Class 20 - April 8
Materialism

I. Lessons from Behaviorism and Dualism

Behaviorism was the first contemporary materialist program.
Early behaviorists identified mental states with behavior.
Jerry Fodor's chess-player example revealed a worry about behaviorist accounts of mental states couched in terms of mental causation.
Often, our overt behavior is the result of long causal chains of private thoughts.
But the quiet chess player shows no overt behavior to distinguish among the distinct thoughts, to guide the train of thought.

> A man may be angry, but give no bodily sign; he may think, but say or do nothing it all (Armstrong 297).

Mature behaviorists added dispositions to behave.
On the dispositional view, I don't cry because I am sad; my sadness is the disposition to cry.
Pain is not an introspective state but the disposition to scream or cry or say, "That hurts!"
I don't say that the apple looks red because I see red.
My seeing red just is my statement, and other related behaviors.
Even for the dispositional behaviorist, the question of why I scream when I am in pain remains without an internal account.

I mentioned that we can divide mental states into two kinds: occurrent sensations (qualia, or qualitative states) and intentional states.
The intentional states are also called propositional attitudes or representational states: I believe that..., I want that..., I intend that..., where the ellipsis is filled-in with a proposition.
The dispositional behaviorist's account is more plausible for intentional states.
My belief that it is cold does seem at least correlated with a particular range of behaviors.

Qualitative states, like being in pain, though, are only awkwardly accommodated by dispositional interpretations.
Here is an argument from Hilary Putnam against the behaviorist's account of qualitative states.
Consider two people whose motor nerves are cut, but only one of whom has cut pain fibers.
Kick them both, hard.
One feels pain, and the other does not.
But they have the same behavior, even potential behavior.

In order to distance themselves from Cartesian-style dualists, whose theories posit immaterial substances, behaviorists denigrated all private, internal processes.
But there do seem to be internal workings of our minds, mental states not apparent in behavior.
Furthermore, neuroscientific results seem to imply that some understanding of the brain is relevant to our understanding of minds.
The brain seems to have an important causal connection with my mental states.

Dualism, which focused exclusively on an internal account, inevitably meets a problem of interaction.
It would be nice to find a theory of mind which could accommodate both the internal states with their

causal connections to behavior in a scientific framework.
That's Armstrong's program.

> We can give an account of man *in purely physico-chemical terms* (Armstrong 295).

This theory of mind is called mind-mody materialism, or the identity theory.

II. Identity Theory

Armstrong interprets mental states naturally, like the dualist, as causes of behavior.
Overt acts like speech are not constitutive of a mental state.
My saying, "I am in pain," is not being in pain.
It is a report of my mental state.
My mental state contributes to my speech and other behavior as a cause.

> If somebody speaks and acts in certain ways it is natural to speak of this speech and action as the *expression* of his thought (Armstrong 298).

Armstrong agrees with the behaviorist that mental states are linked to behavior.
We might identify mental states in terms of the behaviors they cause.
Similarly, we might identify sugar-maple trees by their sap.
Still, the sap is not the tree and the behavior is not the mental state.

> Two separate but interlocking lines of thought have pushed me in the same direction. The first line of thought is that it goes profoundly against the grain to think of the mind as behavior. The mind is, rather, that which stands behind and brings about our complex behavior. The second line of thought is that the Behaviorists' dispositions, properly conceived, are really states that underlie behavior, and, under suitable circumstances, bring about behavior. Putting these two together, we reach the conception of a mental state as *a state of the person apt for producing certain ranges of behavior...* The mind is properly conceived as an inner principle, but a principle that is identified in terms of the outward behavior it is apt for bringing about (Armstrong 299b).

The identity theorist's claim is that every mental state is strictly identical with a physical state.
Identity theory simply claims that sensations and other mental states are brain processes.

III. Identity Theory and Theoretical Identity

The identification of mental states with brain states is supposed to be like any other common theoretical identification, or reduction, in science.
Compare it to two paradigmatic theoretical identifications.

> Lightning is electrical discharge.
> Water is H_2O.
> Mental states are brain states.

People were once ignorant of the nature of lightning and water.
Scientific theories were eventually proposed which identified their real natures.

Now we know that lightning is electrical discharge and water molecules are hydrogen and oxygen.
Similarly, people currently do not know, for example, that their pains are really stimulations of C-fibers in their brains.
Now that we know these things, we can use the old terms (lightning, water, pain) as shorthand (for electrical discharge, H_2O, and C-fiber stimulation).

One objection to the claim that mental states are brain processes involves analyzing the meanings of mental state terms.
The meaning of 'John has a toothache' clearly differs from the meaning of 'John has stimulation of his c-fibers'.
Some philosophers conclude that the differences in meaning entail that the identification of pain with c-fiber stimulation is unacceptable.
If that inference were acceptable, then all theoretical identifications would be incorrect.
Lightning would not be electrical discharge since the meaning of 'lightning' differs from the meaning of 'electrical discharge'.
Water would not be H_2O.
Thus, the kind of identification invoked by the identity theorist, like that used in other scientific reductions, must not be a kind that entails that terms for the reduced phenomenon have the same meaning as terms for the more basic phenomena.

Similarly, notice that we can know about our mental states without knowing about our brain states.
I can know that I am happy or in pain without knowing anything about my brain.
Some philosophers claim that our ability to know about our mental states without having any knowledge of our brain states shows that the identity theory must be false.
Indeed, Descartes defends precisely just this point.

In contrast to these two objections, the identities of mental states with brain states are supposed to be empirical matters, the possible results of scientific research.
The fact that people know about their pains without knowing about their c-fiber stimulations, and the fact that the meanings of the two kinds of terms seem to differ, are all irrelevant to the identification.
Some identity theorists call these identifications contingent, to characterize the fact that we know about them in different ways.
Calling an identity contingent, though, is contentious, as we will see.
First, let's look more closely at the empirical or scientific nature of the identity.

IV. Materialism, Dualism and Science

Identity theory synthesizes the inner states of dualism with the physical evidence of behaviorism.
Both identity theory and dualism provide inner causes, unlike the behaviorist.
If we abandon behaviorism on the basis of its inability to account for internal states, there is a question about which inner-state theory to choose.
One advantage of materialism over dualism is parsimony.
Suppose that materialism and dualism can account for our mental lives equally well.
We must decide between them on the basis of other aspects of the theories like simplicity and elegance and parsimony.

Compare the debate between identity theory and dualism to the invocation of the Gosse hypothesis in defense of creationism.

The Gosse hypothesis says that no scientific evidence like fossils or carbon dating could weigh against creationism since the whole history of the world could be created at any moment.
Even my memories are no evidence of a world history, since they could be created all at once by an omnipotent God.
Perhaps the world was created just now.

Similarly, consider the following two hypotheses:

> H1: Space-time is curved.
> H2: Space-time is flat, though all the evidence will make it appear curved.

Or, consider:

> H3: There are no ghosts.
> H4: There are ghosts, but they can never appear to us.

Or, consider:

> H5: Evolution is correct.
> H6: Creationism is correct, though the evidence supports evolution.

The question of whether to accept H1 or H2, or H3 or H4, or H5 or H6, is not empirical.
We resolve such questions by appealing to the methods of science, including parsimony and the desire to unify theories.
The question of whether the mind is the brain or the heart is empirical.
But, the question between dualism and materialism is not empirical.
It is, like the choices between H1 and H2, or between H2 and H4, methodological, or conceptual.

Armstrong defends methodological scientism by the fact that science can answer questions.

> Before Copernicus, Galileo and Harvey, Hobbes remarks, "there was nothing certain in natural philosophy." And, we might add, with the exception of mathematics, there was nothing certain in any other learned discipline... Since that time, ti has become a commonplace to see new doctrines, sometimes of the most far-reaching kind, established to the satisfaction of the learned, often within the lifetime of their first proponents. Science has provided us with a method of deciding disputed questions (Armstrong 296a).

Science is a domain of settled facts which replace speculation.
Scientists make errors, of course, but consensus leads to progress.
We should try to settle some facts about the mind by seeing what science says, especially neuroscience.

V. A Modal Objection

[Note: this long section is mainly an aside. It contains a neat statement of an objection to identity theory, and discusses some really cool recent work in philosophy, but it could easily be skipped.]

Whatever neuroscience has to say about brains, yummy brains, the philosophical identification of minds and brains requires arguments beyond the empirical data.

In response to worries about semantic and epistemological differences between mental states and brain states, the identity theorist may claim that theoretical identities are contingent.
Contingency is a modal property.
To understand so-called contingent identity, we must understand modal properties.
A modal property, in the relevant sense, is any characteristic that an object could have (possible properties) or must have (necessary properties).
An actual property of an object is contingent if it is possible for the object not to have that property.
For example, I am contingently the height I am.
But it seems to be necessary that I have my parents.
Necessary properties are historically called essences.
Scientists explore actual (non-modal) properties.
Philosophers explore possible properties, often by engaging in thought experiments.
There has been a lot of work on the question of theoretical identifications and essences, and on the related notion of rigid designation, in the last thirty years, involving metaphysics and the philosophy of language.

To see more how theoretical identifications are supposed to be contingent, consider whether it is possible for pain not to be a burst of neural activity.
It seems possible for pain, or any mental state, not to be identical with a particular brain state.
Similarly, since people once did not know that lightning and water were electrical discharge and H_2O, respectively, it seems wrong to say that they are necessarily identical.
For all people knew, it could have turned out that water is something entirely different.
So the identifications seem contingent.

If pain were necessarily identical with, say, a burst of activity of some specific regions of the central cortex, then it might be thought that we should be able to translate statements about pain into statements about neural activity.
The two kinds of statements might also be taken to have the same logical structure.
Consider the identifications of bachelors and unmarried men; or of Superman and Clark Kent.
Any entailments form sentences containing one of the pair, it might be argued, would equally follow from sentences containing the other of the pair.
But mental-state sentences and brain-state sentences are different in both meaning and logic.
The claim that the identities are contingent allows the identity theorist to concede differences in meaning and logic, without denying the identity.

Saul Kripke has criticized identity theory for identifying mental states with brain states, since even if they share actual properties, they have different possible properties.
Here's the relevant work of Kripke's; the most relevant portion is pp 8-11.
Kripke's claim is that if two objects are identical, then they share all properties, including modal properties.

> Let 'A' name a particular brain state, and let 'B' name the corresponding brain state, or the brain state some identity theorist wishes to identify with A. *Prima facie*, it would seem that it is at least logically possible that B should have existed (Jones's brain could have been in exactly that state at the time in question) without Jones feeling any pain at all, and thus without the presence of A... If A and B were identical, the identity would have to be necessary (Kripke, *Naming and Necessity*, p 146).

Mental states can not be identical to brain states, Kripke argues, since mental states and brain states have different possible properties.

Kripke introduces the notion of rigid and non-rigid designators to support his rejection of the claim that mental states and brain states are contingently identical.
A rigid designator is something that names the same object in all possible worlds, in all counterfactual circumstances.
Consider:

A. The number of planets is greater than seven.

A is true, but it might be false.
Contrast that claim with

B. Nine is greater than seven.

B is necessarily true, since 'nine' rigidly designates a particular number.
A is only contingent because 'the number of planets' designates a different number in different possible worlds.
Kripke argues that names are rigid designators.
Even if the facts about a person were different, the person's name still refers to that person.
We saw this claim in Kripke's discussion of the Queen of England.
'Queen Elizabeth II' is a rigid designator, since it refers to her in all possible worlds (in which she exists).
Kripke also uses the examples of Ben Franklin and Nixon.
Even if 'the inventor of bifocals' referred to some one else, 'Ben Franklin' would still refer to Ben Franklin.
Thus, 'the inventor of bifocals' refers non-rigidly, whereas 'Ben Franklin' refers rigidly.

In another case, Kripke tells a story about a possible discovery about Kurt Gödel.
Gödel actually proved that arithmetic, and other related theories, are essentially incomplete.
That is about all that most people know about him.
So, when people use the name Gödel, they mean to refer to the person who proved the incompleteness theorems.
What if Gödel actually stole the theorems from his neighbor Schmidt?
If people mean 'the person who proved incompleteness' when they use 'Gödel', then they really are referring to Schmidt, by the name 'Gödel'.
But, says Kripke, that's just not the case.
When people use 'Gödel', they refer to Gödel, no matter who actually proved incompleteness.
Names are rigid designators.

Notice that identity statements between rigid designators must be necessary.
Let's say that a is identical with b.
For example, 'Russell' is identical with 'Professor Marcus'.
Then, in any possible world, 'a' refers to a, and 'b' refers to b.
So, there are no possible worlds in which a is not identical to b, nor where 'a = b' is false.
There are possible worlds in which 'Ben Franklin is the inventor of bifocals' is false because 'the inventor of bifocals' refers, in any possible world, to the actual inventor of bifocals.
In some possible worlds, Franklin was not the inventor of bifocals.
But, in all possible worlds Franklin was Franklin.
'Russell = Professor Marcus' is true in all possible worlds, even though there are some possible worlds in which I did not become a college professor, because we use 'Professor Marcus' in this world to refer to me, in all possible worlds.

The relevance of rigid designation to our topic is due to Kripke's claim that theoretical terms are rigid designators.
For example, he argues that 'heat' and 'molecular motion' are rigid.
'Heat' is rigid, since in counterfactual situations in which people or Martians did not feel warmth when putting their hands near fires, we would not say that they did not feel heat.
We would say that they get a different sensation from heat than the one that we get.
Even if there are no people to feel it, fire heats up the air around it.
Heat thus rigidly designates molecular motion.
Similarly, 'molecular motion' is also a rigid designator, referring to the same thing, the motion of molecules, in all possible worlds.
Thus, the identification of heat with molecular motion is necessary.
That there are people who feel heat in a certain way is contingent.
Our skin could be constructed differently, say made of asbestos.
But we should not confuse the contingent property of heat (that people feel it in a particular way) with a necessary property of heat (that it is molecular motion.)

Kripke's broader claim is that all theoretical identity statements are, in fact, necessary identities, not contingent identities.
Since theoretical identity statements are necessary, the identification of pain states with brain states must also be necessary.
For, Kripke claims that pain is a rigid designator.
Nothing could be a pain if it did not hurt in the way that pains do.
Similarly, if 's' designates a brain state, it does so rigidly.
The identity of any two rigid designators must be necessary, since neither term could refer to anything other than its referent.

The problem distills to this: it seems possible that pain could be something other than a particular state of the brain.
If it is possible that pain is not a state of the brain, then the identity of the two must be contingent.
But, theoretical identifications must be necessary.

1. The identification of mental states and brain states must be either contingent or necessary.
2. Since mental states and brain states refer rigidly, the identification can not be contingent.
3. Since it is possible that mental states are not states of the brain, the identification can not be necessary.

Thus, mental states and brain states must not be identical.

I will not pursue this objection any further, but it would make a good paper topic.

VI. Qualitative States and Consciousness

Kripke's criticism of identity theory is closely related to a problem about consciousness.
The identity theorist has an uneasy relationship with consciousness.
On the one hand, the identity theorist embraces conscious states to criticize behaviorism.
On the other, the identity theorist wants to eliminate any remaining traces of substance dualism.
So, Armstrong needs a materialist account of conscious mental states.

The problem is that brain states seem intrinsically to lack consciousness.
They are just collections of neural firings.
So, the identity theorist seems to share a problem with the behaviorist.
Both the materialist and the behaviorist seem to lack an account of the way our mental states feel to us.

J.J.C. Smart, another (Australian) identity theorist, argues that sensations are nomological danglers.
'Nomological' refers to the laws of science.
As natural science proceeds, it subsumes increasingly many phenomena under its laws.
To claim that something is not amenable to scientific explanation is to leave a nomological dangler.

There are two problems with Smart's approach.
First, by denigrating mental states, the identity theorist's criticism of behaviorism is weakened.
If internal states are illusory, then behaviorism is still viable.

Second, a difference between sensations and brain processes seems to remain.
According to the identity theory, brain states and mental states need not share all properties.
The property of appearing red is not the same as the property of a set of neuron firings which correspond with that appearance, even if there is only one event in question.
The fact that I see yellow, instead of say black, when I look at a banana, will figure causally in the explanation of why I ate the banana instead of throwing it away.
But, if mental properties and physical properties are distinct, the identity theorist seems committed to nomological danglers, too.

> We are conscious, we have experiences. Now can we say that to be conscious, to have experiences, is simply for something to go on within us apt for the causing of certain sorts of behavior? Such an account does not seem to do any justice to the phenomena. And so it seems that our account of the mind, like Behaviorism, will fail to do justice to the first-person case (Armstrong 300a-b).

Armstrong attempts to provide an account of the missing elements of the materialist's theory.
Consider an action that we perform absent-mindedly, like driving a car while not thinking about it, or walking.
Armstrong says that the missing elements of the materialist's theory are just like the experiences we neglect when not paying attention to walking or driving.
Our conscious awareness, in those cases, is missing.
But all that comes to is that we are not thinking about what we are doing.

> In general, we can think of perceptions as inner states or events apt for the production of certain sorts of selective behavior towards our environment...My proposal is that consciousness, in this sense of the word, is nothing but *perception or awareness of the state of our own mind* (Armstrong 301a).

Consciousness is just thinking about thinking.
The missing elements of the identity theory are just some kinds of thoughts, more neural firings.
This suggestion of David Armstrong has been developed by David Rosenthal, into the HOT theory of consciousness: conscious thoughts are higher-order thoughts, thoughts about thoughts.
The question about the proposal is whether Armstrong's analogy between consciousness and perception yields a satisfying theory of consciousness or whether there is something first-personal left out.
We will return to this question next week.

VII. Multiple Realizability and Identity Theory

A last argument against identity theory, one which will impel us to our final theory of mind, is that it is chauvinistic.
A chauvinistic theory is too narrow.
Here, the worry is that identity theory attributes mental states only to creatures with human brains.
Some psychological states, like fear, seem shared by animals.
Further, we can imagine organisms that are made of completely different kinds of material but which have the same kinds of mental states that we do.
Aliens made of silicon, instead of carbon, could have pains, color sensations, and beliefs and desires, without having brains like ours.
If pain is a specific state of a brain, as the identity theorist claims, the aliens can't have pain.
This claim is generally known as multiple realizability: mental states appear to be realizable in multiple kinds of material.

We will consider three problems for the identity theorist arising from multiple realizability.
To start, let's distinguish token physicalism from type physicalism.
Token physicalism says that every instance, or token, of a mental state is identical with a token of a physical state.
Token physicalism is naturally taken as a denial of dualism.
For, if token physicalism is right, then there are no mental states that can not be explained by physical facts.

Type physicalism is the further claim that every type of mental state is identical with a type of physical state.
According to type physicalism, we will be able to find specific physical states that correspond to any mental state, like pain, or the sensation of seeing red, or the belief that aliens live on Mars.
Type physicalism is a stronger, more contentious claim than token physicalism.
Identity theory is both token and type physicalism.
Consider a list of descriptions of brain states: which neurons are firing, which have recently fired, which inputs are lit up.
Call each possible physical state of the brain S_n.
The identity theory will be made up of a series of clauses, like:

> x has a toothache iff x is in brain state S_{412}
> x is seeing blue iff x is in brain state S_{7583}
> x believes that snow is white iff x is in brain state S_{9238}
> etc.

We should be able to find the specific brain states that always correspond to toothaches and earaches and perceptions fo all colors and odors and all of our beliefs and desires.

The first problem for identity theory from multiple realizability is that we might want to attribute sensations like ours to aliens or machines or animals who do not share our brain structures.
If mental states are realizable in different kinds of brain states, the identity theory is chauvinistic.

A second and related problem is that even human brains do not all work the same way.
My brain state, when I see blue, will be different from your brain state, when you see blue.

Type physicalism seems most plausible for mental states that correspond to occurrent sensations.
But, it seems highly implausible for intentional states.
It is hard to believe that the belief that the Mets will win the World Series this year corresponds to the same exact brain state in every one who believes it.
Even for occurrent sensations, our brains have differences arising from our diverse experience and development.
So, it is reasonable to believe that there will be no single S_n to correspond to the same belief in different people, in the way that heat always corresponds to kinetic energy.
Instead of the clauses above, identity theory will have the following sorts of clauses;

> x_1 has a toothache iff x_1 is in brain state S_{412}
> x_2 has a toothache iff x_2 is in brain state S_{6224}
> x_3 has a toothache iff x_3 is in brain state S_{91}
> ...
> So, x has a toothache iff $x=x_1$ and is in S_{412} or $x=x_2$ and is in S_{6224} or $x=x_3$ and is in S_{91} or ...

We call a theory which includes clauses like the last one disjunctive, since it says that one mental state is identified with any of a variety of physical states.
The claim that human brains realize their states in different ways is called neurological equipotentiality, and is often attributed to the early-twentieth century psychologist Karl Lashley.
In support of the Lashleyan claim, consider that language is normally processed in the left hemisphere for righties, but people with damage in the left hemisphere may process language in their right hemisphere.

There is a third multiple-realizability problem closely related to equipotentiality.
The identity theorist lacks a relational construal of mental states.
A relational construal sorts mental states according to the relations among stimuli and responses.
For the identity theorist, we sort, or type, mental states according to their physical properties.
Consider how we would respond to the discovery that two disparate mental states, say a leg cramp and the belief that chocolate pudding is tasty, had the same physical instantiations, i.e. were correlated with the same brain states.
The following two claims would turn out, empirically, to be part of our theory of the mind of x.

> x has a leg cramp iff x is in brain state S_{3313}
> x believes that chocolate pudding is tasty iff x is in brain state S_{3313}

If we sort mental states according to their content, how they seem to us, it does not matter that the pudding belief and the cramp sensation are instantiated by the same brain state.
They are two strikingly different mental states.
But, for the identity theorist, we would have to say that they are the same state.
For, mental states are just brain states.
The following inference would be thus inevitable, for the identity theorist.

> x has a leg cramp iff x believes that chocolate pudding is tasty.

In other words, the mental state of having a leg cramp would be exactly the same mental state of believing that chocolate pudding is tasty.
By typing mental states according to their physical realizations, rather than by their first-person content, we leave open the possibility of making some wacky, unacceptable identifications.

Token physicalism can survive problems of multiple realizability, since even if we all have different brain states corresponding to relevantly similar mental states, they are all still physical states.
Token physicalism says, at heart, that disjunctive theories are perfectly acceptable.

The behaviorist, like the type-physicalist, is already committed to a disjunctive theory.
There are no unique behaviors that correspond to particular mental states.
Some people react to a painful stimulus by screaming, others by wincing, others by stomping about.
So, the behaviorist's identity sentences will look like: a thing is in pain iff it exhibits behaviors B_1, or B_2, or B_3, or...
If multiple realizability was not a problem for the behaviorist, maybe the identity theorist can also try a disjunctive theory, and hold on to token physicalism.
The identity theorist would then correlate pain with any of a variety of brain states, so that we can have pain-in-a-robot, and pain-in-a-Martian, and pain-in-a-chimp, and pain in a sea slug, etc.

One problem with the disjunctive approach is that it is incompatible with the idea that a given behavioral state or brain state may realize different psychological properties at different times.
The same behavior, say squinting, might be evidence of pain, or concentration, or blurry vision, or...
Similarly, we saw that the identity theorist's non-relational construal of mental states entailed that the same brain state may be correlated with different mental states.
Thus, on either disjunctive approach (the behaviorist's or the identity theorist's), we have lots of disjuncts on both sides of the equations.
If we have long disjunctions on both sides, we do not seem to be getting anywhere.
The fact that the behaviorist was liable to multiple realizability criticisms does not show that these criticisms are superable.
Rather, it shows that multiple realizability was a problem for the behaviorist as well.

VIII. Conclusions

We divided mental states into two kinds: occurrent states (sensations) and intentional states.
The identity of mental states with brain states is more plausible for occurrent mental states, for sensations, than it is for beliefs, or other representational states.
What state of the brain could count as representing my belief that tigers are dangerous animals?

In contrast, behaviorism seems more plausible for intentional states.
Beliefs and desires are at least identified by their correlative behaviors, often.
Still, the chess-playing criticism of behaviorism, that it omits the causal connections among beliefs and desires, remains.

We have now looked at three characterization of the mind:

1. Dualism: the mind is an immaterial substance.
2. Behaviorism: the mind is behavior.
3. Identity theory: the mind is the brain.

Identity theory has parsimony on its side, against dualism, and it is an empirical, scientific theory.
There are good reasons to prefer science to metaphysical speculation, where possible.
But the problems with identity theory, including its anemic analysis of consciousness and the modal and multiple realizability objections, will lead us to one last theory of the mind: functionalism.