

Philosophy 110W - 3: Introduction to Philosophy, Hamilton College, Fall 2007
Russell Marcus, Instructor
email: rmarcus1@hamilton.edu
website: http://thatmarcusfamily.org/philosophy/Intro_F07/Course_Home.htm
Office phone: 859-4056

Lecture Notes, November 6

I. Multiple realizability and identity theory

The main objection to identity theory is that it is chauvinistic: it attributes mental states only to creatures with human brains.

Some psychological states seem shared by animals.

We can imagine organisms that are made of completely different stuff, but which have the same kinds of mental states that we do.

Aliens made of silicon, instead of carbon, may have pains, color sensations, beliefs and desires, etc., and yet not have brains which are in specific ways identical to ours.

If pain is just a specific state of a brain, we must deny that the aliens have mental states.

The problem of chauvinism points more generally to a problem of multiple realizability.

Functionalism is the result of considering both behaviorism and identity theory in the light of multiple realizability.

So, it will be worth our time to explore the problems of multiple realizability in detail.

In order to focus the range of the problems arising from multiple realizability, Fodor distinguishes 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 a 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.

Identity theory is most naturally taken as type physicalism.

The 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}
etc.

Recall that the identity theorist relies on the precedents set by other, paradigmatic theoretical reductions. Mental states are brain states in the same way that heat is energy.

For any theoretical reduction, we specify essential properties of the reduced phenomena, e.g. pain, heat.

So, the heat of a gas is always the average kinetic energy of the molecules of that gas.

Similarly, according to the identity theorist/type physicalist, we should be able to find the specific brain states that always correspond to pain, and all other mental states.

A thing has a toothache iff it is in brain state S_{412} ; a thing is seeing blue iff it is in brain state S_{7583}

Type physicalism seems most plausible for mental states that correspond to occurrent sensations.

Fodor notes that identity theory suffers from problems arising from multiple realizability.

In fact, all type physicalism suffers from multiple realizability problems; see Fodor p 454-5.

If there are different brain states which can correspond to the same mental state, the type physicalist seems to be in trouble.

It seems wildly implausible that the belief that the Patriots will win the Super Bowl this year corresponds to the same exact brain state in every one who believes it.

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.

For one thing, we might want to attribute this belief to aliens or to machines, eventually, who clearly do not share our brain structures.

So, the first problem that multiple realizability raises for identity theory is its chauvinism.

1. Identity theory is chauvinistic.
2. Multiple realizability shows that chauvinism is wrong.

So, identity theory is false.

A second and related problem for identity theory arising from multiple realizability 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.

So, 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 like the one constituted by the last clause disjunctive, since it says that one mental state is identified with any of a variety of physical states.

In a different article, Fodor attributes this claim, called neurological equipotentiality, to the early-20th 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.

Fodor refers to a third multiple realizability problem closely related to equipotentiality.

He says that the identity theorist lacks a relational construal of mental states, p 455.

A relational construal sorts mental states according to the relations among other mental states.

The behaviorist provides a relational construal of mental states by classifying mental states according to 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.

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 is in brain state S_{3313}

x believes that chocolate pudding is tasty iff x is in brain state S_{3313}

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

Another way to put the point is that any theory of the mind should yield mental states that constitute natural kinds.

The issue of natural kinds evokes Goodman's new riddle of induction.

II. Natural kinds and mental states

The notion of natural kinds, and their relations to projectible predicates, comes from Goodman's solution to his new riddle of induction.

Recall that the new riddle of induction posed the question of how we know that grass is green, rather than grue.

All evidence for the greenness of grass was also evidence for its grueness.

Similarly, all evidence that this is a piece of paper is evidence that this is a papod.

One way to describe the success of predictions is to say that predicates like 'green' and 'paper' are projectible, will remain constant through time, whereas predicates like 'grue' and 'papod' are not projectible.

One account of the projectibility of certain predicates is that they refer to natural kinds.

'Green' is a natural kind; 'grue' is an unnatural, deviant philosophical construct.

Scientific theories should refer to natural kinds, but not to deviant, gerrymandered kinds.

Fodor's claim, then, is that mental states are natural kinds, and that ordinary terms which refer to our mental states should be taken as projectible predicates.

Since mental states are natural kinds, theories of the mind must refer to them, not eliminate them.

III. Disjunctive theories

We have looked at three problems for identity theory which come under the heading of multiple realizability.

1. Chauvinism

2. Neurological equipotentiality

3. Non-relational construal of mental states

These problems all apply to the type physicalist; identity theory is a version of type physicalism.

The problems of multiple realizability apply to both the identity theorist and the behaviorist.

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.

Recall that to accommodate multiple realizability, the identity theorist would have to adopt a disjunctive theory of mental states:

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 ...

Token physicalism says, at heart, that such disjunctive theories are perfectly acceptable.

Compare the disjunctive account with the behaviorist's correlations between mental states and behaviors. There are no unique behaviors that correspond to particular mental states.

Some people react to the same 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...

That is, the behaviorist is already committed to a disjunctive theory.

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, 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.

IV. Functionalism

Functionalism was designed to avoid the problems we have seen with behaviorism and identity theory, but maintain the useful insights of both theories.

The functionalist takes behaviorism's attributions of mental states based on behaviors,

According to both the behaviorist and the functionalist, we type mental states according to behavior, not according to the qualities available by introspection.

But, functionalism rejects behaviorism's disavowal of internal states, and its reductionist, eliminativist, program.

Behaviorism tried to reduce mental state language to behavior language, with the goal of eliminating any apparent references to immaterial substance.

In its effort to avoid the problem of mental causation, the behaviorists threw out the baby with the bath water: they eliminated mental causes instead of providing a way to understand them.

Consider Fodor's two Coke machines.

The behaviorist Coke machine can not accommodate internalist descriptions, ones that do not refer to

stimulus and response.

The mentalist Coke machine distinguishes states even when there are no behavioral differences.

Note that functionalism is compatible with substance dualism, since it makes no claim about where and how mental properties are instantiated.

(Fodor also discusses problems with a distinct form of behaviorism, logical behaviorism.

We have not read the logical behaviorists, but Fodor is really referring to Ryle, who I would group with later Wittgenstein and the linguistic philosophers, for these purposes.

Ryle's claim was essentially that we could accept radical behaviorism, if we supplemented it with an account of internal mental states understood as dispositions to behave.

We can distinguish (at least) two kinds of causation: dispositional causation and event-event causation.

The glass broke because it was fragile (dispositional) and because it fell on the floor (event-event).

Fodor points out that the linguistic behaviorist's account makes all mental causation dispositional.

But, really we want event-event accounts of mental causation.)

In parallel fashion, the functionalist adopts from identity theory the legitimacy of mental states and an acceptance of the causal connections among them, p 455.

The functionalist dispenses with identity theory's unacceptable chauvinism.

One good way to think about functionalism is on analogy with a computer: the mind is the software of the brain.

Mental states are computational.

(There are non-computational versions of functionalism, which we will ignore, here.)

Just as the same software can be run on different hardware, the same mental states can be instantiated by distinct physical (or, even, non-physical) systems.

The mind is not identified with any particular hardware.

A computer has certain states, some of which manifest in behavior, others of which just affect other internal states of the machine.

Each of these states is controlled by the program of the computer, its machine table.

The total state of the object (or system) will be the state of the whole system at a particular time.

For a computer, the total state will include what processes are running, what output is going to the screen or the speakers, and which switches are open and closed on the circuit board.

For a person, the total state will include memories, beliefs, desires, occurrent sensations, and conscious thoughts.

The program of the computer, its software, is instantiable in various different particular machines, the hardware.

Similarly, being in pain, or seeing blue, or believing that the moon is made of cheese, are functional states of an organism.

According to functionalism, two things are in the same mental states if, and only if, they have the same state of their programs.

Recall that we said that mental states divide into two classes: qualitative states and intentional states.

Fodor expresses a worry that qualitative states like pain do not have functional descriptions (Q14 on the Reading Guide).

In fact, qualitative states are a serious problem for functionalism, which we will discuss in a short while.

But, on the surface, the functionalist has no problem starting the account.

The challenge for the functionalist is to specify the functional role that each qualitative state plays in the causal history of the person.

That is, what is the functional role of pain?

Well, it is preceded by some causal antecedent, say a piano falling on one's toe.

It is related to other mental states, like fear and anguish.

It engenders certain behavioral responses, like jumping around and shouting, crying, or running away.

More importantly, here, consider intentional states.

Each intentional state is a relation between a believer and a thing believed, which we will call a proposition.

So, my belief that the cat is on the mat is a relation between me and the proposition that the cat is on the mat.

(Fodor oddly calls this a three-way relation among me, my belief, and the content of my belief.)

The proposition that I believe has semantic content: it is about something, or it represents something.

In particular, it is about the cat, and the relation between the cat and the mat.

Similarly, symbols, like words and sentences, have semantic content.

Thoughts are like sentences in that they are about something.

My belief (indeed any of my intentional states) are about representations.

So, intentional states are computational.

“There is no computation without representation” (p 458).

We have seen representational theories before, with Descartes, Locke, and Berkeley.

Fodor mentions Hume's characterization of the representational theory.

The contemporary heir is the computational theory of the mind.

But, instead of thinking about representation as resemblance, we think of it more generally.

There is a causal stimulus, internal relations among mental states, and behavioral responses.

Specifying these elements will yield a functional characterization of intentional states.

The functionalist provides a technical device to specify the functional role of any mental state according to its causal role in our lives.

V. Ramsey sentences and causal-role definitions of mental states

Recall that both the identity theorist and the behaviorist had problems of multiple realizability.

The functionalist avoids these problems by identifying each mental state with the relevant properties of that state, like its interactions with other mental states, and the behaviors of people in that mental state, while eliminating reference to irrelevant particulars, like brain states.

A thing is in pain iff it has been affected in certain relevant ways, and if it has other concomitant mental and behavioral states (wincing, crying), which are causally related to it.

Functionalists eliminate irrelevant vocabulary from the theoretical identity sentences of a formal theory of mental states to achieve the desired level of abstraction.

The functionalist constructs Ramsey sentences.

A Ramsey sentence removes specific references to the particular causal structures (say, brain states) at work in our mental life, and replaces them with claims that something has this causal role.

Imagine a scientific description of your whole life: your experiences, your various mental states and how they are connected, the (presumably causal) relationship between your body, including your brain, and those mental states, the resulting behavior.

Replace references to the specifically mental parts of this theory, references to pains, and color terms and beliefs, with variables.

The resulting theory provides a functional, causal-role definition of your mental states.

Pain is whatever has the place that pain has in your life.

It is preceded by physical or emotional blows, and succeeded by characteristic behavior: sometimes avoidance, and sometimes valiant confrontation.

It engenders certain other mental states, fear or anger or resignation, all of which have their own causal-role definitions.

The resulting Ramsified, functionalist theory defines mental states in terms of their functional roles.

x is in pain iff x has been affected by the kinds of things that cause pain, has other mental states that generally accompany pain, and exhibits the kind of behavior that are associated with pain.

In sorting mental states according to behavior and causal connections with other mental states, functionalism makes identity conditions on mental states very fine-grained.

Unless the machine tables of two organisms match up completely, they can never match up at all.

For, if even one state differs, it throws the whole isomorphism off.

It is unlikely that the causal-role definition of pain in any particular case will look exactly like the causal-role definition in any other case.

But, it will be similar, in many ways, especially if we take the entirety of our lives into account.

The functionalist can appeal to similarity relations among such definitions for a definition of mental states, themselves.

VI. Problems with functionalism

Type-identity statements (what it is to be a thing of a certain type) must be made with reference to the appropriate regularities.

Type-identity of elementary particles will be made in terms of charge, because, presumably, charge is an element of the basic physical laws.

Type-identity of species will be made in terms of genetic constitution.

Type-identity of water will be made in terms of molecular constitution.

Similarly, it seems that type-identity of mental states must be made in terms of psychological laws.

Recall that Fodor criticized the identity theorist for lacking a relational account of mental states.

The identity theorist sorted mental states in terms of brain states, which led to difficulties of multiple realizability.

Behaviorists also had problems sorting mental states, since they did so according to observable criteria, which do not do justice to the internal states.

The Cartesian sorts mental states in the right way, according to psychological regularities which hold among our mental states.

But, since we lack third-person access to the dualist's mental states, the Cartesian lacks key elements of a scientifically legitimate theory: verifiability, replicability, etc.

Functionalists seem to have an advantage over these other positions, because they sort mental states according to their causal roles.

Still, functionalism has deficiencies.

The most serious problems with functionalism concern the account of qualia, as Fodor notes, p 457.

The first criticism is called the problem of inverted qualia.

It appears in Locke's *Essay* (Book 2, Chapter 32, §15).

The general idea is that two people could be identical in their behavior, and indeed in their functioning, and yet not share the same phenomenal experience.

A variation of the problem arises from mere differences in physiology.

My eyes are perhaps a bit bigger or smaller than yours.

Perhaps you have more rods or cones, which are the physical basis for color perception.

Why should I believe that my sensation of red matches yours?

In fact, since I am color blind, we have no reason to believe that we have the same perceptions.

But, we do not share the same functions, either.

The problem arises for two people who do see all colors.

One person's experience might be more vibrant, or brighter, or slightly shifted to the left.

The more startling problem is how to understand the status of inverted qualia.

Take two normal sighted people, who agree on a whole range of color ascriptions.

What if every time one saw red, the other saw purple; every time one saw blue, the other saw green?

They could still use the same terms; they would be functionally isomorphic.

But, they would be having different qualia.

The problem for functionalism is that if there are cases of inverted qualia, then people with the same functional states are in different mental states.

And, there seems to be no way to deny the possibility of inverted qualia.

So, functionalism fails to capture the nature of our mental states.

Further, the situation can be even worse.

David Chalmers has written in defense of property dualism from considerations of zombies.

Zombies are organisms which function just as we do, but which have no phenomenal experience.

This is also known as the problem of absent qualia.

While the possibility of zombies might seem outlandish, the privileged access we seem to have to our mental states eliminates any possibility of ruling zombies out.

Another absent qualia argument, due to Ned Block, involves the Chinese nation.

The brain is essentially a collection of neurons, which discharge impulses from one to another.

Neurons fire, and induce other neurons around them either to fire or not to fire.

The story is more complicated, of course, but the differences appear only to be a matter of degree, not of kind.

The basic picture of neurons transmitting information like electrons passing along a circuit board is apt.

Imagine that we have mapped the brain, and it contains 1.3 billion neurons.

This is a fiction, but only by a factor of about a hundred - there are about a hundred billion neurons in the brain.

Now, we can set up the people of China to act as this billion-neuron brain.

We can give each person the instructions to act as a neuron does, transmit information in the way that our neurons do, to other people.

Essentially, we make a mock brain out of the Chinese nation.

The brain can be attached to a human sensory organs via radio signals from the receptor nerves.

That is, we would, according to this thought experiment, remove a person's brain, and attach an artificial processing system made out of China.