

Jackson and Churchland: Notes on mind and body Philosophy 1 Spring 2006

Recall the puzzle:

1. The human body is a material thing.
2. The human mind is a spiritual thing.
3. Mind and body interact.
4. Spirit and matter do not interact.

The four claims all look plausible, but cannot all be true. The set of claims is inconsistent. Any three of the four claims could be true together, but at least one must go. Which one?

Behaviorism dissolves the puzzle, roughly by denying both 2 and 3. The behaviorist denies that there is causal interaction between mind and body. See the Skinner reading, also Fodor. Skinner distinguishes *methodological behaviorism* and *radical behaviorism*; Fodor introduces the idea of *logical behaviorism*. Fodor raises difficulties for behaviorism.

The identity theorist asserts that mental events such as being in pain are identical with brain events, physical processes occurring in the brain. The identity in question is asserted to be type-type identity not merely token-token identity. The latter claims that any particular instance of a mental event will be identical with some particular instance of a physical event. The former view is a stronger claim. It is that any particular kind or type of mental event is identical to some particular kind or type of physical event. Smart gives the example of the lightning flash we see in the sky and electrical discharges between clouds. These two types of occurrence, we learn from science, are identical. There are not two correlated separate and independent events, the lightning flash and the electrical discharge; there is just one event. In a similar way, we learn from chemistry that water is identical to H₂O.

The identity theorist in effect responds to the puzzle of the relationship of mind and body by accepting 1, 3, and 4 and denying 2.

In Monday's lecture I said the identity theory goes hand in glove with functionalism; the idea that mental concepts are to be analyzed in terms of functional roles.

In Wednesday's reading, Fodor describes functionalism as an alternative to the type-type identity theory and a superior alternative to it. According to Fodor, functionalism easily accommodates the phenomenon of *multiple realizability*, but the type-type identity theory founders in dealing with this phenomenon.

Multiple realizability is the idea that the same mental states might be occurring in different beings even though their physiologies are very different. Humans suffer pain, and so do many other animals, we think, including dolphins. But the physical processes in dolphins and in other animals that give rise to pain might for all we know be very different from the physical processes that give rise to or are linked to pain in dolphins. And it could be the case that creatures from outer space, call them Martians, have immensely different physiologies but share a complex mental life with us. Martians might be intelligent, creative, fun-loving, compassionate, irritable, just as some humans are. Martians might also feel pain. These stories about creatures from outer space make sense; they are logically coherent. Some such stories might even be true.

According to Fodor, the functionalist can allow multiple realizability but the type-type identity theorist cannot. If the pain type event is identical with the type of brain event that involves (whatever pain does involve in human brains, let's call it) C-fibers firing, then such creatures as dolphins and Martians cannot experience pain, according to the identity theorist, because they lack the appropriate brain process, C-fibers firing. The functionalist can avoid this human chauvinism and allow that the same type of mental event can be realized in different kinds of creatures in very different ways.

Functionalism does not care what kind of stuff is involved in mental events; functionalism's claim is about how the stuff (whatever its nature) is organized. Fodor observes, "The problem with type physicalism is that the psychological constitution of a system seems to depend not on its hardware, or physical composition, but on its hardware, or program. Why should the philosopher dismiss the possibility that silicon-based Martians have pains, assuming that the silicon is properly organized? And why should the philosopher rule out the possibility of machines having beliefs, assuming that the machines are correctly programmed? If it is logically possible that Martians and machines could have mental properties, then mental properties and neurophysiological processes cannot be identical" (p. 455).

I described the functionalist analysis of mental concepts and the idea that mental events and physical events (brain processes) are type-type identical as part of a package deal response to the puzzle of the relationship between mind and body. Fodor says functionalism and type-type identity theory are incompatible. What gives?

We need to distinguish two versions of functionalism:

Nonrigid or analytical functionalism: Mental concepts refer contingently to whatever states happen as a matter of fact to occupy (or fill, or realize) certain functional roles. It is analytically true that to be a mental state of a given kind is simply to occupy a certain functional role, but it is contingently true of any particular inner state that it is a mental state of that kind.

So, pain might not have been pain. That is to say, it might have turned out that what we discovered was in fact the physical process in the brain that is identical to being in pain was in fact something else—a soul substance as Descartes affirmed, for instance. Empirical science, not conceptual analysis, tells us what in fact realizes the roles in terms of which any particular mental state concept is characterized.

Rigid functionalism. Mental concepts refer to functional states themselves—to the state of being in a state with a certain functional role. It is both analytically and necessarily true of a given mental state that it manifests itself in certain relations to behavior and to other mental states. Mental states are not identified with their role realizers, their physiological bases.

In another terminology, according to nonrigid functionalism, mental concepts like pain are not rigid designators. According to rigid functionalism mental concepts like pain are rigid designators.

Multiple realizability? The type-type identity theorist can say that pain in humans is identical to (say) C-fibers firing and pain in dolphins is identical to (say) D-fibers firing and pain in Martians is identical to the compression of fluid in their little tubes (if their

mental mechanism is hydraulic) or to their silicone dancing (if their mental mechanism is silicone-based) or whatever. Pain simpliciter is whatever realizes the specified causal roles in the appropriate population. Here the type-type identity theorist is helping herself to the framework of nonrigid functionalism. The story looks to be coherent.

It should be pointed out that “identity theory” may be somewhat misleading as a label for the kinds of identities claimed by the type-type identity theorist yoked to nonrigid functionalism. Consider

Water is identical to H₂O

And

Lightning is identical to electrical discharges between clouds.

Water and lightning are rigid designators, so these claims, if true, are necessarily true—true in all possible worlds.

But on the view being considered,

Pain in humans is identical to C fibers firing

Is contingently true, not true in all possible worlds. *Pain* here is a nonrigid designator.

Pain (the thing we experience, that hurts) might have turned out not to be pain (C fibers firing or whatever is the best neurophysiological account of the relevant brain process).

Fodor says “mental states are distinguished by their having what are called either qualitative content or intentional content.” Functionalism seems able to account for intentional content (see Fodor on this). But qualitative content is a problem for all varieties of functionalism.

Qualitative content is the subjective quality of experience. Fodor introduces the idea by asking us to imagine that we look at a blank wall through a red filter, and then look at the same blank wall through a green filter. Something in our experience changes, and this is an example of qualitative content.

Fodor: “The reason qualitative content is a problem for functionalism is straightforward. Functionalism is committed to defining mental states in terms of their causes and effects. It seems, however, as if two mental states could have all the same causal relations and yet could differ in their qualitative content” (p. 457). For example, we could imagine two persons who react behaviorally and verbally in identical ways to color stimuli, but the qualitative content of their color experiences is inverted or switches—so that what one person sees when she looks at a ripe tomato or a fire engine is what another sees when he looks at leaves on trees in springtime. One sees green when the other sees red and red when the other sees green. This makes sense; we can imagine it, it might be true. But the functionalist cannot allow for the possibility; it cannot be. So much the worse for the functionalist analysis.

Jackson presses the problem of qualitative content as a problem for any version of physicalism—the idea that a complete physical description of the universe with nothing added entails any other truths there are. The psychological truths supervene on these physical facts; the psychological truths are made true by the physical truths.

Jackson says No. Consider Mary, a scientist who knows all the physical truths there are to know about color. She is a color expert. However, she has lived all of her life in a room containing only black and white objects. She has never experienced anything colored. Yet she knows everything there is to know that is physical about color. Now imagine her leaving her black and white enclosure and encountering the brightly colored

world for the first time. She gains new knowledge—she learns what it is like to experience color. She learns something brand new about visual experience. But then her prior knowledge was incomplete. Jackson, concludes, “But she had *all* the physical information. Ergo there is more to have than that, and Physicalism is false” (p. 492). He also says, in the same spirit, “Tell me everything physical there is to tell, . . .you won’t have told me about the hurtfulness of pains, the itchiness of itches, pangs of jealousy, or about the characteristic experience of tasting a lemon, smelling a rose, hearing a loud noise or seeing the sky” (p. 490).

Jackson reasons as follows:

1. Mary knows everything there is to know about brain states and their properties.
2. It is not the case that Mary knows everything there is to know about sensations and their properties.
3. So, it is not the case that sensations and their properties are identical to brain states and their properties.

In the terms of the puzzle stated at the beginning of these notes, Jackson affirms both 1 and 2. What then of 3 and 4? Jackson proposes that qualia, experiences with subjective qualitative content, are causally inert with respect to the physical world. So feeling hungry does not cause me to eat, contrary to what we might have thought. Qualia are caused by physical events but themselves cause nothing. Jackson considers several objections against this epiphenomenalist position and finds the objections inflict no damage on the position.

He considers three objections.

1. How can one sensibly deny the obvious facts that show mind and body causally interact? The sensation that I feel hot causes a physical event, my hands loosening my shirt collar. This happens in a million different ways. —Jackson replies in a way that echoes the behaviorist denial that there is mind-body causal interaction. If we can explain behavior without invoking qualia, then the conjecture that qualia cause physical events is otiose.
2. “According to natural selection the traits that evolve over time are those conducive to physical survival. We may assume that qualia evolved over time”—so they should be conducive to survival, which they can’t be if they are causally inert. —Jackson replies that “we should expect any evolved characteristic to be either a conducive to survival or a by-product of one that is so conducive” (p. 493). See polar bear example to illustrate the point.
3. How can we know about other minds except by inferring from people’s behavior to their mental states? But this inference would seem to be incorrect unless qualia can cause behavior. If my feeling hot does not cause me to loosen my shirt collar, then you can’t infer from seeing me loosen my shirt collar that it is likely that I am feeling hot. Jackson replies that the epiphenomenalist can allow that qualia are caused by physical events even though they cause no physical events. So if the temperature is hot, and I observe in myself that hot temperatures cause me to loosen my shirt collar and are accompanied by my feeling hot, and I then see others loosen their shirt collars and the like when the temperature is hot, I have some basis for inferring maybe they too are feeling hot, like me.

Eliminative materialism is a version of physicalism that is revisionary with respect to our common-sense understanding of our mental life. *Physicalism* here is the doctrine that the universe contains nothing over and above the substances and properties

acknowledged in the physical sciences (the physical sciences as they will become when perfect and complete). As to what eliminative materialism eliminates, we need to distinguish some common-sense ideas of our mental life. Sensations such as we experience when we feel pains, see after-images and the like, do not have propositional content. In contrast, many of our mental concepts are concepts of *propositional attitudes*, attitudes toward propositions. Consider the proposition that it is raining here now. To be in the propositional attitude of belief with respect to that proposition is to believe that it is raining here now, to be in the propositional attitude of desire is to want that it is raining here now, and so on for wishing, fearing, hoping, and so on. The eliminative materialist holds that the propositional attitudes do not exist. There is no such thing as believing, desiring, hoping, wishing, fearing, and so on, as we now conceive them.

Our common-sense psychology in which believing, desiring, and so on figure prominently is infected with a false empirical theory, called "folk psychology" by the eliminativist. Talk of belief and desire is on the same footing as talk of witches and of phlogiston. Talk of witches involves presupposing a theory of witchcraft that turns out to be radically false. There never were women who consorted with the devil and killed their neighbors' cows by magical chants and spells, so there never were any witches. We once believed that there was a substance called *phlogiston* that was given off in combustion and oxidation. After the work of the chemist Lavoissier, we now see phlogiston theory as radically false; so we now think there never was any such thing as phlogiston. (Recall the earlier course reading in which Kuhn discusses this example for his own purposes.) According to eliminative materialism, the same goes for beliefs and desires. They are like phlogiston and witches; they do not exist. The common-sense psychological terms such as belief and desire are bound up with a common-sense psychological theory to the effect that belief combines with desire to cause behavior. This is an empirical theory, albeit vague and wooly, but we can see it is doomed to be supplanted by a future biologically based scientific psychology. The task of trying to figure out how belief and desire figure into our developing scientific account of our mental life is a dead end enterprise that should be abandoned. So says the eliminative materialist.

Eliminative materialism can sound paradoxical. This theorist believes there are no beliefs—how can that make sense? But the eliminative materialist does not deny that something is going on, that we now crudely describe in the language of desire and belief. When the dust settles and the scientific psychology of the future is known to us, we will use those concepts of common-sense pragmatic simplifications of them. Instead of desire we will have some new concept, call it desire₁, and instead of belief we will have some new concept, call it belief₁. The eliminative materialist believes₁ there are no beliefs as we now conceive them. This is not incoherent.

The eliminative materialist is a type-type identity theorist about sensations and perceptual experiences. So this account is vulnerable to the qualitative content objection to functionalism and the Jackson knowledge argument for qualia and against physicalism. Putting this point more neutrally, the eliminative materialist needs a response to the Jackson objection.

Here are some tries. (1) In the black and white room Mary knows everything there is to know if she knows all the physical facts, but when she steps into the colored world she learns a new way to express what she knows. Reply: The claim is that Mary learns new

facts not merely new ways of talking about them. (2) There is a difference between knowing that and knowing how. Whatever propositional knowledge you have, you don't necessarily have all of the knowledge how---like the knowledge of how to ride a bicycle or how to swim. When Mary leaves her black and white environment and negotiates the colored world she gains no new propositional knowledge but some new know-how. Reply: Granted, Mary will gain some know-how, but she also gains some knowledge that, for example, the knowledge that this is how red ripe tomatoes look and what it is like to have this type of color experience. (3) the opacity of knowledge. Just as you can know that Smith is in the room without knowing that the world's greatest swindler is in the room, because even though Smith is the world's greatest swindler, you don't know that, so you can know everything about brain states without knowing everything about sensations even though sensations are identical to brain states. Reply: In the Smith example, there are things Smith doesn't know, and this ignorance is involved in explaining how we cannot infer from the fact that he knows that P to the fact that he knows that $P \rightarrow$ even though P implies $P \rightarrow$. But in the Mary example, there is no ignorance to make possible this kind of slippage.---There are other arguments that purport to refute his argument. For the record, Jackson first proposed the knowledge argument in 1986, and today he does not accept the conclusion that we should reject physicalism. He says we should hold onto physicalism, so something must be wrong with his argument, even if we cannot now figure out what exactly goes wrong with the argument.

Eliminative materialism also opposes functionalism. The nonrigid functionalist specifies the concept of (for example) *desire* by specifying common-sense causal roles constitutive of it. This will be a list of platitudes or things we all know such as that if you desire something and believe that a certain course of action will get you what you desire (and this course of action will not prevent you from getting other things you desire more strongly) then typically you will behave in such a way that your desire will be satisfied if what you believe is true. The full list of these platitudes about what tends to lead to desire and the causal interactions of desire with other mental states and about what desire tends to lead to specifies the concept. What realizes or occupies these roles IS desire. Of course desire like almost all concepts is a cluster concept, so if enough of the properties in the cluster specifying concept X apply to a thing, then the thing is an X, even if it lacks some properties in the cluster.

The eliminative materialist says these platitudes, the bulk of them, are just false, not true, so the concept of desire as we commonly understand it has no integrity, and will and should disappear with the progressive march forward of psychological scientific knowledge.

The functionalist says No. The platitudes that go into specifications of the network of causal roles paired with any particular mental concept are empirical claims all right, but they assert only what is obviously true and what is virtually certain not to be superseded by the future march of science. Science will explain how it is that belief that beer is in the fridge coupled with desire for beer typically leads to beer-seeking behavior. Science will not explode this platitude as a false claim. Not every claim that is contingent and known by empirical experience stands a chance in a zillion of being proved false.