A Brief Review of Materialist Theories of Mind

The question these theories are trying to answer is what the mind is, what consciousness consists of, what experiences are (for example, what the experience of seeing red is, or the experience of being in pain is), and what beliefs and desires are. In sum, they are trying to give an analysis of *mental states* (where a mental state can be an experience, a belief, a desire, etc…).

I. Non Materialist Theories of The Mind
1. **Dualism:** the view that (at least some) mental states are not physical states.

> *Problems:* (a) dualism is inconsistent with physicalism – the view that everything that there is, is physical, which many people are sympathetic to. (b) Even if you're not a physicalist, a lot of people think that the physical world is causally closed – that is, that every physical state can be explained fully by appeal to other physical events. But, it seems as if mental states are the sorts of things that cause physical states (for example, my desire to eat ice cream causes me to eat ice cream). Unless we give up on the claim that our mental states can cause physical states (as Jackson does), it looks like we have to give up on causal closure of the physical world.

II. Materialist Theories of the Mind
Some general issues with materliaism: the Mary argument (also called "the knowledge argument – it looks like maybe you can know all the physical facts without knowing facts about what it's like to see red), subjectivity (facts about our mental life seem to essentially involve a subjective or first person point of view; facts about physics seem to essentially involve an objective or third person point of view).

2. **Behaviorism:** The theory that all mental states are dispositions towards behavior: each mental state is identical to some disposition to behave.

> *Problems:* (a) the matching problem/the circularity worry – it doesn't look like we can match each mental state with a behavioral disposition that doesn't itself appeal to other mental states. For example, what behavioral disposition is associated with the belief that there is a tiger in front of me? Well, it's not always going to be the same thing. For example, if I am *afraid* of the tiger, I might run away, or if *love* tigers I might try to pet it, or if I *desire* a tiger I may try to catch it and bring it home. However we spell out the behavioral disposition is going to have to appeal to other mental states, and thus the analysis becomes circular. (b) Access – it looks like we can know what our own mental states are without knowing what our behavioral dispositions are. (c) Counterexamples – can't there be people with mental states that don't have the corresponding behavioral dispositions? Like very stoic people? Similarly, can't people have a behavioral disposition without the corresponding mental state? (For example, I want people to think I like coffee so I drink coffee as much as possible, whenever it's offered to me, etc but secretly I hate coffee). (d) Causation – don't our mental states cause our behavioral dispositions? Surely I'm disposed eat ice cream because I like the taste of ice cream – but if liking ice cream is just being disposed to eat it, we can't explain why I'm disposed to eat ice cream by appealing to the fact that I like it (that would be to say that I'm disposed to eat ice cream because I'm disposed to eat ice cream).

3. **The Identity Theory:** The theory that mental states are physical states of the brain, for example: "Pain is C-fiber stimulation."

Problems: The big problem for the identity theory is the chauvinism problem, though the problems raised for materialism in general, and some of the problems raised for behaviorism (like access) are problems the identity theory as well. Chauvinism – isn't it possible that there are aliens, or other beings, that can feel pain and have experiences even if they don't have C-fibers, or brains made of organic material?

4. **Functionalism:** Functionalism is the theory that mental states are internal states of a system that are specified by their causal relations to sensory inputs, behavioral outputs and other mental states. For example, being in pain is a state that is frequently caused by damage to tissue and is such that if you're in that state, and believe that you can get out of that state by doing A, (and a bunch of other conditions hold), you will do A, and then move to a different state S….

      Problems: (a) The China Objections (see other handout). (b) The possibility of inverted qualia or zombies. If you think it's *conceivable* that two beings could have the same functions and capacities with regard to color vision, but what looks red to one looks green to the other and vice versa you can't be a functionalist. Why? Because even if the two people you're imagining are *physically* different, they're *functionally* the same. (The same inputs cause the same outputs). And if mental states are functional states, then two people who are *functionally* the same, must have the same experiences. And if two people have the same experiences it can't be that the experience one gets from looking at a ripe tomato differs from the experience the other gets when looking at the ripe tomato. The zombie objection is similar – it says: isn't it conceivable that there be beings who are functionally like us but don't have any experiences? If so, you can't be a functionalist because if they're functionally like us they're mentally like us, and so, since we have experiences, they would have to have experiences as well.

III. Materialist Theories of the No-Mind
5. Eliminativism: There are no mental states.

Functionalism Travels to China
(or: Problems with Functionalism)

Who would have thought that nation of China could cause serious difficulties for a philosophical theory of mind? Well, perhaps you wouldn't have. But think again! In what follows, see how China tries to knock functionalism out! But will the attempt be successful? Read on…

I. The Chinese Room Argument (Searle)

Argument #1:

1. According to Functionalism, the man in the Chinese Room understands Chinese. (Assumption)
2. The man in the Chinese Room doesn't understand Chinese. (Assumption)
3. Functionalism is false. (1,2).

Functionalist: This argument against functionalism is horrible! Premise 1 is false! It's just not true that according to Functionalism, the man in the Chinese Room understands Chinese. The man is comparable to a neuron, because he is just part of the system. If you asked the man questions without the book, he wouldn't be able to answer them which means that he *himself* doesn't understand Chinese. Rather, it's the man+book system that understands Chinese.

Searle: Okay, fine. I'll give you a new argument:

Argument #2:

1. According to Functionalism, the man+book system understands Chinese. (Assumption)
2. The man+book system doesn't understand Chinese. (Assumption)
3. Functionalism is false. (1,2)

Functionalist: Searle, you have to do better than that! In Argument #2, like Argument #1, the first premise is still false, but it's false for a different reason! The Chinese Room system as described is too simple. You couldn't mimic the understanding of a language by providing a rule book which matches any Chinese question, with a Chinese answer. Understanding Chinese is a much more complicated property than that! If someone understood Chinese, then, when asked what color the ceiling in the room is, he'd be disposed to say white, if it's white, and blue if it's blue, etc. But the simple book described just provides one answer to any given question. Also, someone who understood Chinese would respond differently if asked the same question multiple times in a row which isn't true of the man with the book. So it's not true according to functionalism that the man+book system understands Chinese.

Searle: Well, functionalist, if you're going to be difficult like that – I'll produce an even fancier argument for you! This time, I'll redescribe the whole system. It's not crucial to my point that the system be described as a book which just matches single answers to single questions. The system could be as complicated as you want. The room could include a whiteboard which stated what state the system was in. The book would then have instructions like this: If you're in state X, and you get input Y, then shift to

state Z, and reply with output O. I'll call this system the super-man+book+whiteboard system. Here's my new argument:

Argument #3:

1. According to Functionalism, the super man+book+whiteboard system understands Chinese. (Assumption)
2. The man+book+whiteboard system doesn't understand Chinese.
3. Functionalism is false.

Functionalist: Okay Searle. You finally seem to understand my theory and what it's committed to! I no longer have any objections to premise 1. The problem with Argument #3 though, is that premise 2 is false! Once you've made the system complex enough, I see no reason to think that it doesn't understand Chinese. Premise 2 certainly isn't as obvious as second premises in the other two arguments you gave.

Searle: Well, here's where we disagree. I think there are principled reasons for denying premise 2 even in Argument #3. And my reasons are this: no matter how complicated a system you make, the system described is purely syntactic: in other words the rules governing the system are just rules about manipulating symbols. But understanding Chinese requires more than manipulating symbols! You have to know what those symbols stand for! And you won't know what the symbols stand for, just by virtue of being able to manipulate them. Syntax alone is not sufficient for semantics!

Functionalist: Well I think you're wrong! Syntax can be sufficient for semantics! If the rules for symbol manipulation are complicated I enough, I claim that the system *will* have an understanding of semantics as well as syntax. For according to my view, to know what some symbol stands for is just to be such that…(functional description). The system you gave satisfies that functional description and so the system knows what the symbol stands for.

II. The China Brain Argument (Block)

Argument:

1. According to functionalism, the China Brain Robot can experience pain. (Assumption)
2. The China Brain Robot can't experience pain. (Assumption)
3. Functionalism is false. (1,2)

Functionalist: Hmm…okay, your first premise looks more promising than Searle's original first premise – but I'm still not so sure about it. For example, it seems to me like the system would work too slowly to be the functional equivalent of a being in pain.

Block: I have two things to say in response. First, it's not at all obvious to my why the speed of the system is relevant. You, the functionalist, have characterized being in pain as being in a state specified by its causal relations to inputs, outputs and other mental states. There's nothing in there about SPEED. Second, even if you want to add some sort of speed requirement to your functionalist theory, all I need

is that it be possible *in principle* to implement the China Brain. It's not important that in the actual world, it couldn't process very quickly. We can imagine a China Brain where all the people in China are on some sort of crazy amphetamine or, are just way smarter than people in our world. This brain could process as quickly as a human brain. So if it makes you happier imagine the China Brain as being like that. Then you will certainly have no problem with my first premise.

Functionalist: Okay. I'm convinced. But now that I've thought about it a bit more, I'm not sure why I even started attacking premise 1. Obviously, the problematic premise is premise 2! It's not clear to me that there is anything wrong with saying that the China Brain Robot is in pain. It's not as if the individual Chinese people would have to be in pain; it's the system as a whole – and as a functionalist, I have no problem with believing that the system is experiencing pain.

Block: Okay – you're weird. It just *obvious* that that this system couldn't be in pain!

Functionalist: Not so fast. I don't think it's as obvious as you say it is. If you saw this robot walking around acting exactly as a human would, I don't think it would be so clear to you that he wasn't in pain. But anyway, if you're going to claim that the China Brain Robot isn't in pain, I challenge you to find some property that a human in pain has, that the China Brain Robot doesn't, and explain why this property is important for being in pain. Just saying it's obvious isn't enough! At least when I was arguing with Searle, he gave me a REASON for thinking that the super man+book+whiteboard system didn't understand Chinese. He said that the system doesn't have any knowledge of semantics, whereas people who understand Chinese do.   Now as you know, I ultimately disagreed with Searle about whether the system has semantic knowledge, but my point is, at least he TRIED to find a principled way of distinguishing his system from a person that understands Chinese. If you're argument is going to work, you need to do something similar. Explain to me WHY the China Brain Robot couldn't be in pain.

Block: Hey, let's not get personal here. Do you really think that I can't do as well as Searle? I think there are loads of differences between a human in pain and the system in pain. For example: A human being is not spread out over huge spaces; a human being's neurons are not conscious beings; I could go on and on if you want.

Functionalist: Don't trouble yourself. Recall that I asked you to provide me with a *relevant* difference. Can't we imagine enormous aliens who feel pain whose brains are even bigger than the nation of China? Can't we imagine aliens who are in pain and when we look in their brain it turns out that the neurons are actually teeny little people? Size and what kind of thing is acting as a neuron seems totally irrelevant to whether something can feel pain, even by your own lights.

24.00 Problems in Philosophy
Fall 2010