



ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
САНКТ-ПЕТЕРБУРГ

PHILOSOPHY OF MIND

Sergei Levin
Сергей Михайлович Левин
slevin@hse.ru

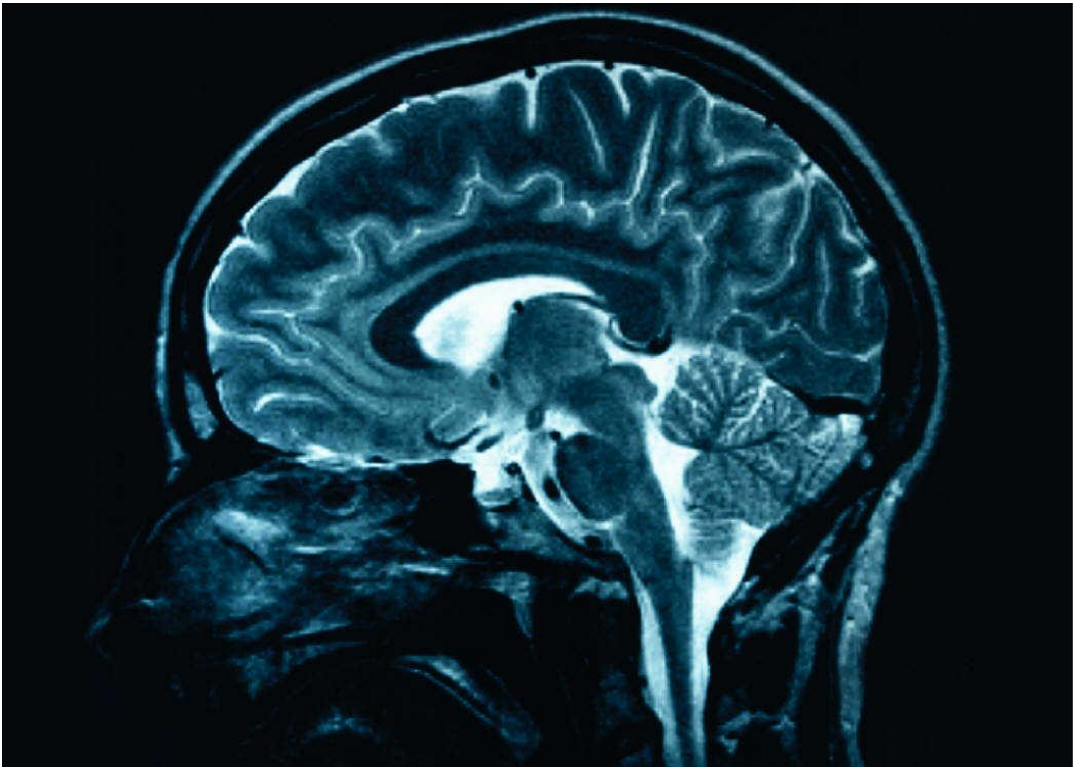
Introduction to Philosophy

* Presentation is for educational purposes only. I do not claim authorship for all texts and pictures in the presentation.

Overview

- What is the Mind?
- What does it mean to have a mind?
- How are the mind and body related?
- How can a mind emerge from purely physical processes (i.e., mind-body problem)?
- How exactly do neurobiological processes in the brain give rise to consciousness?
- Is the mind a computer program? Could a machine think?

What is the Mind? What does it mean to have a mind?



Theories of mind and consciousness



A field that exists in its own parallel "realm" of existence outside reality so can't be seen
(Substance Dualism)



Consciousness and its states (belief, desire, pain) are simply functions the brain performs
(Functionalism)



A physical property of matter, like electromagnetism, just not one we know about
(Property Dualism)



All matter has a psychic part. Consciousness just the psychic part of our brain
(Pan Psychism)



Simply, mental states are physical events that we can see in brain scans
(Identity Theory)



A sensation that "grows" inevitably out of complicated brain states
(Emergent Dualism)



Literally just behaviour. When we behave in a certain way, we appear conscious
(Behaviourism)



An accidental side-effect of complex physical processes in the brain
(Epiphenomenalism)



Not sure. But quantum physics, over classical physics, can better explain it
(Quantum Consciousness)



The sensation of your most significant thoughts being highlighted
(Cognitivism)



Consciousness is just higher order thoughts (thoughts about other thoughts)
(Higher Order Theory)



A continuous stream of ever-recurring phenomena, pinched like eddies into isolated minds
(Buddhism)

1. Substance dualism
2. Identity theory
3. Functionalism



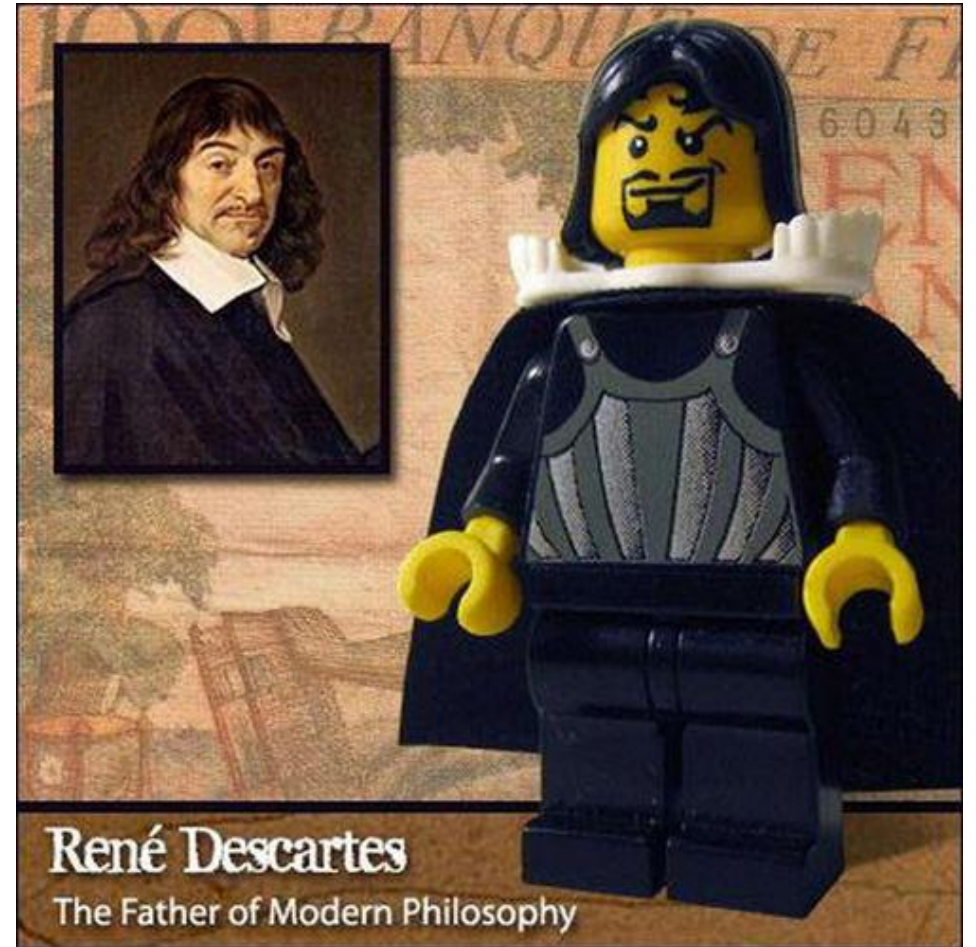
Substance dualism

The mind and body are two different substances.

There are two fundamental kinds of stuff – mental stuff (minds) and physical stuff (such as bodies).

the mind is a completely distinct substance from matter. Matter is easily described: it is measurable, has dimensions, can be touched and seen, sometimes smelt and tasted, divided, destroyed and altered. Mind, however, can almost be defined as the opposite of this – in fact, one of the difficulties with Descartes' definition is that mind seems to have almost no positive qualities. It is invisible, without dimensions, immaterial, unchanging, indivisible and without limit.

Res cogitans · Res extensa



(1596-1650)

Substance dualism

“According to substance dualism our brains and bodies are not really conscious. Your body is just an unconscious machine like your car or your television set. your body is alive in the way that plants are alive, but there is no consciousness to your body. Rather your conscious soul is somehow attached to your body and remains attached to it until your body dies, at which time your soul departs. You are identical with your soul and only incidentally and temporarily inhabit this body” (Searle, Mind p. 30).

	Mind	Body
<i>Essence</i>	Thinking (consciousness)	Extension (having spatial dimensions)
<i>Properties</i>	Known directly Free Indivisible Indestructible	Known indirectly Determined Infinitely divisible Destructible

Argument from doubt Dualism

Premise 1. I can doubt the existence of my body.

Premise 2. I cannot doubt the existence of my thoughts (my mind).

Conclusion. Therefore, my mind must be made of something fundamentally different from everything else around me.

(1641) the Second Meditation

Leibniz's law: If two things are the same thing, they must share all the same properties.

Descartes shows that mind and body seem to have different properties, and how, hence, they must be different things.

Objections

- Just because Descartes can *think* of his mind existing without his body, this doesn't mean that his mind *really can* exist without his body. Perhaps there is some metaphysical connection between his mind and body that would make this impossible that Descartes doesn't know about.
- Cp. I think the Masked Man robbed the bank; I don't think my father robbed the bank; Therefore, my father isn't the Masked Man.



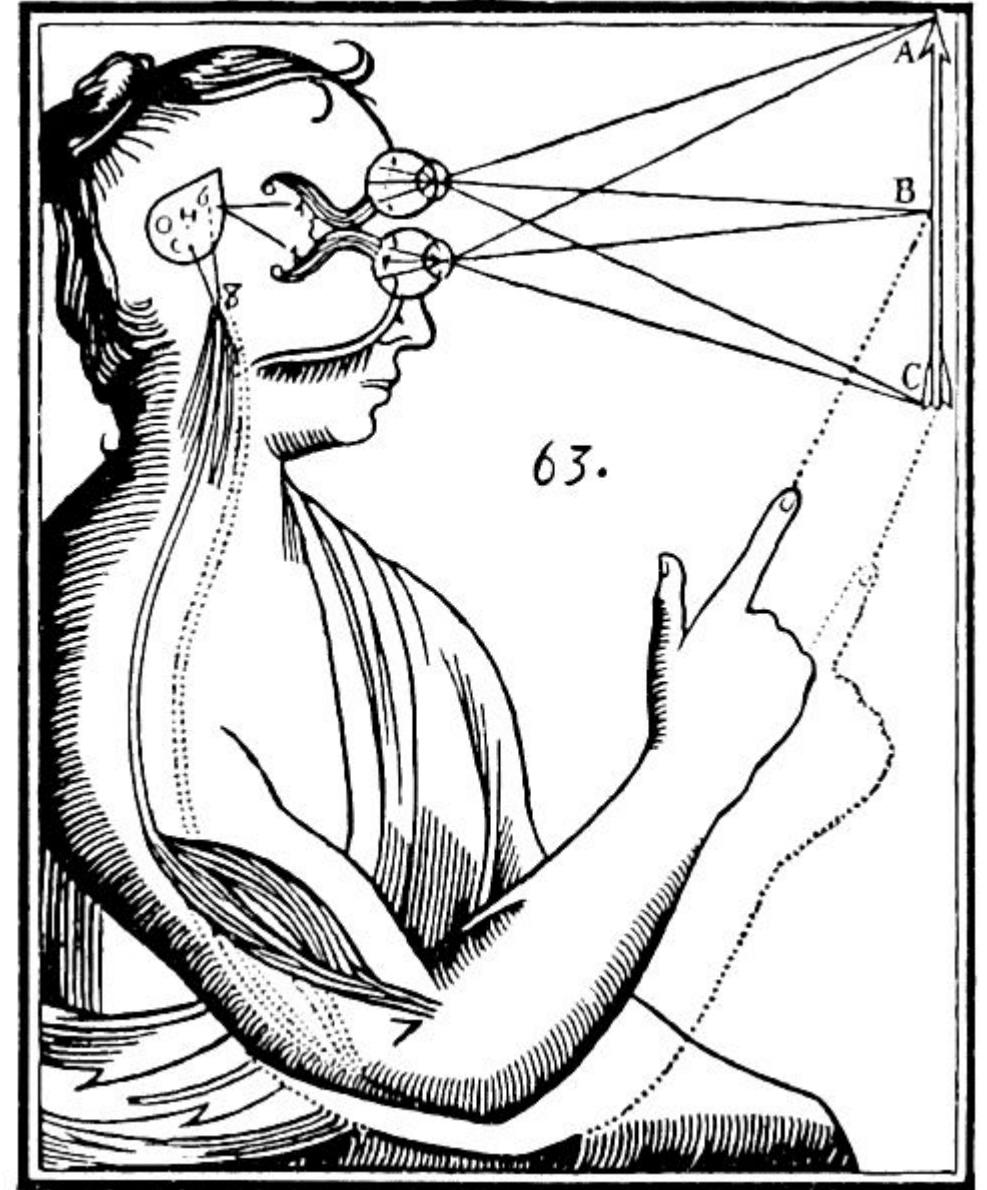
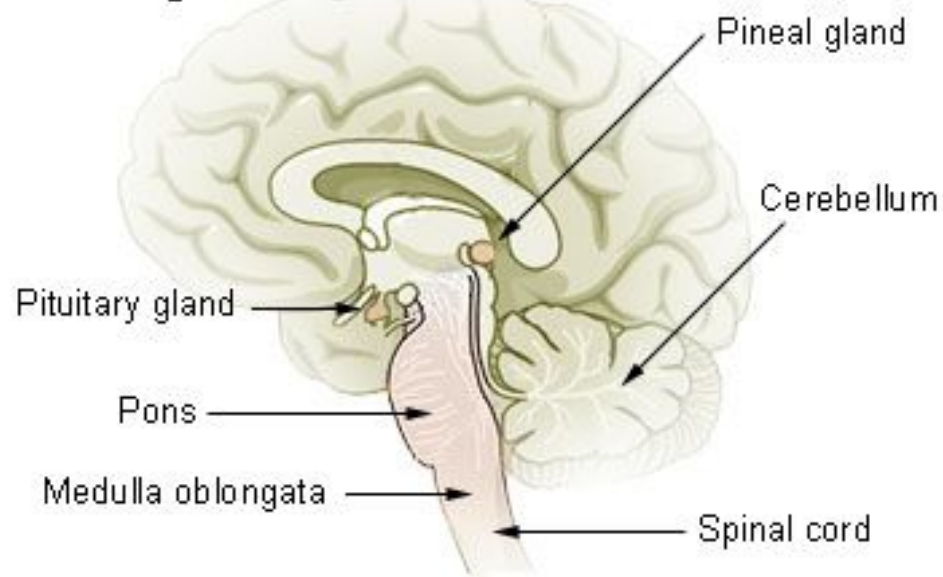
FATHER



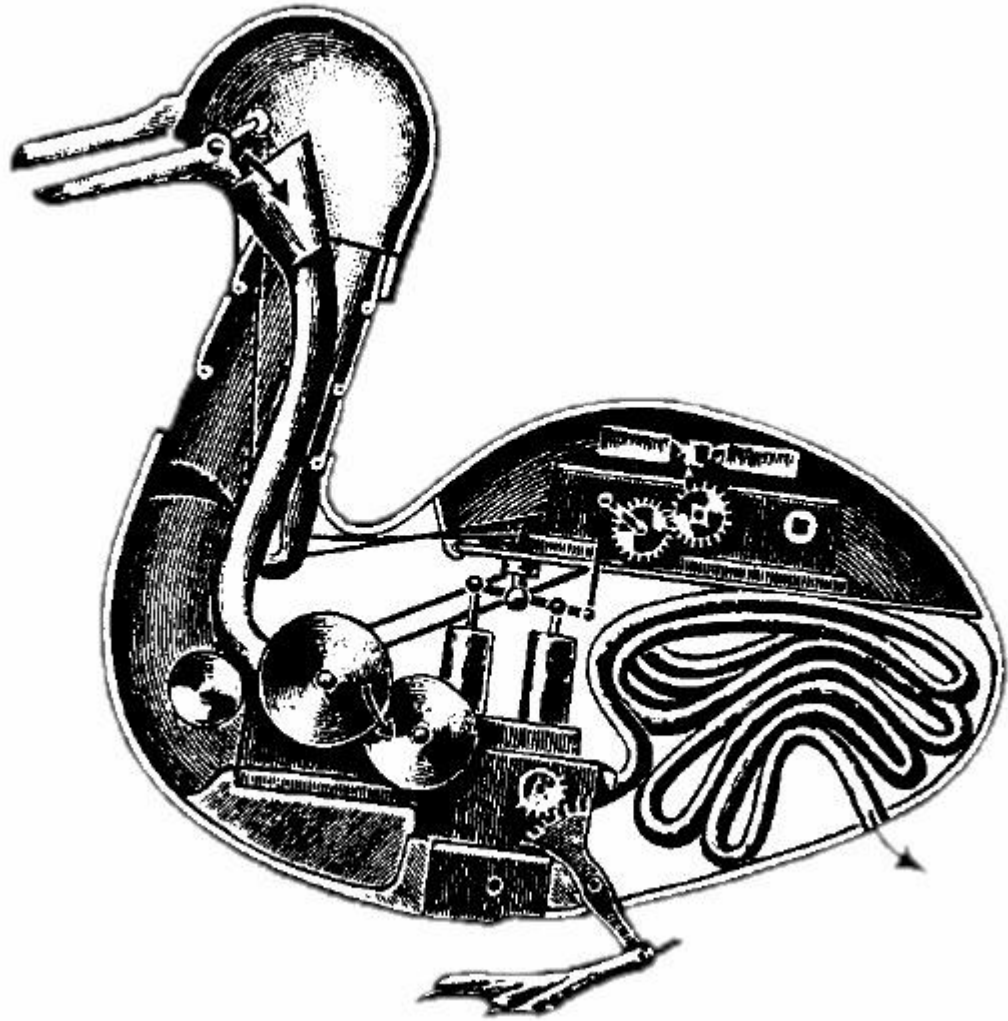
Mind body interaction

[René Descartes](#)' illustration of [mind/body dualism](#). Descartes believed inputs are passed on by the sensory organs to the [epiphysis](#) in the brain and from there to the immaterial spirit

Pituitary and Pineal Glands



Animals are mindless mechanical automatons



Privacy and First Person Authority

If I desire an apple, I know that I have this desire "introspectively." Others can know of my desire only by means of my verbal or non-verbal behavior or, conceivably, by inspection of my brain. (The latter assumes a correlation, if not an identity, between nervous and mental states or events). My linguistic, bodily and neural activities are public in the sense that anyone suitably placed can observe them. Since mental states are private to their possessors, but brain states are not, mental states cannot be identical to brain states. (Rey pp. 55-56).

A closely related argument emphasizes that my own mental states are knowable without inference; I know them "immediately." (Harman, 1973, pp. 35-37). Others can know my mental states only by making inferences based on my verbal, non-verbal or neurophysiological activity. You may infer that I believe it will rain from the fact that I am carrying an umbrella, but I do not infer that I believe it will rain from noticing that I am carrying an umbrella. I do not need to infer my mental states because I know them immediately. Since mental states are knowable without inference in the first person case, but are knowable (or at least plausibly assigned) only by inference in the third person case, we have an authority or incorrigibility with reference to our own mental states that no one else could have. Since beliefs about the physical world are always subject to revision (our inferences or theories could be mistaken), mental states are not physical states.

Problem of Interaction

May 16, 1643, Elisabeth writes,
"tell me please how the soul of a human being (it being only a thinking substance) can determine the bodily spirits and so bring about voluntary actions".

June 20, 1643

"how the soul (nonextended and immaterial) is able to move the body"

Problems with Descartes radical split between the mind and the body:

If mind and body are radically different types of stuff, it is hard to see how they can interact with each other. In particular, it is hard to see how an unextended substance can interact with an extended one.

Yet mind and body do seem to interact in both directions:

1. The mind affects the body: This seems to happen whenever we act. The mind decides to do something and the body does it.
2. The body affects the mind:
 - a. In sense perception, our sense organs seem to affect and produce images in our mind.
 - b. Damage to our brain or the influence of drugs on our body often affects our mind.



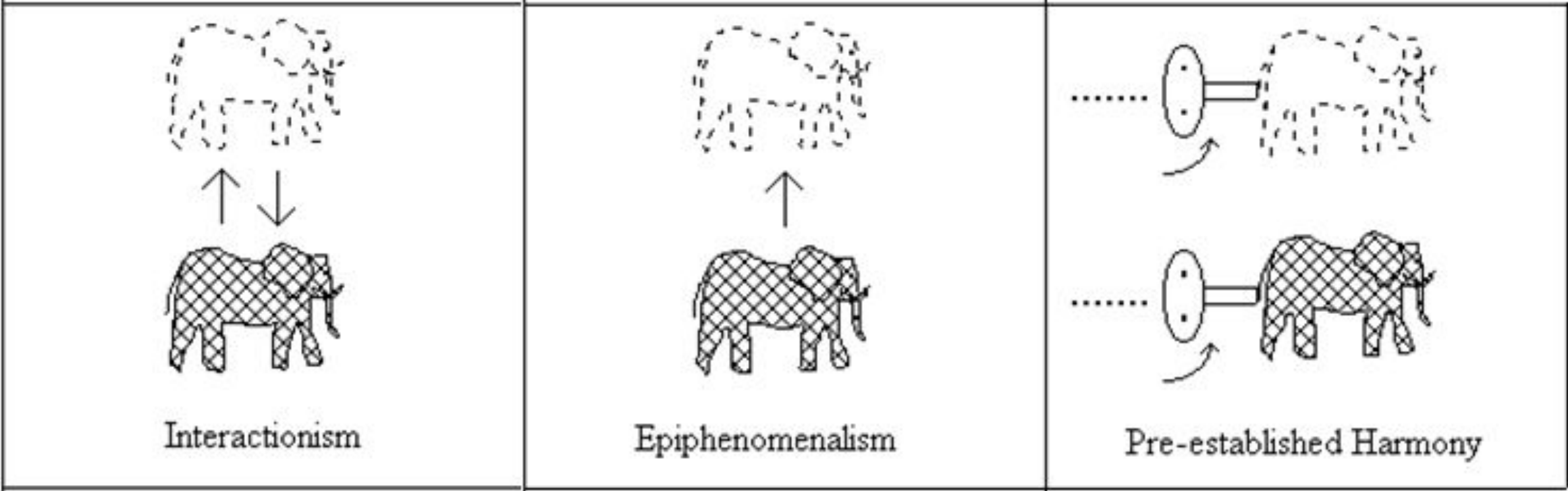
Princess Elisabeth of Bohemia and Palatine

(1618-1680)

Mental causes

- If the mind is just thought, not in space, and matter is just extension, in space, how could one possibly causally affect the other?
- All physical effects have a sufficient physical cause. Nothing physical happens needs a non-physical explanation.
- Mental causes would violate the laws of physics, e.g. law of conservation of energy.

Problem of Interaction



The 21 Grams Theory



SOUL HAS WEIGHT, PHYSICIAN THINKS

Dr. Macdougall of Haverhill Tells
of Experiments at
Death.

LOSS TO BODY RECORDED

Scales Showed an Ounce Gone in One
Case, He Says—Four Other
Doctors Present.

Special to The New York Times.

BOSTON, March 10.—That the human soul has a definite weight, which can be determined when it passes from the body, is the belief of Dr. Duncan Macdougall, a reputable physician of Haverhill. He is at the head of a Research Society which for six years has been experimenting in this field. With him, he says, have been associated four other physicians.

in 1907 Dr. Duncan MacDougall from Massachusetts tried to prove man has a soul by weighing dying people at their death.

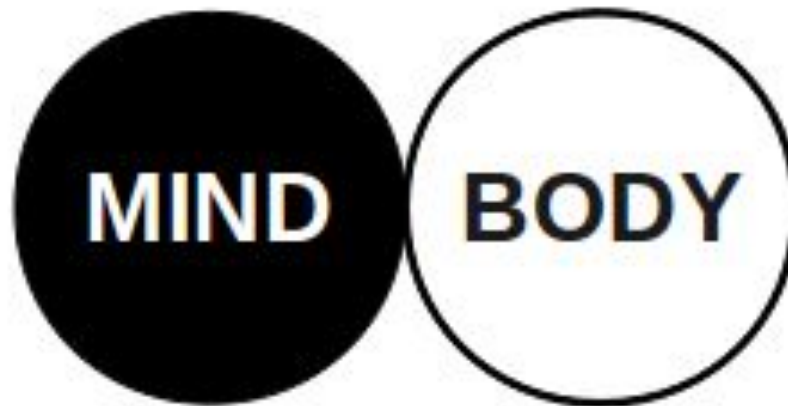
When done he noticed a slight weight change occurred. The weight change was a sudden **3/4 of an ounce** less at the point of death. When 15 dogs were likewise tested there was **no weight change for them**. Duncan MacDougall wanted to prove man was different from an animal by having a soul in an effort to combat *evolution*.

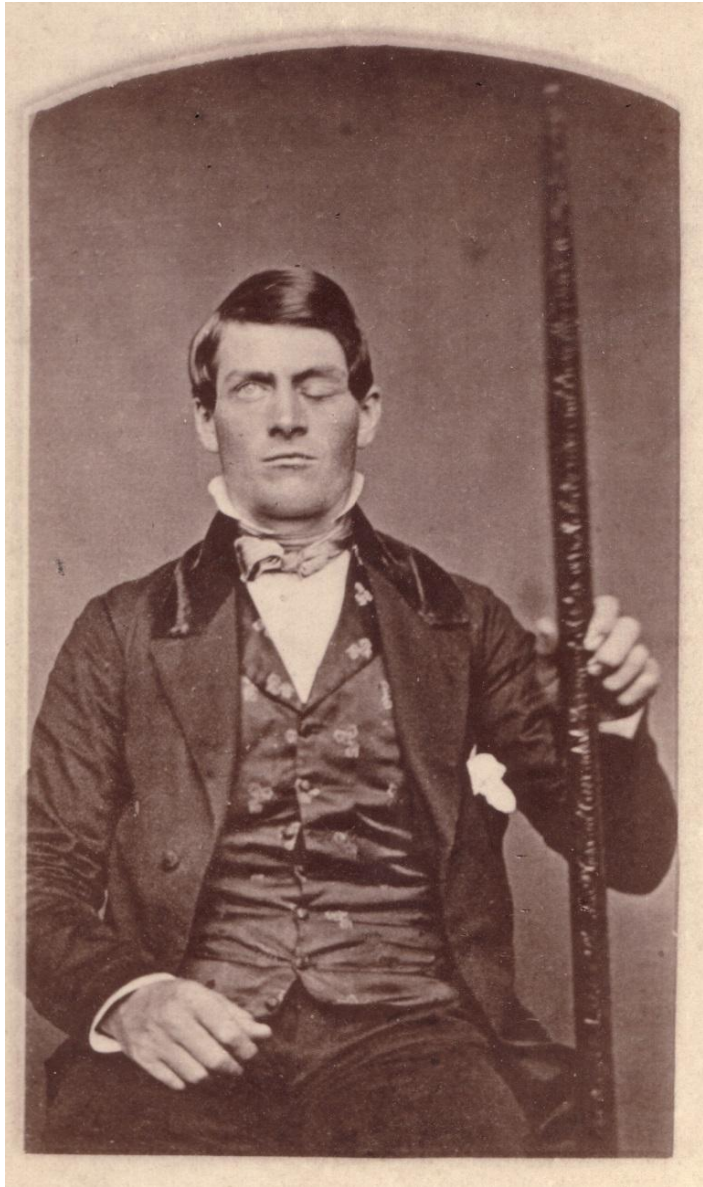
The 21 Grams Theory



Problems for Dualism

- Problem of Interaction
- The Queerness of the Mental
- First Person Authority

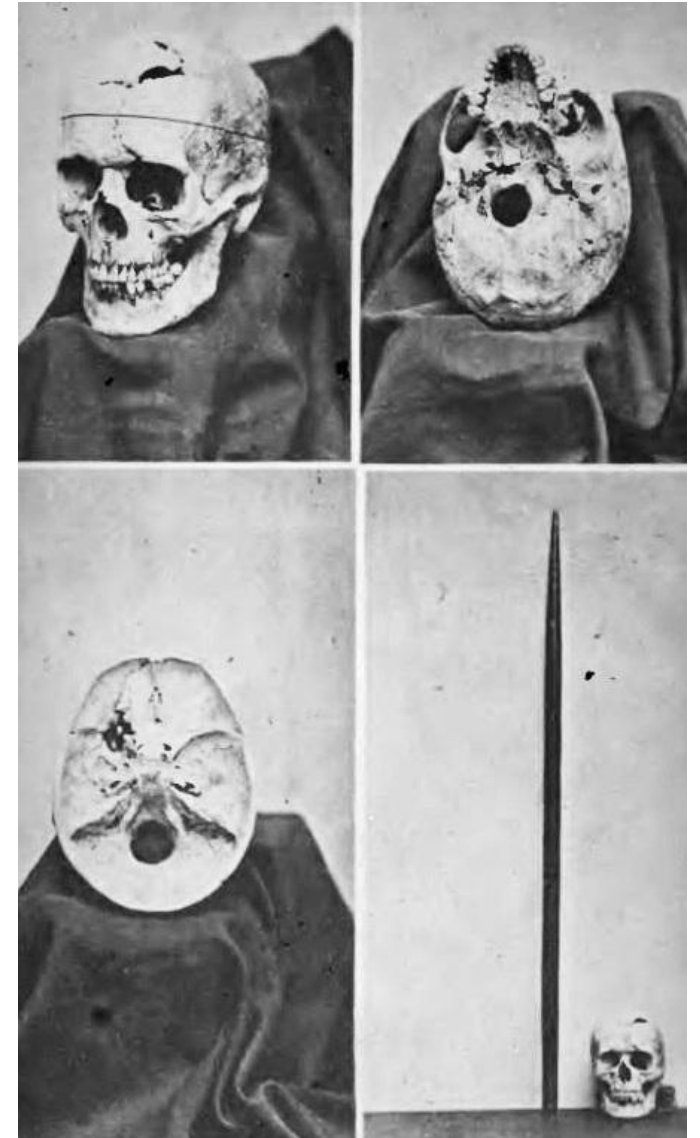




Phineas P. Gage

Phineas P. Gage (1823 – May 21, 1860) was an American railroad construction foreman remembered for his improbable [\[B1\]:19](#) survival of an accident in which a large iron rod was driven completely through his head, destroying much of his brain's left [frontal lobe](#), and for that injury's reported effects on his personality and behavior over the remaining twelve years of his life—effects so profound (for a time at least) that friends saw him as "no longer Gage".

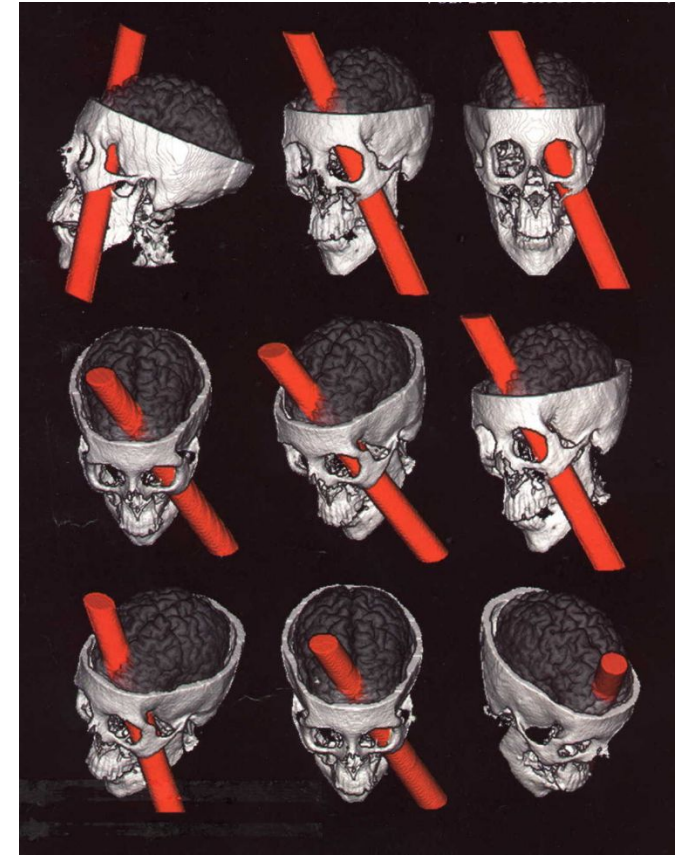
On September 13, 1848, the then 25-year-old Gage was working as the foreman of a crew preparing a railroad bed near Cavendish, Vermont. He was using an iron tamping rod to pack explosive powder into a hole. Unfortunately, the powder detonated, sending the 43 inch long and 1.25 inch diameter rod hurtling upward. The rod penetrated Gage's left cheek, tore through his brain, and exited his skull before reportedly landing some 80 feet away



The effect of an iron bar through the head on personality

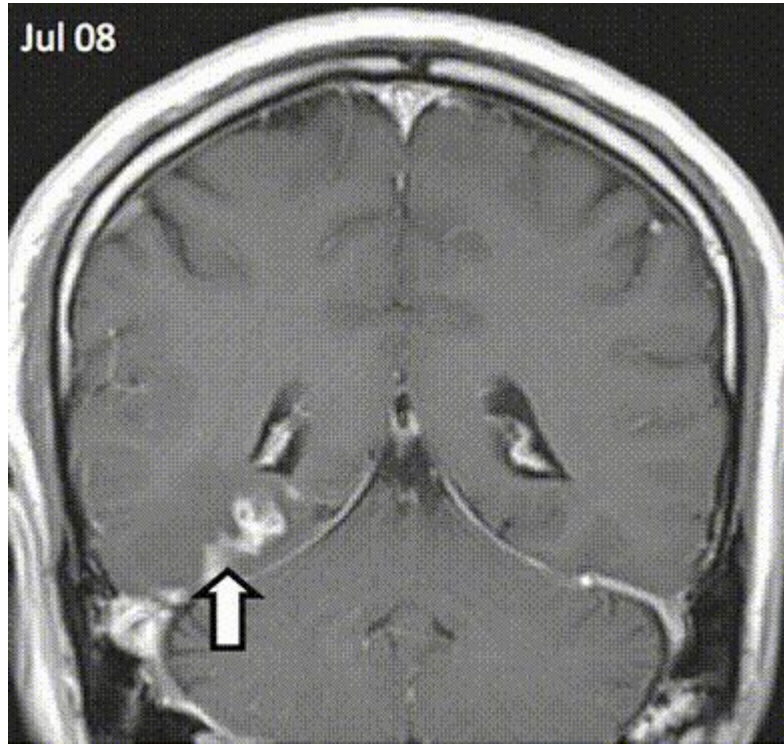
Most popular accounts of Phineas Gage describe him as having undergone profound personality changes because of his injury. He is often reported as having permanently lost his inhibitions, so that he started to behave inappropriately in social situations. Some reports state that he became violent and "uncontrollable", and even that he started to molest children.

** We actually know next to nothing about Gage's personality before the injury, so it is difficult to understand exactly how it changed afterward, and the story is further complicated by our incomplete knowledge of the extent of his injury. Despite this, the case of Phineas Gage has been used and abused ever since it first appeared.*



From Hana Damasio et al., "The Return of Phineas Gage: Clues about the brain from the skull of a famous patient" *Science*. 1994, 264, 1102-1105.

Tapeworm vs Brain



A 50-year-old Chinese man was admitted to a UK hospital complaining of headaches, seizures, an altered sense of smell and memory flashbacks.

Over the next four years, further MRIs recorded the abnormal region moving across the man's brain (see animation), until finally his doctors decided to operate. To their immense surprise, they pulled out a 1 centimetre-long ribbon-shaped worm.

Journal reference: Genome Biology, DOI: [10.1186/s13059-014-0510-3](https://doi.org/10.1186/s13059-014-0510-3)

http://www.newscientist.com/article/dn26597-watch-a-tapeworm-squirm-through-a-living-mans-brain.html?utm_source=NSNS&utm_medium=SOC&utm_campaign=hoot&cmpid=SOC%7CNSNS%7C2014-GLOBAL-hoot#.VG9qGvmsVH4

Physicalism

Physicalism is the thesis that everything is physical, or as contemporary philosophers sometimes put it, that everything supervenes on the physical. The general idea is that the nature of the actual world (i.e. the universe and everything in it) conforms to a certain condition, the condition of being physical. Of course, physicalists don't deny that the world might contain many items that at first glance don't seem physical — items of a biological, or psychological, or moral, or social nature. But they insist nevertheless that at the end of the day such items are either physical or supervene on the physical.

there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect (Davidson, 1970, 214).

Physicalism is true at a possible world w iff any world which is a physical duplicate of w is a duplicate of w *simpliciter*.

Physicalism or materialism?

Materialism

Materialism – only matter exists ->
-> philosophy of mind: mind is material.
Material?

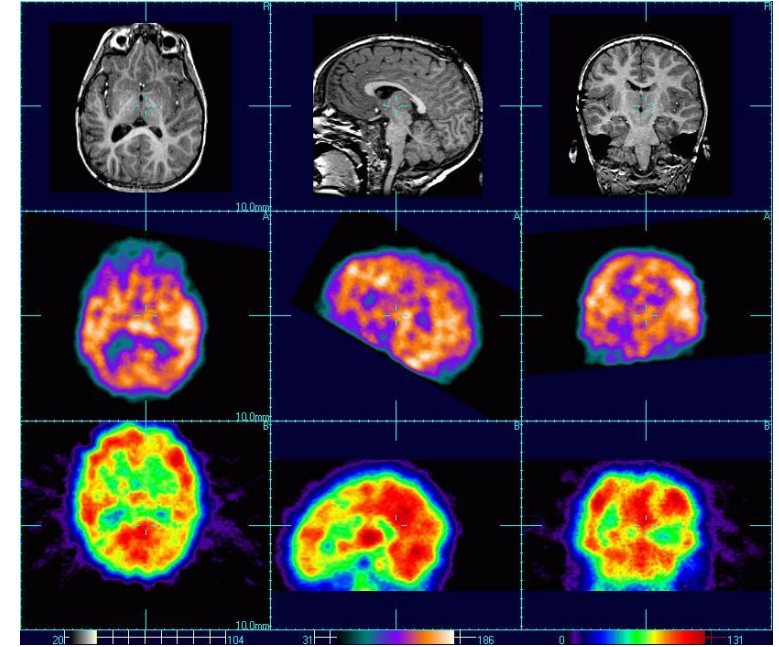
Physicalism

Physicalism – only physical exists. ->
-> philosophy of mind: mind is physical.
Physical?

Physical may be described by contemporary physics or the best physics in the future

The Mind/Brain Identity Theory

The identity theory of mind holds that states and processes of the mind are identical to states and **processes of the brain**. Strictly speaking, it need not hold that the mind is identical to the brain. Consider an experience of pain, or of seeing something, or of having a mental image. The identity theory of mind is to the effect that these experiences just *are* brain processes, not merely *correlated with* brain processes.



Identity Theory

Identity theory is a family of views on the relationship between mind and body. Type Identity theories hold that at least some types (or kinds, or classes) of mental states are, as a matter of contingent fact, literally identical with some types (or kinds, or classes) of brain states. The earliest advocates of Type Identity—U.T. Place, Herbert Feigl, and J.J.C. Smart, respectively—each proposed their own version of the theory in the late 1950s to early 60s.

David Armstrong made the radical claim that *all* mental states (including intentional ones) are identical with physical states, that philosophers of mind divided themselves into camps over the issue.

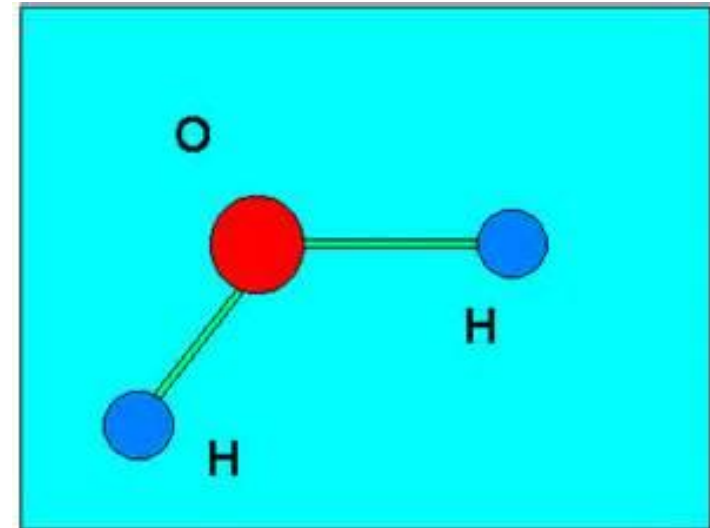
Identity Theory

“Pain” and “the firing of C-fibres” both refer to the same thing.

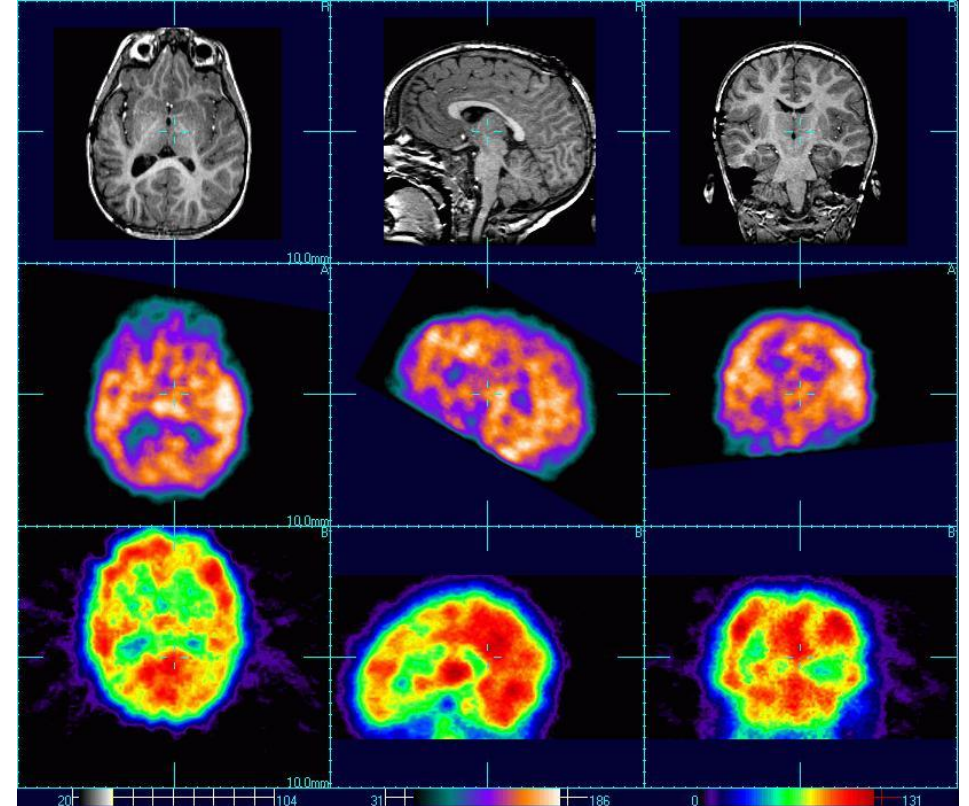
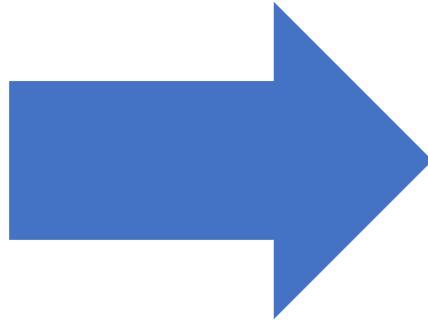
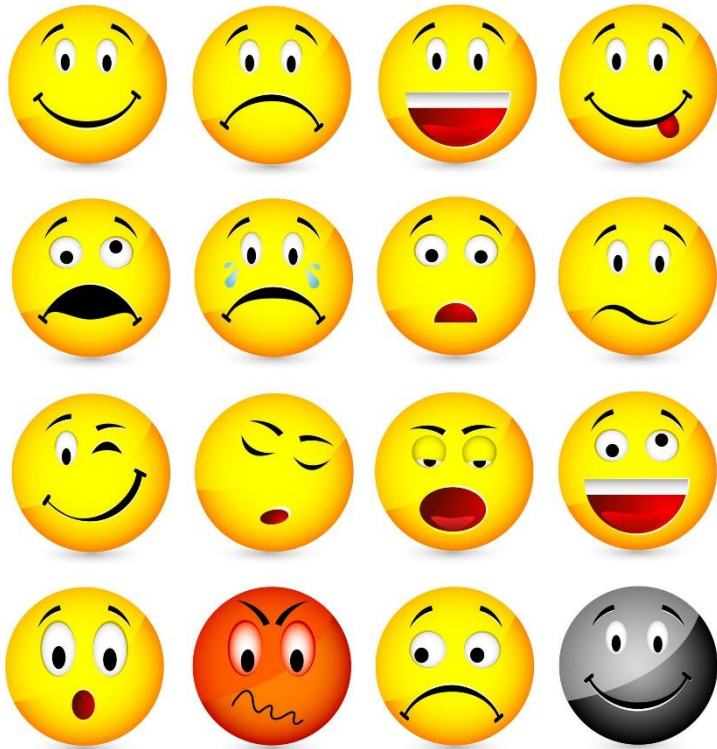
Compare

Water is H₂O

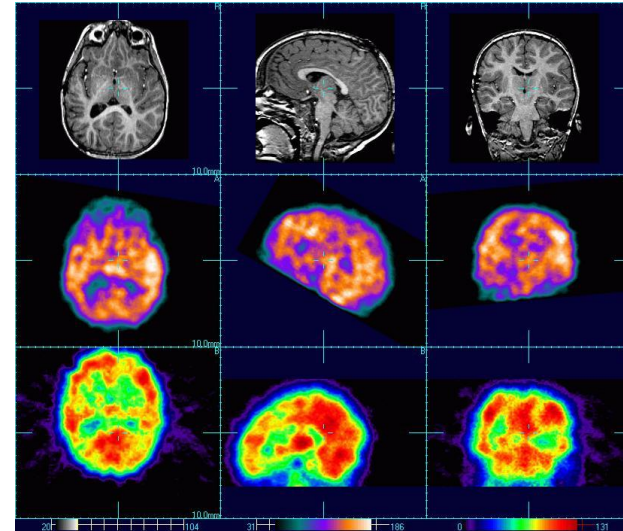
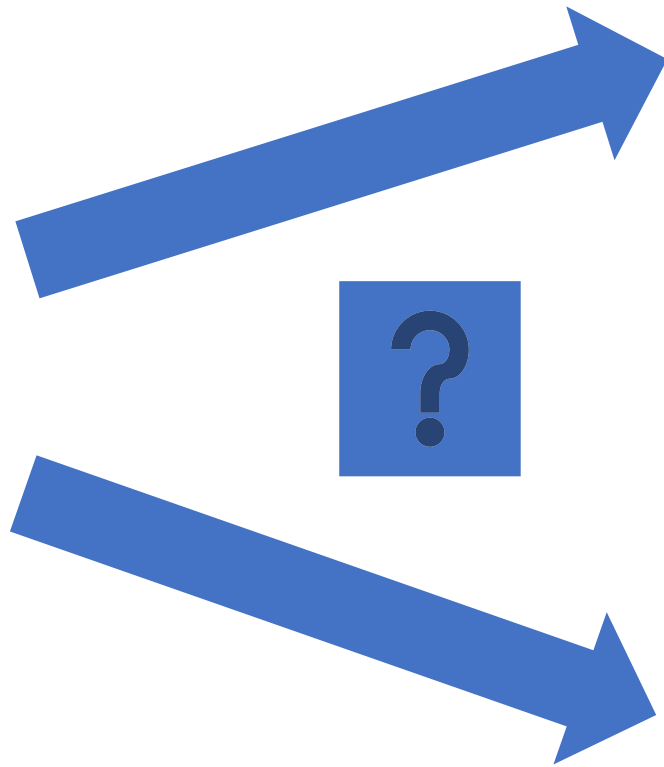
For any type of mental states M, there is some type of brain state B such that M and B are numerically identical



Reduction of folk physiology to neuroscience

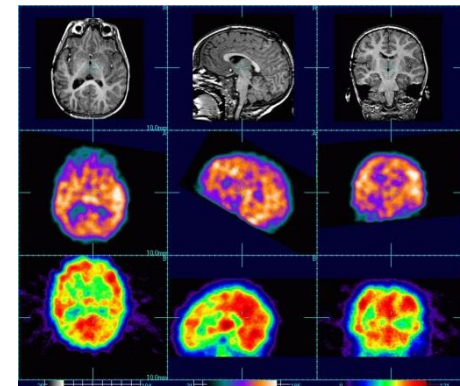
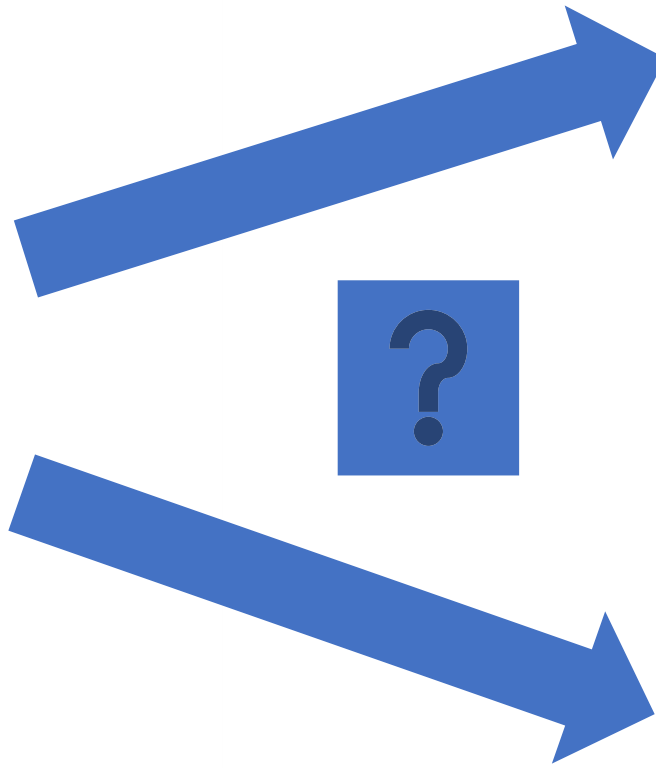


Identity Theory claims to be empirical



Benefits of the identity Theory

1) Mind/Body Correlation



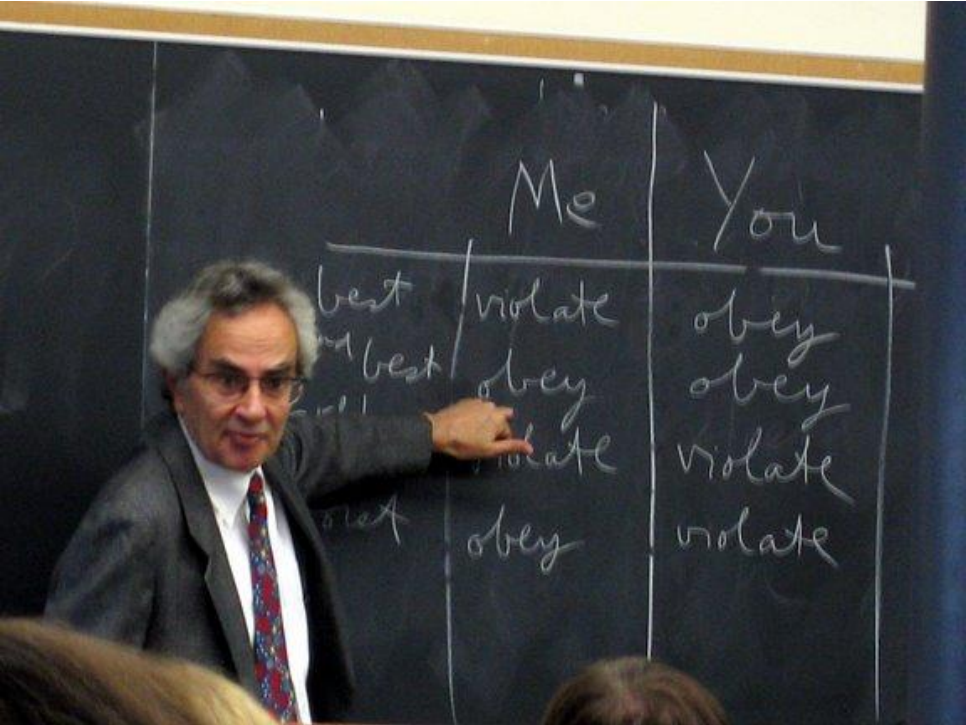
Benefits of the identity Theory

2) Historical parallels: commonsense phenomena have often been reduced by biology, chemistry, etc.

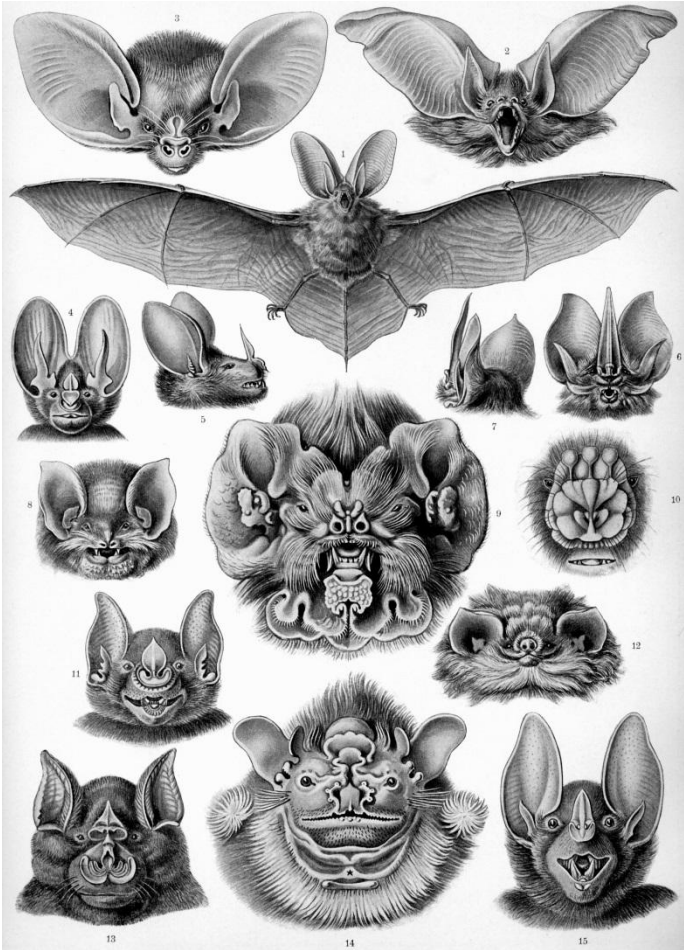
Vitalism is an obsolete scientific doctrine that "living organisms are fundamentally different from non-living entities because they contain some non-physical element or are governed by different principles than are inanimate things". Where vitalism explicitly invokes a vital principle, that element is often referred to as the "vital spark", "energy" or "élan vital", which some equate with the soul.

Objections to the identity theory

What is it like to be a bat?



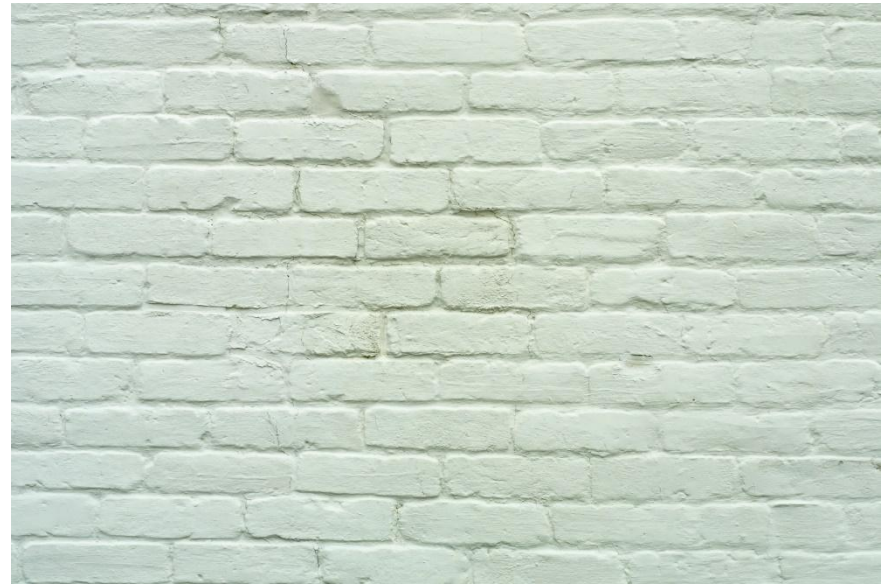
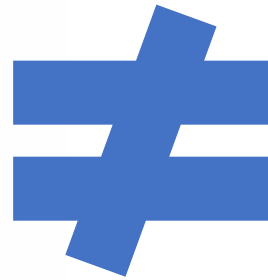
Thomas Nagel (1937-)



Objections to the identity theory

- 1) Mental states appear to have many properties that physical states lack

“Phenomenological fallacy”(U. Place)



Objections to the identity theory

2) Philosophical zombies (David Chalmers)

A philosophical zombie or p-zombie in the philosophy of mind and perception is a hypothetical being that is indistinguishable from a normal human being except in that it lacks conscious experience, qualia, or sentience. For example, a philosophical zombie could be poked with a sharp object, and not feel any pain sensation, but yet, behave exactly as if it does feel pain (it may say "ouch" and recoil from the stimulus, or say that it is in intense pain).

A **neurological zombie** that has a human brain and is generally physiologically indistinguishable from a human.



Philosophical zombies responses

- 1) Circularity. To believe in P-zombie is to believe that identity theory is fals.
- 2) *Supposing that by an act of stipulative imagination you can remove consciousness while leaving all cognitive systems intact--a quite standard but entirely bogus feat of imagination--is like supposing that by an act of stipulative imagination, you can remove health while leaving all bodily functions and powers intact. If you think you can imagine this, it's only because you are confusedly imagining some health-module that might or might not be present in a body. Health isn't that sort of thing, and neither is consciousness.*

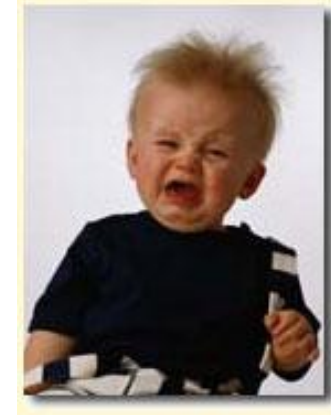
The Unimagined Preposterousness of Zombies

commentary on T. Moody, O. Flanagan and T. Polger, *Journal of Consciousness Studies*, vol. 2, no. 4, 1995, pp. 322-326. Daniel C. Dennett

Objections to the identity theory

3) Multiple realizability

Mental properties cannot be *identical* to physical properties because the same mental property can be 'realized by' different physical properties, e.g. the brain states that relate to pain are different in different species, but pain is the same mental state.



Putnam, Hilary. 1967b. "The Mental Life of Some Machines." In *Intentionality, Minds, and Perception*. Hector-Neri Castañeda, ed. (Detroit: Wayne State University Press), 177-200. Reprinted in Putnam 1975a, 408-4

Multiple realizability

The multiple-realizability thesis implies that mental types and physical types are correlated one-many not one-one. A mental state such as pain might be correlated with one type of physical state in a human and another type of physical state in, say, a Martian or pain-capable robot. This has often been taken to imply that mental types are not identical to physical types since their identity would require one type of mental state to be correlated with only one type of physical state.

Multiple realizability

1. Mental types are multiply realizable;
2. If mental types are multiply realizable, then they are not identical to physical types;
3. If mental types are not identical to physical types, then psychological discourse (vernacular or scientific) is not reducible to physical theory.

Functionalism

According to functionalism, mental states are identified by what they do rather than by what they are made of.

Consider, for example, mouse traps. Mouse traps are devices for catching or killing mice. Mouse traps can be made of most any material, and perhaps indefinitely or infinitely many designs could be employed. The most familiar sort involves a wooden platform and a metal strike bar that is driven by a coiled metal spring and can be released by a trigger. But there are mouse traps designed with adhesives, boxes, poisons, and so on. All that matters to something's being a mouse trap, at the end of the day, is that it is capable of catching or killing mice.



Defining functions

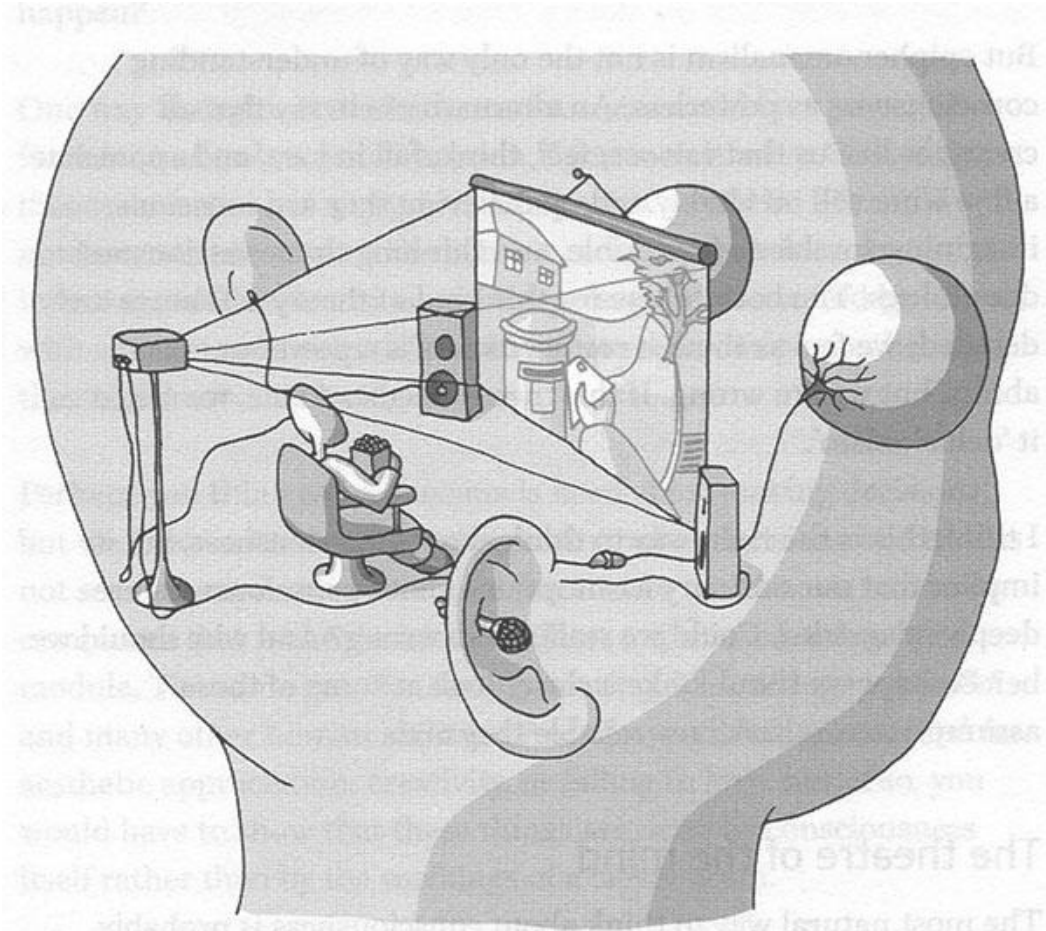
Functionalism claims that the nature of mental states is determined by what they do, by how they function. So a belief that it is sunny, for example, might be constituted in part by its relations to certain other beliefs (such as that the sun is a star), desires (such as the desire to be on a beach), inputs (such as seeing the sun), and outputs (such as putting on sunglasses.) Now consider the other beliefs and desires (in the above example) that partially constitute the nature of the belief that it is sunny. In the strongest versions of functionalism, those beliefs and desires are themselves functional states, defined by their relations to **inputs**, **outputs**, and **other mental states** that are in turn functionally constituted; and so on.

Machine state functionalism

Coca-cola machine:

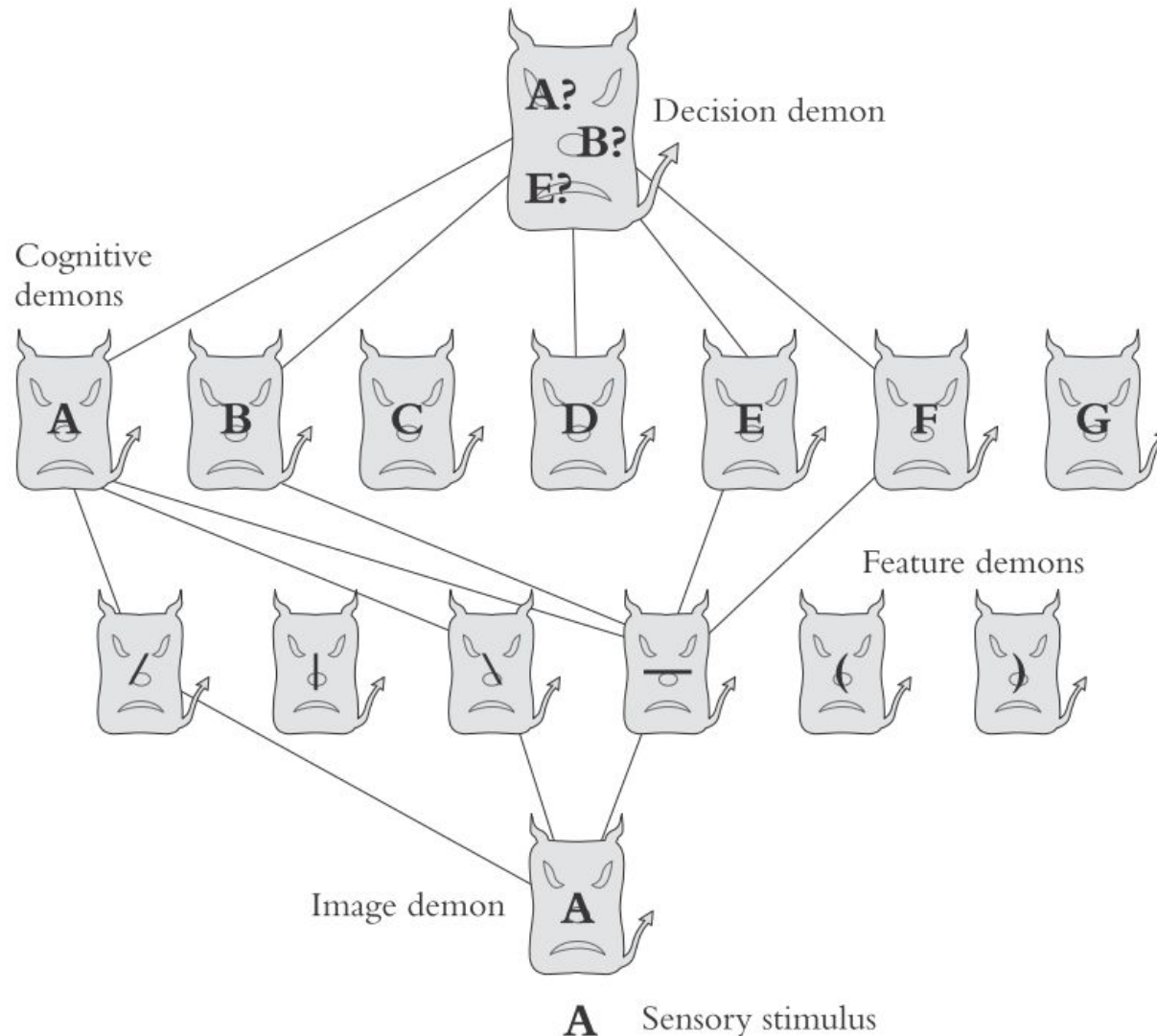
	S1	S2
50p	- change to S2	- deliver coke - change to S1
£1	- deliver coke - stay in S1	- deliver coke - deliver 50p - change to S1

Cartesian theater



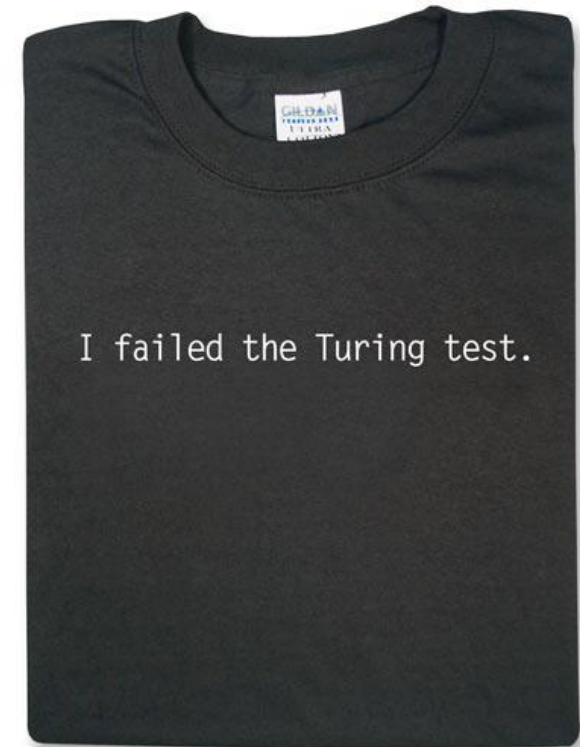
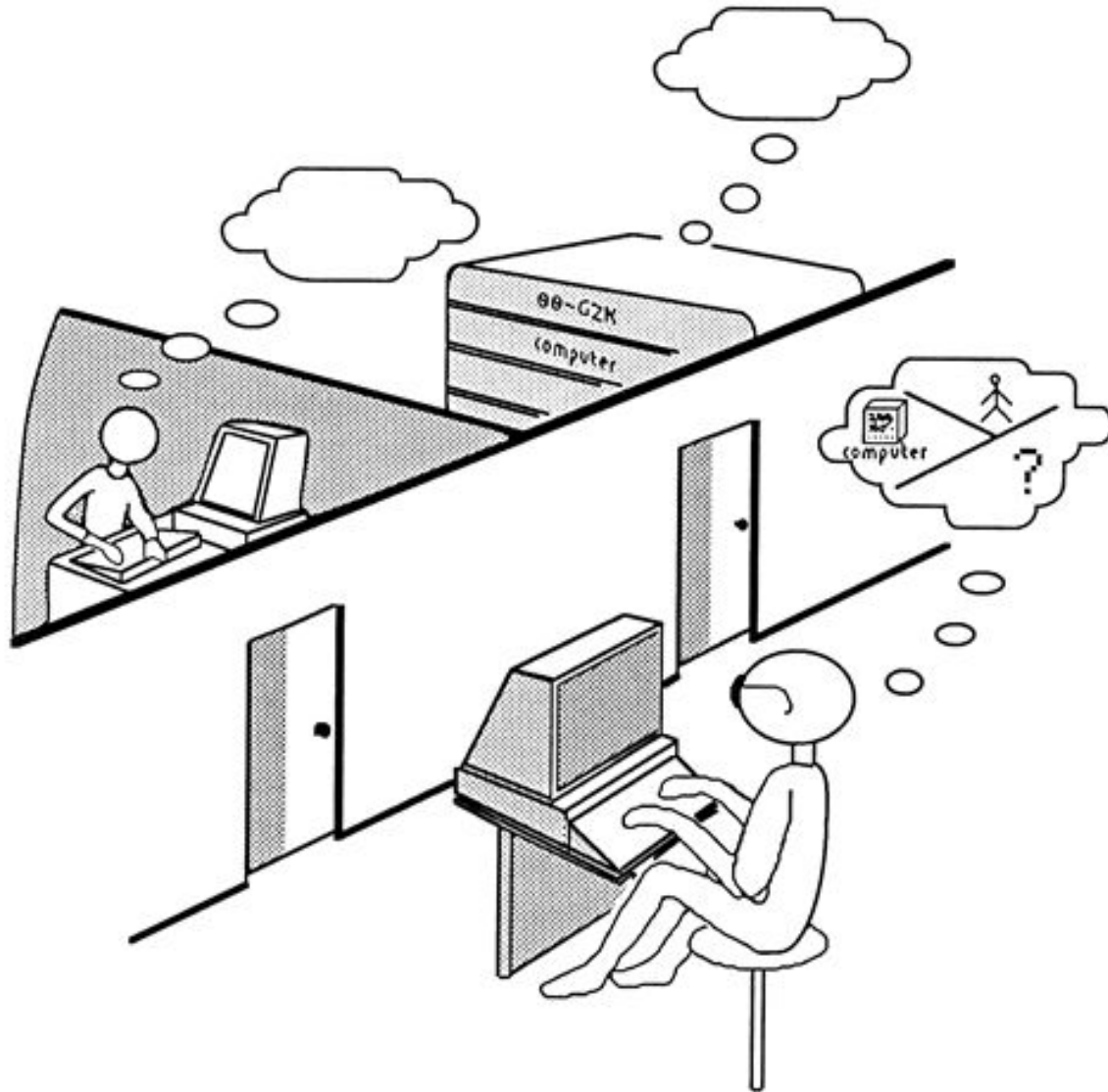
4. I feel as though I am somewhere inside my head looking out – that I experience the outside world through my eyes and ears, imagine things in my mind's eye, and direct my arms and legs to walk me down the street and post the letter. But the brain cannot work this way. This is Dennett's mythical Cartesian theatre.

Pandemonium model of Oliver Selfridge (1959)



The method of functional decomposition, a method which, put simply, explains a cognitive capacity by decomposing it into constituent parts, and specifying the causal relationships between the parts, as well as decomposing each part into further constituents, and so on (Cummins 1975).

Turing Test ("The Imitation Game")



Chines room - objection to the functionalism



If you see this shape,
"什麼"
followed by this shape,
"帶來"
followed by this shape,
"快樂"

then produce this shape,
"爲天"
followed by this shape,
"下式".

Searle, J. (1980) ` Minds, Brains and Programs

