

4. Identity Theory

Martín Abreu Zavaleta

June 2, 2014

1 Introducing Identity Theory

Smart wants to focus on a certain kind of mental states, namely, what we may now call *conscious* states. Examples of these kinds of mental states are all sorts of sensations, like *pain*, or Smart's "roundish, blurry-edged after-image which is yellowish towards its edge and orange towards its centre." (Smart, p. 60) ¹

His main theses about these sorts of mental states is a version of what some people later called *Identity Theory*. The basic tenet of the view is that mental states *are just* certain kinds of physical states. Strictly speaking, the view can be stated in more general ways—for instance, that all mental properties are just physical properties—but let's focus on the view that Smart has in mind.

Smart thinks that mental states—he only talks about sensations, but he must think that his claim can be generalized to all sorts of mental states—*are just* brain processes of a certain kind. He says:

Let me first try to state more accurately the thesis that sensations are brain processes. It is not the thesis that, for example, "after-image" or "ache" means the same as "brain process of sort X". It is that, in so far as "after-image" or "ache" is a report of a process, it is a report of a process that *happens to be* a brain process. It follows that the thesis does not claim that sensation statements can be *translated* into statements about brain processes [...] All it claims is that in so far as a sensation statement is a report of something, that something is in fact a brain process. (p. 62)

When I say that a sensation is a brain process or that lightning is an electric discharge, I am using "is" in the sense of strict identity. When I say that a sensation is a brain process or that lightning is an electric discharge I do not mean just that the sensation is somehow spatially or temporally continuous with the brain process or that the lightning is just spatially or temporally continuous with the discharge. (p. 62)

Smart considers some objections to these view. Here we will consider the most interesting ones.

An objection from a priori knowledge

Argument: We can make claims about our sensations without having any knowledge of our brain processes. But if sensations were just brain processes, we would know that they are. Thus, the things we talk about when we describe our sensations are not brain processes.

¹He starts by considering the view that the sentences we thought we used to report or describe such mental states don't really report anything. We won't examine this view.

One way to reconstruct the argument is this:

1. If A and B are the same, then this must be known a priori.
2. For any brain process x and sensation y , we don't know a priori that x is the same as y ; therefore,
3. For any brain process x and sensation y , x and y are not the same thing.

Questions: What do you think about this argument? Is it a good reconstruction of the intuitive characterization? Do you think it's related to Descartes' argument that he is not his body in the second meditation? Why?

For our purposes, it will be worth examining this argument more closely. By now you should understand the notions of necessity and possibility, and in particular the notion of *metaphysical necessity*. You may have also noticed that the argument above requires the notion of something being known *a priori*. What is this? The answer is still controversial, but you can take the following as our working definition:

A priori: Something is knowable a priori (for a subject S) just in case it can be known (by S) without the aid of empirical information.

We should be careful on how to understand empirical information here. For instance, we may think that coming to understand a language requires empirical information, but linguistic knowledge is not the kind of thing that friends of the a priori want to talk about. What they want to rule out are things that you come to know through your senses, like the fact that the grass is green, or that your hair is of a certain color, or even that you look the way you do.

In the past (when Smart wrote his paper), people used to confuse *aprioricity* with *metaphysical necessity*. Though Smart himself seems to make this mistake, he also offers an interesting consideration in response to the argument:

Consider lightning. Modern Physical science tells us that lightning is a certain kind of electrical discharge due to ionization of clouds of water-vapor in the atmosphere. This, it is now believed, is what the true nature of lightning is. Note that there are not two things: a flash of lightning and an electrical discharge. There is one thing, a flash of lightning, which is described scientifically as an electrical discharge to the earth from a cloud of ionized water-molecules. The case is not at all like that of explaining a footprint by reference to a burglar. We say that what lightning really is, what its true nature as revealed by science is, is an electric discharge. (p. 63)

Phenomenal properties

Smart attributes the following argument to Max Black: even if the objection above doesn't show that sensations are not brain processes, it shows that

the qualities of sensations are something over and above brain processes. That is, it may be possible to get out of asserting the existence of irreducibly psychic processes, but not out of asserting the existence of irreducibly psychic *properties*. For suppose we identify the Morning Star with the Evening Star. Then there must be some properties

which logically imply that of being the Morning Star, and quite distinct properties which entail that of being the Evening Star. Again, there must be some properties (for example, that of being a yellow flash) which are logically distinct from those in the physicalist story.

Think about the planet that was known long ago as the evening star and the morning star. Even though the actual evening star and the actual morning star are in fact the same thing, there is a sense in which the *property* of being the morning star is not the same as the *property* of being the evening star.

The first is the property of being the last celestial body visible in the morning. The second is the property of being the first celestial body visible in the morning. These two are clearly different properties, but as it happens, they are actually instantiated by the same celestial body (Venus).

Black's point is somewhat similar: we may accept that sensations *are just* brain processes. That is, whenever we experience pain, say, it's not the case that there are two different things, one of which is a sensation, and another which is a brain process. Black's point is that even if we accept this, we can insist that there are two different kinds of *properties*, one which is purely mental, and some others which are completely physical.

In particular, even though our sensation might be a brain process, it has a property that is not physical, to wit, there is something the sensation feels like. For instance, when we experience a pain, our sensation has the property of feeling a certain way. This property, Black claims, is not a physical property.

The view that Smart attributes to Black is usually called *property dualism*. It claims that there may be only one kind of substance, but there are at least two kinds of properties. For some reason, Smart claims that Black's view is *epiphenomenalism*, which we will discuss later.

2 Separating meanings

Some people, especially by the time Smart, Feigl and others introduced the identity theory, conflated not only metaphysical necessity with apriority, but they also conflated these two with *analyticity*. What is this? The answer is also a matter of controversy, but roughly, philosophers agree that a sentence is *analytic* if it is true in virtue of its meaning. Today, some people deny that any sentence is analytic, but a standard definition of analyticity is this, due to Paul Boghossian:

Analyticity: A sentence *s* is analytic just in case knowledge that *s* is true is required by competence in its use.

The notion of analyticity has been recently attacked, but we need not worry about this for now.

With the notion in place, we can see what the first argument was trying to get at: some people in the past thought that, since identities were necessary, any true identity statement had to be analytic. Since the statement of identity between sensations and brain processes is not analytic, this must mean that it is not true, or so these people thought.

This is why Smart and Feigl make an extra effort to distinguish their view from the view that sensation-words and brain-process-words mean the same. Indeed, it is part of their view that the identity statements in question are known only with the aid of experience, and so, they are neither a priori nor analytic:

The “mental” states or events (in the sense of raw feels) are the referents (denotata) of both the phenomenal terms of the language of introspection, as well as of certain terms of the neurophysiological language [...] The identification, I have emphasized, is to be empirically justified, and hence there can be no logical equivalence between the concepts (or statements) in the two languages.

Feigl goes on as to how we may discover that these identities hold:

The empirical character of the identification rests upon the extensional equivalencies, or extensional implications, which hold between statements about the behavioral and the neurophysiological evidence. In our example this means that all persons to whom we ascribe an after-image, as evidenced by certain stimulus and response conditions, also have cerebral processes of a certain kind, and vice versa. (Feigl, pp. 69-70)

3 Arguments for Identity Theory

Smart offered one of the first arguments in favor of the identity theory (or, if you want, against dualism). It is an argument from the simplicity of the identity theory and the complexity of dualism:

Dualism involves a large number of irreducible psychophysical laws of a queer sort, that just have to be taken on trust, and are just as difficult to swallow as the irreducible facts about the paleontology of the earth with which we are faced on Philip Gosse’s theory. (Smart, p. 67)

The main intuition behind the argument is pretty simple, and we may even trace it to Princess Elisabeth. You may recall that at some point in her correspondence with Descartes, Elisabeth said that she would find more plausible to identify the mental with the physical than to understand the kind of links that would allow the mental to cause physical events, and vice versa.

Smart is saying something similar: interactionist dualism posits some laws linking the mental with the physical without giving us any explanation of them. Following standard scientific methodology, this is not a theory we should take seriously, just like we shouldn’t take seriously Philip Gosse’s theory about the beginning of the universe. Sometimes we don’t have direct empirical evidence for a theory, but that doesn’t mean that we can’t decide between it and its competitor.

Most arguments for identity theory appeal to this general intuition. Let’s see one more such argument, often called the argument from *overdetermination*