with this theory too.

Francine: Teacher, isn't one of the biggest problems with identity theories that they assume that the mind is a thing rather than a process?

Teacher: Yes, Francine. Let's get back to that point in a few minutes. I agree with Mary that identity theories are very appealing in their simplicity and potential power, but, unfortunately, they are not without their problems. Let's review three of them. The first stems from what has been called Leibniz's law, which specifies a critical feature of an identity relationship. This law simply says that if an identity relationship is true between A and B, then A and B must have all the same properties or characteristics. If there is a property possessed by A and not by B, then A and B cannot have an identity relationship.

Mary: That makes sense, and it certainly works for the examples you gave. Whatever is true for water must be true for H₂O, and the same goes for lightning and a cloud-to-earth electrical discharge.

Teacher: Yes, but what about mind and brain? As Doug said, the brain is physical and has direction, mass, and temperature, whereas the mind has wishes, intentions, and fears. Can two such different things as mind and brain really have an identity relationship?

Doug: Yes. That is exactly what I have been trying to say. That puts a big hole in your mind-equals-brain and brain-equals-mind ideas, Mary!

Mary: Maybe. But isn't that a pretty narrow view of identity? Perhaps the problem of mind and brain is not like that of lightning and electricity. It might be that mind and brain are one thing, but when you experience it from the inside (as mind) and then from the outside (as brain), it is unrealistic to expect them to appear the same and to have all the same properties.

Teacher: A good point, Mary. What you have raised is the possibility of another kind of dualism that is less radical than the kind proposed by Descartes. One way to think about it is that there are two levels to what we might call identity. Things can be identical at the level of substance and/or at the level of property. Imagine that brain and mind are the same substance but have two fundamentally different sets of properties.

Mary: Please give me an example. I am having a hard time grasping this.

Teacher: Sure. If we take any object, it will have several distinct sets of properties: mass, volume, and color, for instance.

Mary: So mind and brain are two distinct properties of the same substance?

Teacher: Precisely. Not surprisingly, this is called property dualism and is considerably more popular today in philosophical circles than is the substance dualism originally promulgated by Descartes.

Doug: That is quite an appealing theory.

Teacher: I agree, Doug. Let me get on to the second major problem with the identity theory.

Doug: Wait, Teacher! At least tell us, is property dualism consistent with identity theories?

Teacher: A hard question. Most identity theorists suggest that brain and mind have a full theoretical identity just like lightning and cloud-to-earth electrical discharges. But you could argue that property dualism is consistent with an attenuated kind of identity. This might disappoint most identity theorists because it suggests that the relationship of brain to mind is different from that of other aspects of our world in which we are moving from folk knowledge to scientific theories.

Doug: I think I followed that.

Teacher: Let's get back to the second major weakness of the classic identity theory. The commonsense or "strong" form of these identity theories is called a type identity theory. That is, if A and B have an identity relationship, that relationship is fundamental and the same everywhere and always.

Mary: OK. So where does this lead?

Teacher: Well, let's go back to Mr. A's depression. Let's imagine that in the year 2050 we have a super-duper

MRI scanner that can look into Mr. A's brain and completely explain the brain changes that occur with his depression. We can then claim that the brain state for depression and the mind state for depression have an identity relationship.

Mary: OK. Makes sense so far.

Teacher: Now, if Mr. A develops the same kind of depression again in 20 years, would we expect to find the exact same brain state? Or what about if Ms. B comes in with depression. Would her brain state be the same as that seen in Mr. A? Or even worse, imagine an alien race of intelligent sentient beings who might also be prone to depression and were able to explain to us how they felt. Is there any credible reason that we would expect the changes in their brains (if they have a brain anything like ours) to be the same as those seen in Mr. A? Philosophers call this the problem of multiple realizability, that is, the probability that many different brain states might all cause the same mind state (e.g., depression).

Doug: Sounds to me like you are making a pretty strong case against what you call the type identity theory. Is there any other kind?

Teacher: Yes. It is called a token identity theory, and it postulates a weaker kind of identity relationship.

Mary: Weaker in what way?

Teacher: It makes no claims for a universal relationship. It only claims that for a given person at a given time there is an identity relationship between the brain state and the mind state.

Mary: That makes clinical sense. I have certainly seen depressed patients who had very similar clinical pictures and got the same treatment, but one got better and the other did not. This might imply that they actually had different brain states underlying their depression.

Teacher: Precisely. My guess is that if the identity theory is correct, we will find that there is a spectrum extending from primarily type identity to token identity relationships. This will, in part, reflect the plasticity of the central nervous system (CNS) and the level of individual differences. For example, for some subcortical processes, such as stimulation of spinal tracts producing pain, there may be no individual differences, and the type identity theory may apply. It is as if that brain function is hard wired—at least, for all humans. All bets would be off for Martians! On the other hand, for more complex neurobehavioral traits that are controlled by very plastic parts of the human CNS, there may be highly variable interindividual and even intraindividual, across-time differences, so the token identity model may be more appropriate.

Francine: What is the third problem?

Teacher: The third problem is sometimes referred to as the explanatory gap (11) or the "hard problem of consciousness" (12). When most people think about an identity,

they tend to view mind from the outside. For example, recent research with positron emission tomography and functional MRI has produced an increasing number of results supporting identity theories. We can see the effects of vision in the occipital cortex, hearing in the temporal cortex, and the like. We have even been able to see greater metabolic activity in a range of structures, including the temporal association cortex, correlated with the report of auditory hallucinations.

Mary: Yes. My point exactly. This is leading the way to a view of psychiatry as applied neuroscience.

Doug: I also have to agree that this evidence strongly supports some kind of close relationship between brain and some functions of mind.

Teacher: Well, the problem of the explanatory gap is that it is easy to establish a correlation or identity between brain activity and a mental function but much harder to get from brain activity to the actual experience. That is, I don't have too much trouble going from the state of looking at a red square to the state of increased blood flow in the visual cortex reflecting increased neuronal firing. Modern neuroscience has been making an increasingly compelling case for viewing this as an identity relationship. However, I have a lot more trouble going from increased blood flow reflecting neuronal firing in the visual cortex to *the actual experience of seeing red*.

In the philosophical literature, this subjective "feel" of mind is often called "qualia." In a famous essay titled "What Is It Like to Be a Bat?" (13), the philosopher Nagel argued that the qualia problem is fundamental. We will never be able to understand what it feels like to be a bat.

With the qualia problem, we come up against the mind-body divide in a particularly stark and direct way. Even if we got down to the firing of the specific neurons in the cortex that we know are correlated with the perception of color, how does that neuronal firing actually produce the subjective sense of redness with which we are all familiar? Can you have an identity relationship between what seems to be clearly within the material world (neuronal firing) and the raw sensory feel of redness?

I should say that having spoken about this problem with a number of people, I have gotten a wide range of reactions. Some don't see it as a problem at all, and others, like me, feel a sense of existential vertigo when trying to grapple with this question.

Doug: I think this comes close to describing my concerns. Neuronal firing and the sense of seeing red—they just cannot be the same thing.

Mary: Maybe we have finally gotten to the root of our differences. I have no problem with this. When you stimulate a muscle, it twitches. When you stimulate a liver cell, it makes bile. When you stimulate a cell in the visual cortex, you get the perception of red. What's the difference?

Teacher: I am glad that the two of you responded so differently to this question. I would only say, Mary, that the major difference is that with stimulating a motor neuron and producing a muscle twitch, you have events that all take place in the material world. On the other hand, stimulation in the visual cortex produces a perception of the color red that is only seen by the person whose brain is being stimulated. In this sense, it is not the same thing at all. It is hard for some of us to see how the nerve cell firing and the experience of seeing red could be the same thing, at least in the way that lightning and an electrical discharge are the same thing.

Mary: It seems to me, Teacher, that you and these philosophers are just getting too precious. That's the way the world is. We humans with our consciousness are not nearly as special as you think. When you poke a paramecium, it moves away. A few hundred million years later you have big-brained primates, but nothing has really changed. It's all biology. I'm losing patience with all this identity-relationship talk. This sounds too much like abstract psychoanalytic theorizing to me.

Francine: I have been patient for a long time, Teacher. Can I have my say now?

Teacher: Sure, Francine.

Francine: Up until now, all of you have been approaching the mind-body problem the wrong way. Identity theories see the mind as a thing like a rock or a molecule. But the mind isn't a thing; it's a process.

Mary: Help me understand. What is the difference between a thing and a process?

Francine: Sure. There are two ways to answer the question, "What is it?" If I asked you about a steel girder, you would probably tell me *what it is*, that is, its composition and structure. You would treat it as a thing. However, if I asked you about a clock, you would probably say, "Something to tell time"; that is, you would tell me *what it does*. In answering what a clock does, you have given a functional (or process) description.

Mary: I think I get it. So how does this explain how brain relates to mind?

Francine: I think that states of mind are functional states of brain.

Doug: You've lost me! How does a functional state of brain differ from a physical state of brain?

Francine: Think about a computer. You can change its physical state by adding more random-access memory or getting a bigger hard drive, or you can change its functional state by loading different software programs.

Teacher: Let me interrupt for a moment. Francine is advocating what has been called functionalism, which is almost certainly the most popular current philosophical approach to the mind-body problem. Functionalism has strong historical roots in computer science, artificial intelligence, and the cognitive neurosciences.

Mary: Why has it been so popular?

Teacher: Well, maybe we can let Francine explain. But its advocates say that it avoids the worst problems of dualism and the identity theories.

Francine: Let me start with how this approach is superior to the type identity theories. When we think about mental states from a functional perspective, the problems of multiple realizability go away.

Mary: How?

Francine: A functional theory would say depression is a functional state defined by responses to certain inputs with specific outputs.

Doug: Such as crying when you look at a picture of an old boyfriend or having a very sad facial expression when waking up in the morning?

Francine: Correct. Functionalism doesn't try to say that depression is a particular physiological brain state. It defines it at a more abstract level as any brain state that plays this particular functional role, of causing someone to cry, to look sad, etc.

Mary: So functionalism would not be so concerned about whether the basic biology of depression in different humans or humans and aliens was the same as long as the state of depression in these organisms played the same functional role?

Francine: Yes. I find this especially attractive.

Mary: So functionally equivalent need not say anything about biologically equivalent?

Francine: Right. You can see how well functionalism lends itself to cognitive science and artificial intelligence. If brain states are functions connecting certain inputs with certain outputs (stub your toe → experience pain → swear, get red in the face, and dance around holding your toe and cursing), then this kind of state could be realized in a variety of different physical systems, including neurons or silicon chips.

Mary: I think I understand. But are functionalists materialists?

Francine: Can I try to answer this?

Teacher: Sure.

Francine: I think the answer is mostly yes but a little bit no. Functional states are realized in material systems, but they are not essentially material states.

Mary: Can you translate that into plain English?

Francine: OK. Let's go back to clocks. Their functional role is to tell time.

But we could design a machine to tell time that used springs, pendulums, batteries, or even water. In each case, the clock is a material thing, but very different kinds of material things that were nonidentical on a physical level could all have the same function—of telling time.

Teacher: Let me try to clarify. I think Francine is right that functionalism avoids the unattractive feature of dualism, which postulates a nonmaterial mental substance. But it resembles dualism in that it postulates two levels of reality. That is, there is the physical apparatus—I like to think of a huge series of switches—and then there is the functional state of those switches. Let's imagine a computer program that controls the railroad. You have thousands of switches in the form of transistors. Depending on whether one of those thousands of switches is on or off, you send a train to New York or Chicago. Recall that the fundamental physical nature of a transistor or any switch is not changed as a function of which position it is in. What is important are the rules of your program, that is, the functional significance of what that switch means.

Mary: I think I see more clearly. The switch is physical, but the significance of its on-ness or off-ness is really a function of the rules of the system—the program, in this case.

Teacher: Yes. Philosophers would say that the functional status of the switch is *realized* in some physical structure.

Doug: I think that I only dimly see the difference between functionalism and identity theories.

Teacher: Doug, let me try one more approach, and this, for me, is perhaps the most important insight that functionalism has given me. Identity theorists want to equate a specific physiological aspect of brain function with specific mental events. A problem with this is that at a basic level, a lot of what goes on in the brain (ions crossing through membranes, second messenger systems being activated, neurotransmitters binding receptors) is nonspecific. If you looked at the biophysics of cell firing, it would probably look similar whether the neuron was involved in a pain pathway, the visual system, or a motor pathway. So, on one level, I would wager that the functionalists are right: that the specific mental consequences of a brain event cannot be fully specified at a purely physical level (e.g., as ions crossing membranes) but must also be a consequence of the functional organization of the brain. The same action potential could initiate the activity of neuronal arrays associated with a perception of pain, the color red, or the pitch of middle C, depending on where that neuron is located, that is, its functional position in the various brain pathways.

Doug: That helps a lot. Thanks.

Mary: If we accept functionalism, aren't we at risk of sliding back into the whole functional-versus-organic mess? Am I not correct that functionalism might predict that some psychiatric illness is in the software, hence, functional, whereas others might be in the hardware? That has been a pretty sterile approach, hasn't it? I still want to stick up for the identity theories.

Teacher: That is quite insightful, Mary. Maybe we will have time to look quickly at the implications of these theories for psychiatry. Let me briefly outline how the philosophical community has responded to functionalism. I will focus on two main objections. The first, and probably most profound, addresses the question of whether a software program is really a good model of the mind. This approach, known as the “Chinese room problem,” was developed by the Berkeley philosopher John Searle (14). It goes like this. Assume that you are part of a program designed to simulate naturally spoken Chinese, about which you know absolutely nothing. You have an input function, a window in your room through which you receive Chinese characters. You then have a complex manual in which you look up instructions. You go to the piles of Chinese characters in this room and, carefully following the instruction manual, you assemble a response that you produce as the output function of your room. If the manual is good enough, it is possible that to someone outside the room it might appear that you know Chinese—but, of course, you do not.

Mary: So, what is the point?

Teacher: The point that Searle is making is that software is a bad model of the mind because it is only rules with no understanding (or, more technically, “syntax without semantics”). An aspect of mind that has to be taken into account in any model of the mind-body problem is that minds understand things. You know what the words “box,” “love,” and “sky” mean. Meaning is a key, basic aspect of some critical dimensions of mental functioning.

Francine: But there have been a number of strong rebuttals to this argument!

Teacher: I know, Francine, but if we are going to get done with these rounds, let me outline the second main problem with the functionalist approach: it defines mind-brain operations solely in terms of their functional status.

Doug: Yes. So, where is the problem?

Teacher: Well, for example, say I am faced with a color-discrimination task: having to learn which kind of fruit is bitter and poisonous (let’s say green) and which is sweet and nutritious (let’s say red). The state of my perception of color in this context is only meaningful to a functionalist because it enables me to predict taste and nutritional status.

In what is called the “inverted-spectrum problem,” if the wiring from your eye to your brain for color were somehow reversed so that you saw green where I saw red and vice versa, from a functional perspective, we would never know. You would learn just as quickly as I would that green fruit is bitter and to be avoided and red fruit is good and can be eaten safely. However, our subjective experiences would be exactly the opposite. You would have learned to associate the subjective sense of redness (which you would have learned to call greenness because that is what everyone else would call it)

with fruit to be avoided and greenness with fruit to be eaten.

Mary: I think I see.

Teacher: We need to move on to another of the most puzzling aspects of the mind-body problem.

Mary: How many more are there? My head is starting to swim.

Teacher: Hold on, Mary. Just a few more minutes, and we’ll be done. The first issue we need to talk about has a bunch of names, which I will call the problem of intentionality. If we reject dualistic models and accept one of the family of identity theories or functionalism, how do we explain that when I want to scratch my nose, amazingly, my arm and hand move and my fingers scratch?

Doug: Isn’t your big term “intentionality” another word for free will?

Teacher: Sort of, Doug, but I don’t want to get into all the ethical and religious implications of free will. I would argue that it is an absolutely compelling subjective impression of every human being I have spoken to that we have a will. We can wish to do things, and then our body executes those wishes. This phenomenon, which in the old dualistic theory might be called mind-brain causality, is pretty hard to explain using identity and functionalist theories.

Mary: The eliminative materialists have a solution: that the perception of having a will is false.

Doug: Isn’t there any philosophically defensible alternative to this rather grim view?

Mary: I’d be more interested in a scientifically defensible alternative. But I’m a little confused. If we accept the identity theory, aren’t we then saying that brain and mind are the same thing? So, if the brain wishes something—has intention, to use your words—then so does the mind.

Teacher: Technically you’re right, Mary. But here is the problem. How do carbon atoms, sodium ions, and cAMP have intentions or wishes?

Mary: Hmm. I’ll have to think about that.

Teacher: Although there are several different approaches to this problem, I want to focus here on only one: that of emergent properties and the closely associated issues of bottom-up versus top-down causation.

Francine: Can you define emergent properties for us?

Teacher: Sure. But first we have to review issues about levels of causation. Most of us accept that there are certain laws of subatomic particles that govern how atoms work and function. The rules for how atoms work can then explain chemical reactions, the rules that explain biochemical systems like DNA replication, and these in turn can explain the biology of life. I could keep going, but I think that you get the basic idea.

Mary: Yes. So what does this have to do with the mind-body problem?

Teacher: The concept of emergent properties is that at higher levels of complexity, new features of systems emerge that could not be predicted from the more basic levels. With these new features come new capabilities.

Francine: Can you give us some examples?

Teacher: Sure. One example that is often used is water and wetness. It makes no sense to say that one water molecule is wet. Wetness is an emergent property of water in its liquid form. Probably a better analogy is life itself. Imagine two test tubes full of all the constituents of life: oxygen, carbon, nitrogen, etc. In one of them, there are only chemicals—no living forms—and in the other, there are single-celled organisms. You would be hard pressed to deny that although the physical constituents of the two tubes are the same, there are not some new properties that arise in the tube with life.

Doug: Couldn't you say the same things about family or social systems, that they have emergent properties that were not predictable from the behavior of single individuals?

Teacher: Yes, Doug. Many would argue that. One critical concept of emergent properties is that all the laws of the lower level operate at the higher level, but new ones come online. So, the question this all leads to is whether we can view many aspects of mind, such as intentionality, consciousness, or qualia, as emergent properties of brain.

The theory of emergent properties can challenge traditional scientific ideas about the direction of causality. Traditional reductionist models of science see causation flowing unidirectionally through these hierarchical systems, from the bottom up. Changes in subatomic particles might influence atomic structure, which in turn would affect molecular structure, etc. But no change in a biological system would affect the laws of quantum mechanics. However, if we adopt this perspective on the mind-body problem, it is very difficult to see how volition could ever work.

Doug: I think I need an example here to understand what you're driving at.

Teacher: Let's look at evolution. Most of us accept that life is explicable on the basis of understood principles of chemistry. However, life is a classic emergent property. Evolution does not work directly on atoms, molecules, or cells. The unit of selection by which evolution works is the whole organism, which will or will not succeed in passing on its genes to the next generation. So your and my DNA are in fact influenced by natural selection acting on the whole organisms of our ancestors. Thus, in addition to the traditional bottom-up causality we usually think about—DNA produces RNA, which makes protein—DNA itself is shaped over evolutionary time by the self-organizing emergent properties of the whole or-

ganism that it creates. That is an example of top-down causality. But, critically, this hypothesis is nothing like dualism. Organisms are entirely material beings that operate by the rules of physics and chemistry.

Mary: So you see this as a possible model for the mind/brain?

Teacher: This is one of the main ways that people try to accommodate two seemingly contradictory positions, that dualism is not acceptable and is probably false and that the mind/brain truly has causal powers, so that human volition is not a fantasy as proposed by the eliminative materialists.

I need to see a patient in a few minutes. But before we end, we have to touch on two more issues. The first returns us to where we started, with dualism. As I said before, few working scientists today give much credence to classical Cartesian substance dualism, although property dualism does have some current adherents. However, there is a third form of dualism that may be highly relevant to modern neuroscience, especially psychiatry.

Francine: What is that, Teacher?

Teacher: It is called explanatory dualism and might be defined as follows: to have a complete understanding of humans, two different kinds of explanations are required. Lots of different names have been applied to these two kinds of explanations. The first can be called mental, psychological, or first person. The second can be called material, biological, or third person.

Doug: Aren't those just different names for Descartes' mental and material spheres?

Teacher: Yes. But with one critical difference. Descartes spoke of the existence of two fundamentally different kinds of "stuff." Technically, he was talking about ontology, the discipline in philosophy that examines the fundamental basis of reality. Explanatory dualism, by contrast, deals with two different ways of knowing or understanding. This is a concern of the discipline of epistemology, or the problem of the nature of knowledge.

Mary: Can you explain that without all the big words?

Teacher: A fair question! Explanatory dualism makes no assumptions about the nature of the relationship between mind and brain. It just says that there are two different and complementary ways of explaining events in the mind/brain.

Doug: To accept explanatory dualism, do you have to accept Descartes' substance dualism?

Teacher: No. In fact, explanatory dualism is consistent with identity theories or functionalism. Let's assume that the token identity theory about Mr. A's depression is true; that is, the serotonergic dysfunction in certain critical limbic regions in his brain is his depression. Explanatory dualism suggests that even if these brain and mind states have an identity relationship, to understand these

states completely requires explanations both from the perspective of mind (perhaps the psychological issues that Doug first raised at the beginning of our discussion) and the perspective of brain. Neither approach provides a complete explanation.

Doug: That is very attractive. This makes it possible to believe that mind and brain are the same thing and yet not deny the unique status of mental experiences.

Teacher: Yes, if you accept explanatory dualism.

Mary: Isn't this theory a bit unusual in that most events in the material world have only one explanation? We wouldn't think that you would have one explanation for lightning and another for earth-to-cloud electrical discharges. Why should events in the brain be different?

Teacher: I agree, Mary. Explanatory dualism suggests that there is something unique about mind-brain events that does not apply to material events that do not occur in brains, that they can be validly explained from two perspectives, not just one.

Doug: What is appealing about explanatory dualism is that it seems to describe what we do every day when we see psychiatric patients, despite all these discussions, and it may be that mind really is brain. I am still impressed with the basic fact that we have fundamentally different ways in which we can know brains versus minds. One is public and the other private. Getting back to Mr. A, the optimal treatment for him requires me to be able to view Mr. A's depression from both the perspectives of brain and mind. We need to be able to view the depression as a product of Mr. A's brain to consider whether his disorder might be due to a neurologic and endocrine disease and to evaluate the efficacy and understand the mode of action of antidepressant medications. But to provide good quality humanistic clinical care, and especially psychotherapy, I need to be able to use and develop my natural intuitive and empathic powers to understand his depression from the perspective of mind, thinking about his wishes, conflicts, anger, sadness, and the impact of life events in addition to autoreceptors, uptake pumps, and down-regulation.

Teacher: There is a lot that is very sensible in what you say, Doug. I would add that it is critically important for us to understand both the strengths and limitations of and the important differences between knowing our patients from the perspective of brain versus from the perspective of mind.

Doug: I agree.

Teacher: One last issue, and then I really have to go. When we look at major theories in psychiatry, like the dopamine hypothesis of schizophrenia or trying to tie Mr. A's depression to dysfunction in the serotonin system, what assumptions are we making about the relationship between mind and brain?

Mary: It is pretty clearly materialistic at least in the sense that changes in brain explain the mental symptoms of these syndromes.

Teacher: So these theories embody the assumption of brain-to-mind causality?

Doug: I am not clear on this. Do these biological theories of psychiatric illness assume a causal or an identity relationship between mind and brain?

Teacher: Good question, Doug. If you listen to biological researchers closely, they actually use causal language quite commonly. They might say, "An excess of dopamine transmission in key limbic forebrain structures *causes* schizophrenia."

Francine: Do they mean that, or do they actually mean that an excess of dopamine transmission *is* schizophrenia?

Teacher: I am not sure. My bet is that most biological psychiatrists prefer some kind of identity theory. I wonder if they use causal language because Cartesian assumptions about the separation of material and mental spheres are so deeply rooted in the way we think.

Mary: So they may not be very precise about the philosophical assumptions they are making?

Teacher: That's my impression.

Mary: What about the multiple-realizability problem? Are they likely to assume type identity theories that imply that a single mental state (e.g., auditory hallucinations) has an identity relationship with a single brain state or token identity states that suggest that multiple brain states might produce the same mental state, such as hearing voices?

Teacher: My guess is that most researchers suggest that token identity models are most realistic. They would be more likely to call it "etiologic heterogeneity," but I think it is the same concept in different garb.

Doug: After all, we know that hearing voices can arise from drugs of abuse, schizophrenia, affective illness, and dementia.

Mary: What about eliminative materialism? That would have pretty radical implications for the practice of psychiatry!

Teacher: Yes, it would, Mary. If you took that theory literally—that mental processes are without causal efficacy, like froth on the wave—then any psychiatric interventions that are purely mental in nature, like psychotherapy, could not possibly work.

Doug: We have a lot of evidence that psychological interventions work and can produce changes in biology. That would be strong evidence against eliminative materialism, wouldn't it?

Teacher: That's how I see it, Doug.

Francine: What about functionalism? Certainly theories of schizophrenia and affective illness have pointed toward

defective information processing and mood control modules, respectively.

Teacher: This gets to a pretty basic point. Etiologic heterogeneity aside, are specific forms of mental illness “things” that have a defined material basis or abnormalities at a functional level, like an error in a module of software?

Francine: This gets at what you said before. Functionalism is different from identity theories in that it implies abnormalities are possible in psychiatric illness at two levels: the functional “software” level that affects mind or the material “hardware” level that affects brain.

Teacher: Yes. I am ambivalent about that implication. It suggests two different pathways to psychiatric illness. Is it helpful to ask whether Mr. A developed depression with a normal brain that was “misprogrammed” perhaps through faulty rearing or because there is a structural abnormality in his brain? I’m not sure. I continue to feel that functionalism makes sense if you think about computers and artificial intelligence, but when you deal with brains like we do, I have my doubts. But as I said, this is still the most popular theory about the mind-body problem among philosophers. I have to go. I look forward to seeing you on rounds tomorrow. This was fun.

Doug, Francine, and Mary: Good-bye, Teacher.

Conclusions

The goal of this introductory dialogue was to provide a helpful, user-friendly introduction to some of the current thinking on the mind-body problem, as seen from a psychiatric perspective. Many interesting topics were not considered (including, for example, philosophical behaviorism and the details of theories propounded by leading workers in the field, such as Searle and Dennett), and others were discussed only superficially. Those interested in pursuing this fascinating area might wish to consult the list of references and web sites below.

Recommendations

Web Sites

The *Stanford Encyclopedia of Philosophy* (15) has a number of entries relevant to the mind-body problem. For example, see “Epiphenomenalism,” “Identity Theory of Mind,” and “Multiple Realizability.” See, also, the *Dictionary of Philosophy of Mind* (16).

David Chalmers has compiled a very useful list of “Online Papers on Consciousness” as part of a larger web site titled “Contemporary Philosophy of Mind: An Annotated Bibliography” (17).

Further Reading

Bechtel’s *Philosophy of Mind: An Overview for Cognitive Science* (18) is a good overview from the perspective of psychology, although rather technical in places.

A Companion to the Philosophy of Mind (19) is a helpful but somewhat difficult introductory essay followed by short entries on nearly all topics of importance in the mind-body problem. Very useful.

Gennaro’s *Mind and Brain: A Dialogue on the Mind-Body Problem* (20) is a brief, easily understood introduction to the mind-body problem, also in the form of a dialogue.

Churchland’s *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind* (21) is a good introduction to the mind-body problem. Although a strong advocate for eliminative materialism, Churchland fairly presents the other main perspectives. The chapter on neuroscience is dated.

Brook and Stainton’s *Knowledge and Mind: A Physical Introduction* (22) is a charming, accessible, and up-to-date introduction that includes sections on epistemology and the problem of free will.

Priest’s *Theories of the Mind* (23) is a quite useful, albeit somewhat more advanced, treatment of the mind-body problem. Priest takes a different approach from the other books listed here, summarizing the views of this problem by 17 major philosophers, from Plato to Wittgenstein.

Heil’s *Philosophy of Mind: A Contemporary Introduction* (24) is a recent introductory book, with an emphasis on the metaphysical aspects of the mind-body problem. A bit hard to follow in the later chapters.

Searle’s *The Rediscovery of the Mind* (25) is probably the most important book by this influential philosopher who has been very critical of functionalism. He writes clearly and with a minimum of philosophical jargon.

Nagel’s *The View From Nowhere* (26) is a brilliant book-length treatment of the key epistemic issue raised by the mind-body problem: that we see the world from a third-person perspective but ourselves from a first-person perspective.

Hannan’s *Subjectivity and Reduction: An Introduction to the Mind-Body Problem* (27) is a short and relatively clear introduction. The author makes no attempt to hide her views about the problem.

The Nature of Consciousness: Philosophical Debates (28) is probably the most up-to-date of the several available collections of key articles in this area, with an emphasis on problems related to consciousness.

Cunningham’s *What Is a Mind? An Integrative Introduction to the Philosophy of Mind* (29) is a particularly clear and up-to-date summary of the mind/body and related philosophical topics. It is one of the best available introductions.

Audiotapes

If you want to probe the mind-body problem on your way to work, you might want to try Searle’s *The Philosophy of Mind* (30) audiotapes. Searle has his own specific “take” on this problem, but he is down-to-earth and rather accessible for the beginner.

Received Aug. 22, 2000; revision received Dec. 5, 2000; accepted Dec. 21, 2000. From the Virginia Institute for Psychiatric and Behavioral Genetics, the Department of Psychiatry and the Department of Human Genetics, Medical College of Virginia of Virginia Commonwealth University, Richmond. Address reprint requests to Dr. Kendler, Box 980126, Richmond, VA 23298-0126.

Funded in part by the Rachel Brown Banks Endowment Fund.

Jonathan Flint, M.D., and Becky Gander, M.A., provided comments on an earlier version of this article.

References

- Kandel ER: A new intellectual framework for psychiatry. *Am J Psychiatry* 1998; 155:457–469
- Kandel ER: Biology and the future of psychoanalysis: a new intellectual framework for psychiatry revisited. *Am J Psychiatry* 1999; 156:505–524
- Seager W: *Theories of Consciousness: An Introduction and Assessment*, 1st ed. London, Routledge, 1999
- Edelman GM, Tononi G: *A Universe of Consciousness*. New York, Basic Books, 2000
- Damasio A: *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. San Diego, Harcourt Brace, 1999
- Guzeldere G, Flanagan O, Hardcastle VG: The nature and function of consciousness: lessons from blindsight, in *The New Cognitive Neurosciences*, 2nd ed. Edited by Gazzaniga MS. Cambridge, Mass, MIT Press, 2000, pp 1277–1284
- Baynes K, Gazzaniga MS: Consciousness: introspection, and the split-brain: the two minds/one body problem. *Ibid*, pp 1355–1364
- Descartes R: *Meditations on First Philosophy*. Translated by Lafleur LJ. New York, Macmillan, 1985
- Squire LR, Kandel ER: *Memory: From Mind to Molecules*. New York, Scientific American Library, 1999
- Libet B: Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav Brain Sci* 1985; 8:529–566
- Levine J: Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly* 1983; 64:354–361
- Chalmers D: Facing up to the problem of consciousness. *J Consciousness Studies* 1995; 2:200–219
- Nagel T: What is it like to be a bat, in *Mortal Questions*. New York, Cambridge University Press, 1979, pp 165–180
- Searle JR: Minds, brains, and programs. *Behav Brain Sci* 1980; 3:417–424
- Stanford University, Metaphysics Research Lab, Center for the Study of Language and Information: *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu>
- Eliasmith C (ed): *Dictionary of Philosophy of Mind*. <http://www.artsci.wustl.edu/~philos/MindDict/index.html>
- Chalmers D: *Online Papers on Consciousness*. <http://www.u.arizona.edu/~chalmers/online.html>
- Bechtel W: *Philosophy of Mind: An Overview for Cognitive Science*. Hillsdale, NJ, Lawrence Erlbaum Associates, 1988
- Guttenplan S (ed): *A Companion to the Philosophy of Mind*. Cambridge, Mass, Blackwell, 1994
- Gennaro RJ: *Mind and Brain: A Dialogue on the Mind-Body Problem*. Indianapolis, Hackett, 1996
- Churchland PM: *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*, revised ed. Cambridge, Mass, MIT Press, 1988
- Brook A, Stainton RJ: *Knowledge and Mind: A Physical Introduction*. Cambridge, Mass, MIT Press, 2000
- Priest S: *Theories of the Mind*. New York, Houghton Mifflin, 1991
- Heil J: *Philosophy of Mind: A Contemporary Introduction*. London, Routledge, 1998
- Searle JR: *The Rediscovery of the Mind*. Cambridge, Mass, MIT Press, 1992
- Nagel T: *The View From Nowhere*. New York, Oxford University Press, 1986
- Hannan B: *Subjectivity and Reduction: An Introduction to the Mind-Body Problem*. Boulder, Colo, Westview Press, 1994
- Block N, Flanagan O, Guzeldere G (eds): *The Nature of Consciousness: Philosophical Debates*. Cambridge, Mass, MIT Press, 1997
- Cunningham S: *What Is a Mind? An Integrative Introduction to the Philosophy of Mind*. Indianapolis, Hackett, 2000
- Searle JR: *The Philosophy of Mind*. Springfield, Va, Teaching Company, 1996 (audiotapes)