

The Mind-Brain Identity Thesis

From Correlation to Identity: Arguments, Interpretations, and Objections

Overview

I. The Identity Thesis

What does it mean to claim that mental states are brain states?

II. Arguments for the Identity Thesis

Simplicity, reduction, and mental causation

III. Interpretation of the Thesis

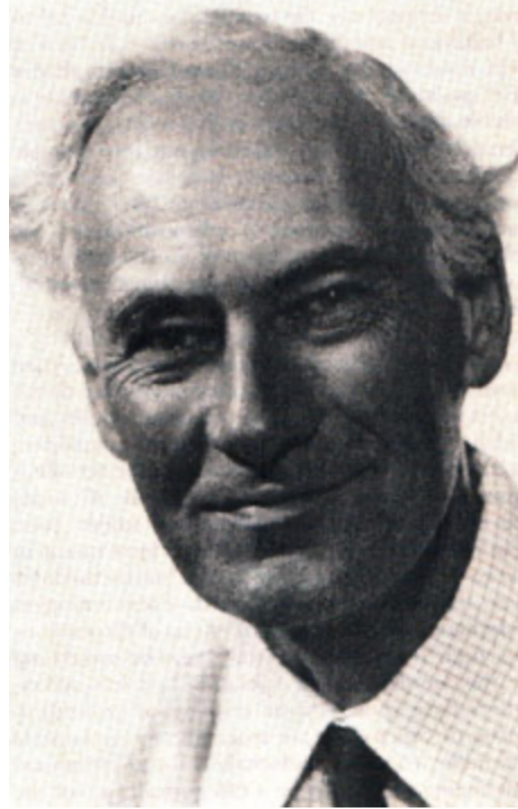
Token vs. type identity; the nature of events

IV. Putnam's Rebuttal: Functionalism

Multiple realizability and the functional state hypothesis

Part I

The Identity Thesis



J. J. C. Smart
1920-2012

Mind-Brain Correlations

The starting point: Empirical correlations

- Brain states can be measured via fMRI, EEG, MEG, intracranial recordings.
- Mental states can be measured via behavior and introspective reports.
- The NCC (Neural Correlates of Consciousness) research program seeks "the minimally sufficient neural conditions for a specific content being conscious."
- Key finding: The fronto-parietal network is consistently implicated in both levels of consciousness and contents of consciousness.

The Identity Thesis: Definition

"It is not the thesis that '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. [...] Sensations are nothing over and above brain processes."

— J.J.C. Smart, "Sensations and Brain Processes"

What the Identity Thesis Claims

The core claim:

Mental states are identical to brain states — not merely correlated with them.

Analogy with scientific identities:

- Water = H₂O molecules
- Lightning = atmospheric electrical discharge
- Temperature = mean molecular kinetic energy

Similarly:

- Pain = C-fiber stimulation (or some brain state)

What the Thesis Does NOT Claim

The identity thesis does NOT claim:

- That 'pain' means the same as 'brain state X' (semantic equivalence)
- That sensation statements can be translated into brain-state statements
- That the logic of mental discourse is the same as that of physical discourse

The identity is empirical (a posteriori), not semantic or definitional.

Part II

Arguments for the Identity Thesis



Mies van der Rohe, Barcelona Pavilion

The Simplicity Argument (Occam's Razor)

"Science is increasingly giving us a viewpoint whereby organisms are able to be seen as physicochemical mechanisms. [...] There does seem to be nothing in the world but increasingly complex arrangements of physical constituents. All except for one place: in consciousness. That these should be correlated with brain processes does not help, for to say that they are correlated is to say that they are something 'over and above.'"

— J.J.C. Smart

Ontological and Theoretical Simplicity

Ontological simplicity:

Identity reduces the number of entities in our ontology.

If pain = brain state X, we don't need two separate things.

Theoretical simplicity:

Identity enables theoretical reduction.

The language of sensations may (in principle) reduce to the language of brain states.

Important distinction:

Reduction $\neq$ Elimination. A reduced theory (like Kepler's laws) is not false — it's conserved within a more fundamental framework.

The Mental Causation Argument

The problem of mental causation:

How can mental states (beliefs, desires) cause physical events (bodily movements)?

The identity solution:

- If mental states are brain states, the mind is a proper part of the body.
- Mental causation becomes brain causation — a scientific problem, not a metaphysical mystery.
- No causal overdetermination is required.

Advantage: Physical causal closure is preserved.

Neuroscientific Support: Memory Systems

Identifying neural mechanisms can refine psychological classifications.

Case study: H.M. (Henry Molaison, 1926-2008)

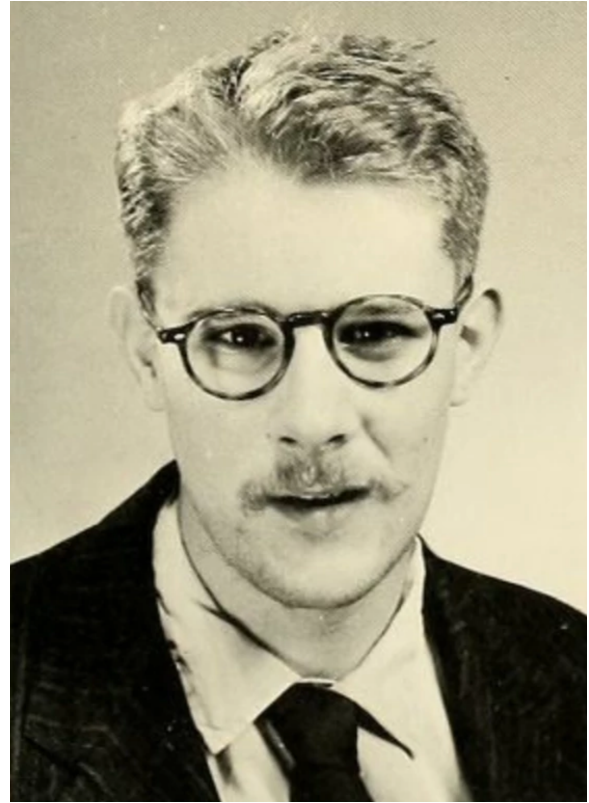
- Medial temporal lobe removal → profound amnesia for new episodic memories
- Short-term (working) memory was spared
- Procedural memory also remained intact

Double dissociations:

- Hippocampal lesions → impaired episodic memory, intact working memory
- Prefrontal lesions → impaired working memory, intact episodic memory

Part III

Interpreting the Identity Thesis



David Lewis
1941-2001

Identity of Meaning vs. Metaphysical Identity

A priori identities (analytic):

"The bachelor = the unmarried man" — justified by meaning alone.

A posteriori identities (empirical):

- "Water = H₂O" — not justified by meaning, but by scientific discovery.
- "Romain Gary = Émile Ajar" — empirical fact, not semantic truth.

Mind-brain identities are empirical, a posteriori:

"An illiterate peasant might talk about sensations without knowing about brain processes, just as he can talk about lightning though he knows nothing of electricity." (Smart)

Necessity and A Posteriority

Smart's original view:

Mind-brain identities are contingent (could have been otherwise).

After Kripke's Naming and Necessity (1972):

- True identity statements are necessarily true.
- "Water = H₂O" is necessary — water couldn't have been something else.
- Yet this necessity is discovered a posteriori.

Implication:

If pain = C-fiber stimulation, this is necessarily true, though knowable only empirically.

Token Physicalism vs. Type Physicalism

Token physicalism:

Every particular mental event is identical to some particular physical event.
But mental properties need not be identical to physical properties.

Type physicalism:

Mental types (properties, kinds) are identical to physical types.
"Being in pain" = "Having C-fibers firing" (as properties).

Logical point:

Token physicalism does not entail type physicalism.

Analogy: Every colored object has a shape, but colors $\neq$ shapes.

Two Views of Events

View 1: Events as concrete particulars

Events are basic entities that can instantiate many properties.

One pain event can be: painful, throbbing, lasting, a brain event, a C-fiber event.

Identity theory: mental and physical properties characterize the same event.

View 2: Events as property instantiations

Event = object x instantiating property P at time t.

Event identity requires property identity.

On this view, identity theory logically implies type physicalism.

Lewis: Reconciling Folk Psychology and Identity

David Lewis's key insight (1972):

Theoretical identifications are not merely posited for parsimony — they are implied by theories.

Lewis's two-step argument:

- (i) Folk psychology implicitly defines mental terms by their causal roles.
- (ii) Neuroscience discovers which physical states occupy those roles.

The identification follows logically:

Mental state M = the occupant of causal role R (by definition, from folk psychology)

Neural state N = the occupant of causal role R (by physiological theory)

∴ Mental state M = neural state N (by transitivity of identity)

Folk Psychology as a Term-Introducing Theory

Lewis's method:

Collect all the platitudes about mental states — their causal relations to stimuli, responses, and each other.

Example platitudes:

- "If someone suffers tissue damage and is alert, they tend to be in pain."
- "If someone is in pain, they tend to wince, groan, and desire relief."
- "Toothache is a kind of pain."

Mental terms are T-terms:

Names of mental states are functionally defined by the causal roles specified in these platitudes. They name whatever occupies those roles.

Lewis vs. Smart: Two Routes to Identity

Smart's argument (Occam's Razor):

Identifications are posited for parsimony — "made, not found."

We choose to identify mental and brain states to simplify our ontology.

Lewis's argument (Functional Definition):

Identifications are deduced, not posited — they follow from meanings + facts.

The meanings of mental terms (from folk psychology) + empirical discoveries (from neuroscience) logically compel identification.

Advantage of Lewis:

Psychophysical identity is not optional; physiology and the meanings of words leave us no choice.

Part IV

Putnam's Rebuttal: Functionalism



Hilary Putnam
1926-2016

Putnam's Approach

Putnam's question: Is pain a brain state?

His strategy:

- Not to argue a priori that pain cannot be a brain state.
- Rather, to argue empirically that another hypothesis is more plausible.

His alternative hypothesis:

Pain is not a physical-chemical state of the brain, but a functional state of a whole organism.

This is the foundation of functionalism in philosophy of mind.

The Multiple Realizability Intuition

Consider the variety of pain-capable organisms:

Humans, primates, mammals, birds, reptiles, amphibians, octopi...

The identity theorist's challenge:

There must be a single neurophysiological kind that is instantiated every time any organism feels pain.

Putnam's objection:

It's implausible that such diverse organisms share one neural kind.

Pain is likely multiply realizable: different physical states in different species.

The Multiple Realizability Argument

Premise 1:

All mental kinds are multiply realizable by distinct physical kinds.

Premise 2:

If a mental kind is multiply realizable, it cannot be identical to any specific physical kind.

Conclusion:

No mental kind is identical to any specific physical kind.

Putnam's picture: Pain is realized by N_h in humans, N_r in reptiles, N_m in Martians — no single physical kind.

The Burden on the Identity Theorist

What the brain-state theorist must demonstrate:

Specify a physical-chemical state such that any organism is in pain if and only if:

- (a) It has a brain of suitable physical-chemical structure
- (b) Its brain is in that physical-chemical state

This state must be:

- Possible for mammalian, reptilian, and mollusc brains (octopi feel pain)
- Impossible for any creature that cannot feel pain
- Universal across any extraterrestrial life capable of pain

The Functional State Hypothesis

Putnam's alternative:

Being in pain is a functional state of a whole organism.

Key elements:

- All pain-capable organisms are Probabilistic Automata.
- Pain-capable organisms have appropriate Functional Organization.
- Pain = a state defined by its role in a Machine Table (inputs, outputs, transitions).

Functional states are specified implicitly — by their causal relations, not by their physical realization.

Why Prefer Functionalism over Identity Theory?

Behavioral evidence:

We identify organisms as in pain based on behavior.

Behavioral similarity suggests functional similarity — not physical similarity.

Cross-species psychological laws:

We expect basic states (hunger, thirst, pain) to have similar transition probabilities across species.

This follows from how we identify these states, not from neural similarity.

The functionalist hope: species-independent psychological laws.

The identity theorist's hope: species-independent neurophysiological laws (less plausible).

Functionalism vs. Behaviorism

Why not behavior dispositions?

Problem 1: Circularity

The disposition can only be specified as "the disposition to behave as if in pain."

Problem 2: Counterexamples

- Two organisms with cut motor nerves: same behavior (none), but one may feel pain.
- One person with cut pain fibers, another suppressing pain: same behavior, different experiences.

Functionalism can specify pain without using 'pain' — by its role in functional organization.

Is Multiple Realizability Empirically Supported?

Challenge to MR:

Science-fiction examples (silicon Martians) are question-begging.

We need actual evidence of multiply realized mental states.

Counter-evidence:

- Animal models of brain function reliably generalize to humans.
- Striking parallels in brain architecture across primates (recent research).



Rewired ferrets (Sharma et al., 2000):

Visual input routed to auditory cortex → ferrets develop visual function there.

But: the auditory cortex comes to structurally resemble visual cortex. Is this true MR?

Summary and Open Questions

The Identity Thesis:

Mental states are brain states — an empirical, a posteriori claim.

Two argumentative strategies:

Smart: Identity is posited for ontological simplicity (Occam's Razor).

Lewis: Identity is deduced from folk-psychological definitions + neuroscience.

Main challenge:

Putnam's multiple realizability argument: mental states are functional, not physical, kinds.

Open questions:

Is MR empirically vindicated, or do neural realizers converge?

Can Lewis's functional definitions accommodate species-relative identities?