

A Romp Through the Philosophy of Mind

Session One: Identity Theory And Why It Won't Work

Marianne Talbot

University of Oxford

26/27th November 2011

Session One: Identity Theory And Why It Won't Work

Session Two: Non-reductive Physicalisms And The Problems They Face

Session Three: If Physicalism Won't Work What Is The Alternative?

Session Four: Are We Asking The Wrong Questions?

Session Five: Question And Answer Session

Session One: Identity Theory And Why It Won't Work:

- **What is the context in which Identity Theory was postulated?**
- **What is Identity Theory?**
- **Why is Identity Theory attractive?**
- **Why should we reject Identity Theory?**

The Context of Identity Theory:

From the mid-1600s until the early 1900s...

... Cartesian Dualism was the 'in' theory...

...it holds that there is a 'real distinction' between mental states and physical states...

...that they *cannot* be identical because...

... their essential properties are utterly different



So Descartes believed that the essence of the mind was the various modes of *thinking*...

...and today we might add that there are mental states whose essence is to have some *quality*...

...whilst the essence of 'body' consisted in various modes of *extension*

It seems obvious that if two states differ in their properties then they cannot be identical to each other...

...so so long as we believed that mental states and physical states...

... differed with respect to their properties – indeed their *essential* properties...

...we believed they must be states of two different kinds

Let's have a look at *why* we might think that mental states are essentially different from physical states.

There is undoubtedly a problem with the belief...

... that mental states are not physical states...

...if mental states are not physical then...

... *how* could they possibly causally interact with physical states?

In the 20th Century this problem started to become intolerable...

...and so the science of psychology was born...

...and with it the attempt to show that mental states...

... are not in fact so very different from physical states...

...indeed mental states – somehow or other – *are* physical states

What is Identity Theory?

The aim of the Identity Theorist is to discover empirical evidence for 'bridge laws' such as:

- Pains are identical to C-Fibre Firings (CFF)
- Believing P is identical to an activation of neural state N (NSN)

for *all* mental states, so *reducing* the mental to the physical...

...Identity Theory is also known as *Reductive Physicalism*

It is important to note that the notion of
'identity' being used here is that of *numerical* not
qualitative identity

- If we say two dresses are identical we mean the dresses are alike in all their qualities, but they are still *different* dresses



- When we say Hesperus is identical to Phosphorus we mean that the planet named by ‘Hesperus’ is *the very same planet* as that named by ‘Phosphorus’



- Two dresses are *qualitatively* identical
- Hesperus is *numerically* identical to Phosphorus

Identity Theory believes that mental state types, such as pains or beliefs that P, are *numerically* identical to physical state types such as CFFs or NSNs

Why is Identity Theory attractive?

Reason One:

- Some of our behaviours seem straightforwardly to be function of events in our environment, others are explicable only on the assumption that they are a function of the way the world *appears* to us rather than the way the world *is*
- This is explicable on the hypothesis that inside our heads there are events that are sometimes a function of events external to us, and sometimes a function of something else: we might think of these events as meaning '*it appears to be the case that P*' which is true if it has been caused by an instance of P and false otherwise.
- We know what goes on inside our heads: neural states activate in response to environmental stimuli and in doing so they causally impact on our behaviours

Reason Two:

We dearly want to think of the mental as causally efficacious

We think of physics as causally closed (i.e. causation does not ‘pop out’ of the physical realm and then come back in again)

Given this, if mental states are not physical states this makes mental causation look problematic

If mental states are physical there is no problem with mental causation

Reason Three:

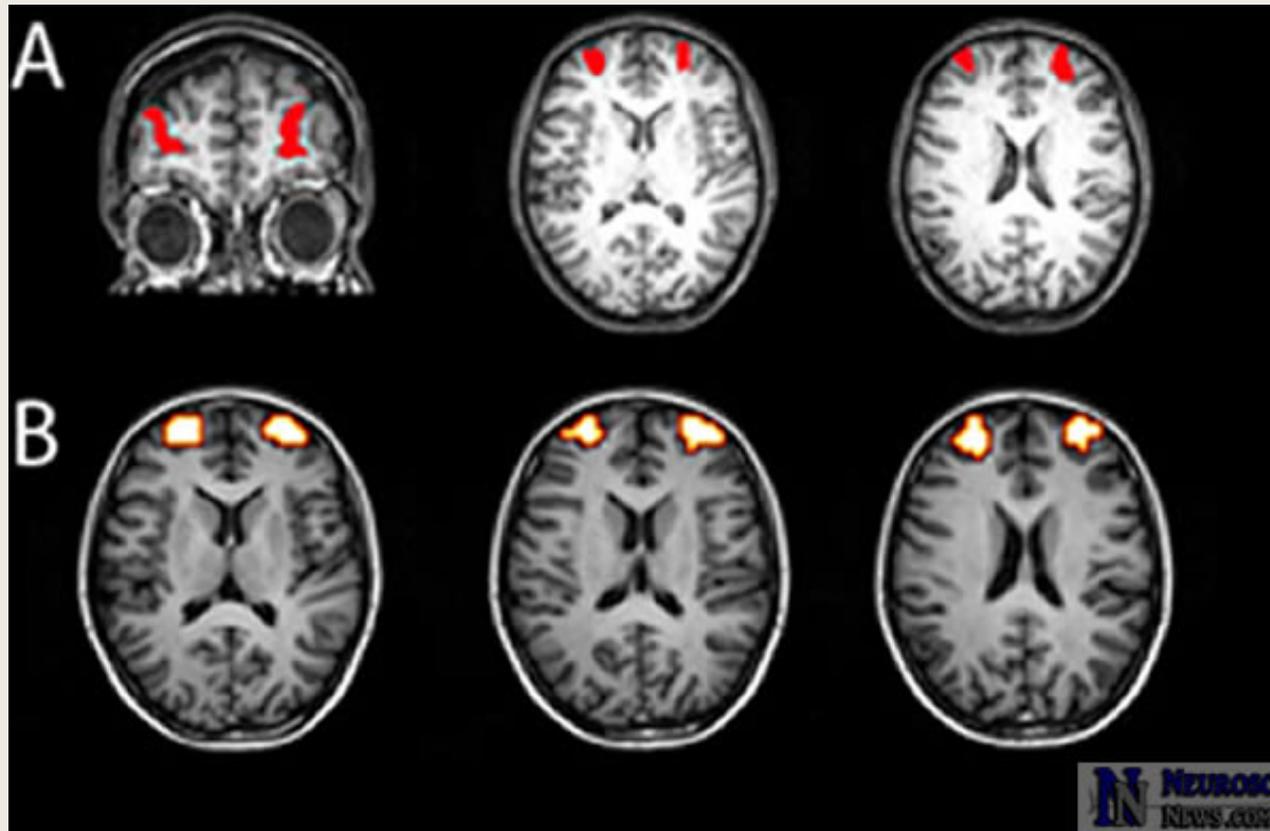
Ockham's razor tells us that we should not multiply entities unnecessarily

Reducing the mental to the physical immediately halves the number of entities we must admit into our ontology

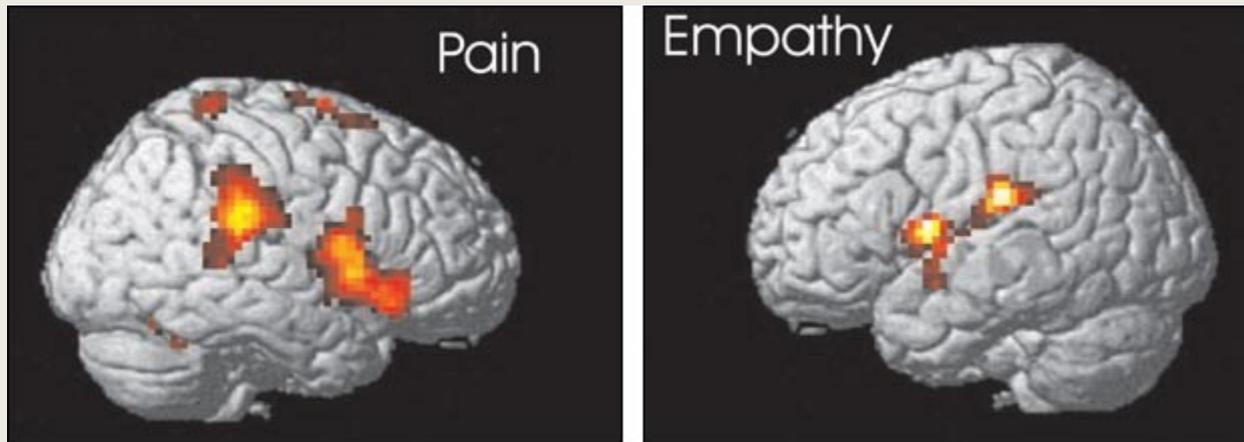
Reason Four:

Science appears to have already found many correlations of precisely the sort we'd expect if mental states *are* physical states

A picture of an fMRI scan showing the subject's thoughts (neurosciencenews.com)



Another fMRI scan showing not only a subject in pain but another subject empathising with someone in pain (msnbc.msn.com)

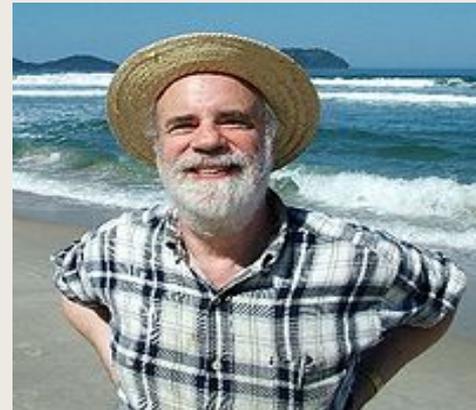


**So we have both philosophical *and* scientific reasons
to embrace Identity Theory**

Why should we reject Identity Theory?

Identity Theory was suggested in the 1950s and early 1960s

In the early 1970s it was blown out of the water by the logician Saul Kripke



It was one of the shortest-lived theories in philosophy

Kripke's argument against Identity Theory

Premise one: Numerical identity is a logically *necessary* relation: if $a=b$, then $a=b$ in *every possible world*

Premise Two: Logic tells us that the relation between mental and physical states is *not* a logically necessary relation

Conclusion: The relation between mental states and physical states is not that of numerical identity

This argument is valid (if its premises are true its conclusion *must* be true)

If we want to reject it, therefore, we must question its premises

Why should we believe Premise One (that numerical identity is a logically necessary relation)?

Answer:

It is a law of logic that everything is numerically identical to itself

It is the famous ‘law of identity’: $a = a$

There is no questioning this law (at least you can try....)

Why should we believe Premise Two (that logic tells us that the relation between mental and physical states is not a logically necessary relation)?

Another law of logic, Leibniz's Law (or the *indiscernibility of identicals*), tells us that if $a = b$, then any property that a has will also be a property that b has

Note: we must relativise this to time to allow for persistence over time

Leibniz's Law tells us that if pain states are numerically identical to CFF, and beliefs that P are numerically identical to NSN, then:

- **any property possessed by a pain will also be a property possessed by CFF,**

and

- **any property possessed by a belief that P will also be a property possessed by NSN**

This means that if pain=CFF and the belief that P = NSN, then:

- there couldn't be *anything* with the property of being a pain that does not have the physical properties of CFF and nor could there be *anything* that is a CFF that isn't also a pain
- there couldn't be *anything* with the property of being the belief that P that does not have the physical properties of a NSN and nor could there be *anything* that is an activation of NSN without being the belief that P

i.e. it would be simply *impossible* for a pain state to exist that isn't a CFF, and similarly *impossible* for there to be beliefs that P aren't NSNs

Problem:

- we have good empirical reason to believe that even *this* is a world in which there are pains that are not CFF: after all are dogs' pains correlated with CFF?
- we simply don't believe that there *couldn't* be a world in which there are beliefs that P that are not NSNs: after all if there are aliens, physically unlike us but mentally similar, why couldn't *they* believe P?

When we consider the properties peculiar to pains and to beliefs...

...the properties without which these states wouldn't be the states they are...

...these properties are not *necessary* to any particular physical state...

....in the way they *are* necessary to mental states

It is simply a fact, empirically established (in the case of sensations like pain) that mental states are *multiply realisable* in that *one and the same type of mental state* can be correlated with *different* physical states

It is also a fact, established by conceptual analysis, that we would attribute beliefs that P to an alien if doing so was the only way to make sense of his behaviour, and we would do this quite irrespective of his physical make-up

Pains simply cannot be numerically identical with CFF if there can be pains that are not CFF or CFF that are not pains

Beliefs that P cannot be numerically identical with activations of NSN if there can be beliefs that P that are not activations of NSN or vice versa

Therefore Identity Theory is FALSE

It is interesting to note...

...that Kripke's argument is an updated...

...and logically grounded...

...version of Descartes' argument for Dualism

Eh?

That all happened rather quickly didn't it?

Surely there is something wrong with this argument given how attractive Identity Theory is?

But if so, how might we object to Kripke's argument?

Objection One:

Even if we can't find bridge laws that are entirely general mightn't we find bridge laws that are *species-specific*?

Response to Objection One:

We'd still have the problem of unusual species-members: if someone seemed from everything he said and did over time (and after covert scrutiny) to be in pain yet to lack CFF would we continue to insist that he couldn't be in pain?

We wouldn't insist our alien *couldn't* believe P would we, or even that great apes (say) *couldn't* believe P?

Objection Two:

We might want to say that the Law of Identity does not *always* hold.

For example: ‘Marianne Talbot = the Director of Studies in Philosophy at OUDCE’ is an identity statement that, though true in *this* world does *not* hold in every possible world.

There are, after all, worlds in which Marianne Talbot is *not* the DoS in Philosophy at OUDCE: any world where someone else got the job.

Why then shouldn’t there be worlds in which pain is not identical to CFF, even though in this world pain IS CFF (and ditto for beliefs)?

Response to Objection Two:

There are indeed identity statements that are not logically necessary (that are *contingent*) but all of them are flanked by at least one *non-rigid designator*.

Non-Rigid Designator: designator that names different things in different possible worlds and that can't therefore track individuals across different possible worlds

Rigid Designator: designator that names the same thing in every possible world and that can therefore track individuals across different possible worlds

If we ask 'could the DoS in Philosophy at OUDCE be anyone other than Marianne Talbot?' the answer is clearly yes because there are worlds in which Marianne Talbot (that very person) exists but *isn't* the DoS in Philosophy at OUDCE

But we know this *only* because ‘Marianne Talbot’ tracks me into worlds where I am *not* the DoS of Philosophy at OUDCE

If an identity statement is flanked by two rigid designators then it *is* necessarily true – true in every possible world

This is because if $a=b$, and ‘a’ and ‘b’ are rigid designators, then if ‘a’ names something in this world that is identical to the thing named by ‘b’ then the thing named by ‘a’ will be identical to the thing named by ‘b’ in every other world too

This means that so long as ‘pain’, ‘CFF’, ‘belief that P’ and ‘neural state n’ are rigid designators Kripke’s argument holds and Objection Two fails

Question:

Are these words rigid designators? Do they name the very same thing in every world in which that thing exists? Can we use these words to track states through different possible worlds?

If not we *can* defuse Kripke's argument.

- Could pains be states that do not feel awful?
- Could CFF or neural state n be states with different *physical* properties?
- Could beliefs that P be states with a content other than P?

Objection Three

We might want to insist that there are identity statements that *appear* to be contingent though they are in fact necessary (i.e. it seems to us possible that they might be false even though they are in fact necessarily true)

For example 'water = H₂O' appears to be contingently true despite being necessary

It appears contingent because it seems possible that it might have turned out that water *wasn't* H₂O

It is nevertheless necessary because given that science has demonstrated that water = H₂O, water is necessarily H₂O

Might 'pain=CFF' ('beliefs that P = NSN') be identity statements that *appear* to be false even though they are in fact necessarily true?

Response to Objection Three:

The identity statement 'water = H₂O' appears contingent (i.e. it seems to be the case it might have been false) because we can imagine being in a situation exactly similar to one in which we experience water but where there is no water to experience (because there is no H₂O).

Is there *any* imaginable situation exactly similar to that of being in pain but where we are *not* in pain?

Is there any imaginable situation exactly similar to that in which we believe P but where we don't believe P?

If not then 'pain=CFF' and 'the belief P =NSN' are not identity statements even *appears* to be contingently false, both appear to be *necessarily* false

And surely it is not possible to imagine either of these things?

There is no gap between appearance and reality in the case of mental states in the way there is with respect to natural kinds like ‘water’.

This means that Kripke’s argument holds and Objection Three fails

Question: Can we imagine situations in which it is with us exactly as if we are in pain yet we are not in pain?

If so we *can* defuse Kripke's argument by means of such situations.

Summary:

Identity Theory is extremely attractive not least because there seems to be copious empirical evidence for it in the correlations science discovers between mental states and physical states.

Logic tells us that the empirical discovery of a *correlation* cannot be taken to be evidence for an *identity* so long as the items said to be identical have different properties.

Conceptual analysis tells us that mental states and physical states have properties that are quite different

It doesn't matter how many correlations between mental and physical states science discovers, none of them can be taken as evidence for an *identity* between mental and physical states unless Kripke's argument can be defused

References:

- Heil, J:** **Philosophy of Mind: A Contemporary Introduction** chapters 1-9
- Chalmers D:** **Philosophy of Mind: Classical and Contemporary Readings** (2002 Oxford University Press)
- Descartes, R:** **The Meditations on First Philosophy** (excerpt) in Chalmers p.10
- Place, U.T:** ‘Is Consciousness a Brain Process’ in Chalmers p.55
- Smart, J.C.C:** ‘Sensations and Brain Processes’ in Chalmers p.60
- Kripke, S:** **Naming and Necessity** (excerpt) in Chalmers p. 329
- Lewis, D:** ‘Psychophysical and Theoretical Identifications’ in Chalmers p. 88
- Block, N:** ‘Troubles with Functionalism’ (excerpt) in Chalmers p. 94
- Davidson, D:** ‘Mental Events’ in Chalmers p. 116
- Kim, J:** ‘The Many Problems of Mental Causation’ (excerpt) in Chalmers p. 170
-
- Millican, P:** Podcast: <http://www.youtube.com/watch?v=7bIS3oRb6ag>