

Class 3 - Identity Theory

I. Mental states and brain states

Behaviorism was the first contemporary materialist program.

The behaviorist identified mental states with behavior.

The sophisticated behaviorist added dispositions to behave.

So, pain is not an introspective state, but the disposition to scream, cry, wince, etc.

The behaviorists were criticized for ignoring, or denigrating, the internal workings of the mind.

Neuroscience seems to imply that some understanding of the brain is relevant to our understanding of minds.

For example, consider [transcranial magnetic stimulation](#) (TMS).

By stimulating neurons, we can alleviate depression in some patients.

That is, we can change people's mental states merely by massaging their brains.

This seems like evidence that the workings of the brain are in some deep way related to our mental states.

Of course, punching me in the gut will affect my mental states, as well.

But, the brain seems to have a deeper causal connection with my mental states.

The behaviorist's appeal to dispositions mitigates the problem of describing mental states with no attached behavior.

I can be in pain but not scream or wince.

If I have a disposition to scream and wince, then the behaviorist can ascribe to me the pain, even with no attached behavior.

But, the behaviorist maintains that any mental state must be identified with at least some disposition to behave.

In general, we can divide mental states into two kinds: occurrent sensations (qualia, or qualitative states) and intentional states.

Place accepts the dispositional account for what he calls cognitive concepts: believing, wanting intending, understanding.

These are the intentional states, which are also called propositional attitudes, or representational states: I believe that..., I want that..., I intend that..., where the hole is completed with a proposition.

(All these terms are not exactly identical, but they are close enough for our purposes.)

That is, the dispositional behaviorist's account might be satisfactory for some kinds of mental states, but not for others.

And the states with which we are most concerned in this class, the qualitative states, seem to resist dispositional interpretations.

In class, Matt brought the problem that a disposition is not a cause or a factor.

According to the dispositional theory, I don't cry because I am sad; my sadness is the disposition to cry.

I don't say that the apple looks red because I see red.

My seeing red just is my statement, and other related behaviors.

In contrast, if somebody speaks or acts in certain ways, it is natural to speak of this speech and action as the *expression* of his thought.

Even for the dispositional behaviorist, the question of why I scream when I am in pain remains without an internal account.

II. Identity theory and the problem of mental causation

The identity theorist, in contrast to the behaviorist, interprets mental states, naturally, like the dualist, as causes of behavior.

The identity theorist thinks of behavior as caused by thought, but not as identical with it, or constitutive of it.

Still, the identity theorist agrees with the behaviorist that mental states are linked to behavior.

Identity theory defines, or picks out, mental states in terms of the behaviors they cause.

The identity theorist's claim is that every mental state is strictly identical with a physical state.

(It is sometimes called brain-state theory, e.g. by Putnam, or central state theory, e.g. by Fodor; or reductionism, e.g. by Churchland.)

Identity theory simply claims that sensations are brain processes.

One obvious objection to the claim that sensations are brain processes is that the meaning of 'John has a toothache' clearly differs from the meaning of 'John has stimulation of his c-fibers'.

"Consciousness and brain processes must be independent entities because the expressions used to refer to them are logically independent..." (Place, 46).

So, the identification of pain with c-fiber stimulation can not be the kind of identification that entails sameness of meaning.

Philosophers ordinarily distinguish between the 'is' of identity and the 'is' of predication.

Certainly, the claim that sensations are brain states involves the 'is' of identity rather than the 'is' of predication.

Place further distinguishes between the 'is' of definition and the 'is' of composition.

Consider:

1. The stages of the moon cause the tides.
2. That table is a packing case.
3. The cloud is a mass of tiny particles.
4. Lightning is electrical discharge.
5. Red is a color.

The identity theorist urges us to see the identification of sensations with brain states like any other common theoretical identification in science, like 4, or like the identity of water with H₂O.

These two cases are paradigmatic theoretical identifications.

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 made of hydrogen and oxygen.

Similarly, people do not know 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, H₂O, and C-fiber stimulation).

So, the dialectic is as follows:

Claim: Consciousness is a brain process. So, 'pain' is 'c-fiber stimulation'.

Objection: The terms 'pain' and 'c-fiber stimulation' are logically independent, so they can not refer to the same object.

Response: There are other cases in which logically independent terms refer to the same object.

Place thus claims that the identities of mental states with brain states are contingent, empirical matters, the results of scientific research.

The fact that people know about their pains, and not about their c-fiber stimulations, and the fact that the meanings of the two kinds of terms seem to differ, and the fact that the expressions are logically independent entities, are all irrelevant to the identification.

III. Dualism and the phenomenological fallacy

The advantage of behaviorism over dualism was Ockhamist: the behaviorist did not posit an independent, immaterial realm.

The identity theory wants the same advantage over dualism.

But, by emphasizing sensations against the behaviorist, the identity theorist appeals to the feeling of mental states, which the dualist makes central.

Place gives us no argument against the reality of qualia, of sensations or raw feels.

Smart, in his article "Sensations and Brain Processes", argues that sensations would be nomological danglers, pp 142-3.

'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.

Place argues against the dualist by trying to eliminate the dualist's motivation, which he calls the phenomenological fallacy.

The phenomenological fallacy seems to support dualism, against identity theory.

It says that our sensations are like projections on an inner screen.

So, the after-image of green has to be the property of some inner screen.

If the mind is the inner screen, and the mind is the brain, then we have to look for some green part of the brain.

But, of course there is no green/blue/pained part of the brain, in the relevant sense.

So, it looks like consciousness could not be a brain process.

Place argues that if we remove the idea that properties of our conscious experience must be found on the mental screen, if we eliminate the idea that we have to find something other than our experience which is green, or in pain, then the idea that consciousness is a brain process remains plausible.

Unfortunately, identifying the phenomenological fallacy may not remove the difficulty.

The difference between sensations and brain processes seems to remain.

Place does not seem to provide a way to eliminate the dangler.

Note that since the identity theorist supposes brain states to be contingently identical to mental states, they need not share all properties.

This point is in their favor, since mental properties seem quite different from physical 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, then even if Smart eliminates mental states in favor of brain states, he has not eliminated mental properties of those brain states.

So, the identity theorist seems committed to nomological danglers, too.

A theory which admits irreducibly mental properties, while denying that there are mental substances, is called property dualism.

To avoid even property dualism, the identity theorist must explain mental properties without appeal to irreducible mental qualities.

Smart calls sensations topic-neutral experiences.

To say that something appears orange, or feels like pain, is just to say that it makes me feel as if I am seeing an orange, or it makes me feel as if I am in a pain state, that I am having an experience which is like other experiences.

But, he argues, there should be no presumption that this experience is an irreducibly mental experience.

To explain sensations, we can appeal merely to the material (including behavioral) connections among mental states.

Mental states are defined by their role in my perceptual experience, not by their (irreducibly) conscious qualities.

IV. Evaluating the identity theory

We divided mental states into two kinds: occurrent states (sensations) and intentional states.

The identity of mental states with brain states may be 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?

Is there some state of the brain, identical in all of our brains, that could stand for this belief?

For non-occurrent mental states, we might correlate mental states with the behaviors they produce.

Mental states would then be internal (brain) states, but they would be picked out, distinguished or sorted, by behaviors that they cause.

Such a solution to the problem of mental properties continues to be topic-neutral, and plausibly physicalistic.

We have now looked at three characterizations 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.

Science is a domain of settled facts.

Scientists make errors, of course, but consensus leads to progress.

So, we should at least try to settle some facts about the mind by seeing what science says.

Against behaviorism, the identity theory gives mental states their appropriate causal role.

But, both the materialist and the behaviorist seem to lack an account of consciousness, of the way our

mental states feel to us.

David Armstrong provides an account of the missing elements of the materialist's theory.

Considers 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.

Consciousness, then, is just thinking about thinking.

The missing elements of the identity theory are just some kinds of thoughts.

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.

We will look at more serious criticisms of identity theory in our next class, with the discussion of functionalism.

There is just one more topic which it will be useful to discuss, for later in the term.

V. Modal claims

Place claims that theoretical identities, like that of sensations and brain processes, are contingent.

Contingency is a modal property.

To understand contingent identity, it is necessary to understand a bit about modal properties.

(This topic will become important later in the term.)

The modal properties with which we are concerned are possibility and necessity .

A modal property is anything 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 (arguably) I necessarily 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 which I discuss below, in the last thirty years, involving metaphysics and the philosophy of language.

To see how theoretical identifications are 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 H₂O, 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.

Place, as we saw, uses the contingent identity of mental states and brain states to block a criticism.

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 from 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.

Place's claim that the identities are contingent allows him 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.

We will return to this criticism of identity theory later in the term, mostly in order to look at its methodology.

If you want to look more closely at the criticism, the relevant selections from *Naming and Necessity* are available on the syllabus, at Class 15.