

# Identity over time

Parfit and Swinburne

- Or it is Hume and Reid?



# 'Numerical' identity

- In general, identity is the relation that each thing bears to itself, and itself only. Thus A and B are identical just in case the set  $\{A, B\}$  has exactly one member.
- This is sometimes called *numerical* identity. I'm not sure why, but numerical identity is fundamental to counting.

# Qualitative identity

- We often talk of ‘identity’ as something different from numerical identity.
- For example, we speak of identical twins, that you and I have the same car, and so on. But of course a pair of twins are *two* people, not one, and your car is not literally the same as mine. (Or we’d be sharing!)
- So philosophers use the term *qualitative* identity to refer to this other meaning of ‘identical’. Two objects are qualitatively identical when they have the same list of properties, or qualities.

# Theories about personal identity

- Bundle theory
  - Persons are fictions. (Buddha, Hume, Parfit)
- Ego theory
  - Persons are substances. Personal identity is simple numerical identity. (Reid, Swinburne)
- “Criterion of identity” theories
  - Memory-character theory (Locke)
  - Bodily identity
  - Brain identity (Bernard Williams?)

# Bundle theory of persons

- The bundle theory says that **a person does not literally exist**, as a substance that persists through time.
  - A so-called “person” is really a fiction, that we create in our imagination out of a series of mental events.

“...there are long series of different mental states and events—thoughts, sensations, and the like—**each series being what we call one life**. Each series is unified by various kinds of causal relation, such as the relations that hold between experiences and later memories of them. Each series is thus like a bundle tied up with string.”

(Parfit)

# Successive “objects”

- Many situations involve a succession of objects that are related by “resemblance, contiguity, or causation”.
  - *Resemblance*: they have similar properties
  - *Contiguity*: they “touch” each other in space and time
  - *Causation*: the later objects arise out of the earlier objects
- In such situations the *imagination* often creates a single object out of the succession, as if the succession really consists of just one object that undergoes changes.
- Hume said that it is convenient to ‘feign identity’ in such cases.

E.g.

“A man, who hears a noise, that is frequently interrupted and renewed, says, it is still the same noise; though ‘tis evident the sounds have only a specific identity or resemblance, and there is nothing numerically the same ...”

(p. 366) N.B. ‘specific identity’ = qualitative identity



# E.g. The Ship of Theseus

- Hume points out, for example, that we talk as if a ship persists over many decades, even if all the matter that the ship is composed of is gradually replaced over that time. The “identity” of the ship over time is a mistake, a fiction.
- The same is true of people, he says. There really is no “mysterious and inexplicable” thing that is “invariable and uninterrupted”.

# E.g. the Bertrand Russell Society

- Suppose that a certain club exists for some time, holding regular meetings. The meetings then cease. Some years later, several people form a club with the same name, and the same rules. We can ask, “Did these people revive the very same club? Or did they merely start up another club which is exactly similar?” Given certain further details, **this would be another empty question.**
- (Parfit, p. 7)

# E.g. Holes

- A hole is dug in the ground, remains for a few days, and is then filled in. A year later another hole is dug in the same place. Is it the same hole as the earlier one?
- Holes in solid ground are fairly stable entities. Holes (e.g. vortices) in water, on the other hand, can come and go fairly quickly. Two of them can collide, and become one. A single vortex can split into two. There doesn't seem to be any meaning to 'identity' here, but only causal continuity.

# David Hume, *Treatise*, Book I.

“For my part, when I look inward at what I call myself, I always stumble on some particular perception of heat or cold, light or shade, love or hatred, pain or pleasure, or the like. I never catch myself without a perception, and **never observe anything but the perception**. When I am without perceptions for a while, as in sound sleep, for that period I am not aware of myself and can truly be said not to exist. ...  
... I am willing to affirm ... **that each of us is nothing but a bundle or collection of different perceptions** that follow each other enormously quickly and are in a perpetual flux and movement. ...”

(Part iv Section 6)

# Buddha

“A sentient being does exist, you think, O Mara? You are misled by a false conception. This bundle of elements is void of Self. In it there is no sentient being. Just as a set of wooden parts receives the name of carriage, so do we give to elements the name of fancied being.”

# *Ego* theory of persons

- A person is a *substance*, a thing that persists through time and can have different properties at different times. **Personal identity is numerical identity**, a primitive notion that cannot be defined.

“... what explains the unity of a person’s whole life is the fact that all of the experiences in this life are had by the same person, or subject of experiences. In its best-known form, the *Cartesian view*, each person is a persisting purely mental thing—a soul, or spiritual substance.”

## E.g. Thomas Reid holds the ego view

“Everyone has a conviction of his own identity as far back as his memory reaches; this conviction doesn’t need help from philosophy to strengthen it, and no philosophy can weaken it without first producing some degree of insanity.”

Thomas Reid, *Essays on the Intellectual Powers of Man* (1785), Essay 3 (Memory).

The identity of a person is a perfect identity; wherever it is real, it admits of no degrees; and it is impossible that a person should be in part the same, and in part different. . . For this cause, I have first considered personal identity, as that which is perfect in its kind, and the natural measure of that which is imperfect.

(Thomas Reid, *Essays on the Intellectual Powers of Man*, essay III, ch. 4 )



“If you ask for a definition of identity, I confess that I can’t give one; it is too simple a notion to admit of logical definition.” (Reid)

“My personal identity, therefore, implies the continued existence of that indivisible thing that I call myself. Whatever this self may be, it is something that thinks and wonders what to do and decides and acts and is acted on. **I am not thought; I am not action; I am not feeling; I am something that thinks and acts and feels.** My thoughts and actions and feelings change every moment; rather than lasting through time they occur in a series; but **the self or I to which they belong is permanent ...**”

(Thomas Reid, *Essays on the Intellectual Powers of Man*, essay III, ch. 4 )

# Teletransporter case

- A nice way to find out whether you hold the ego or bundle theory is to consider the teletransporter.
- “Several writers claim that, if you chose to be teletransported, believing this to be the fastest way of travelling, you would be making a terrible mistake. This would not be a way of travelling, but a way of dying.”
- **Question:** Would the person who arrives at the other end be *you*, or just a *replica*?

# Teletransporter case

- According to the bundle theory, *there is no difference at all* between **you** arriving on Mars, and **a replica of you** arriving there.
- “You do not merely want there to be psychological continuity between you and some future person. You want to *be* this future person. On the Bundle Theory, **there is no such special further fact**. What you fear will not happen, in this imagined case, *never* happens.”

# Part 2

Arguments about split brains

# The ego theory is common sense

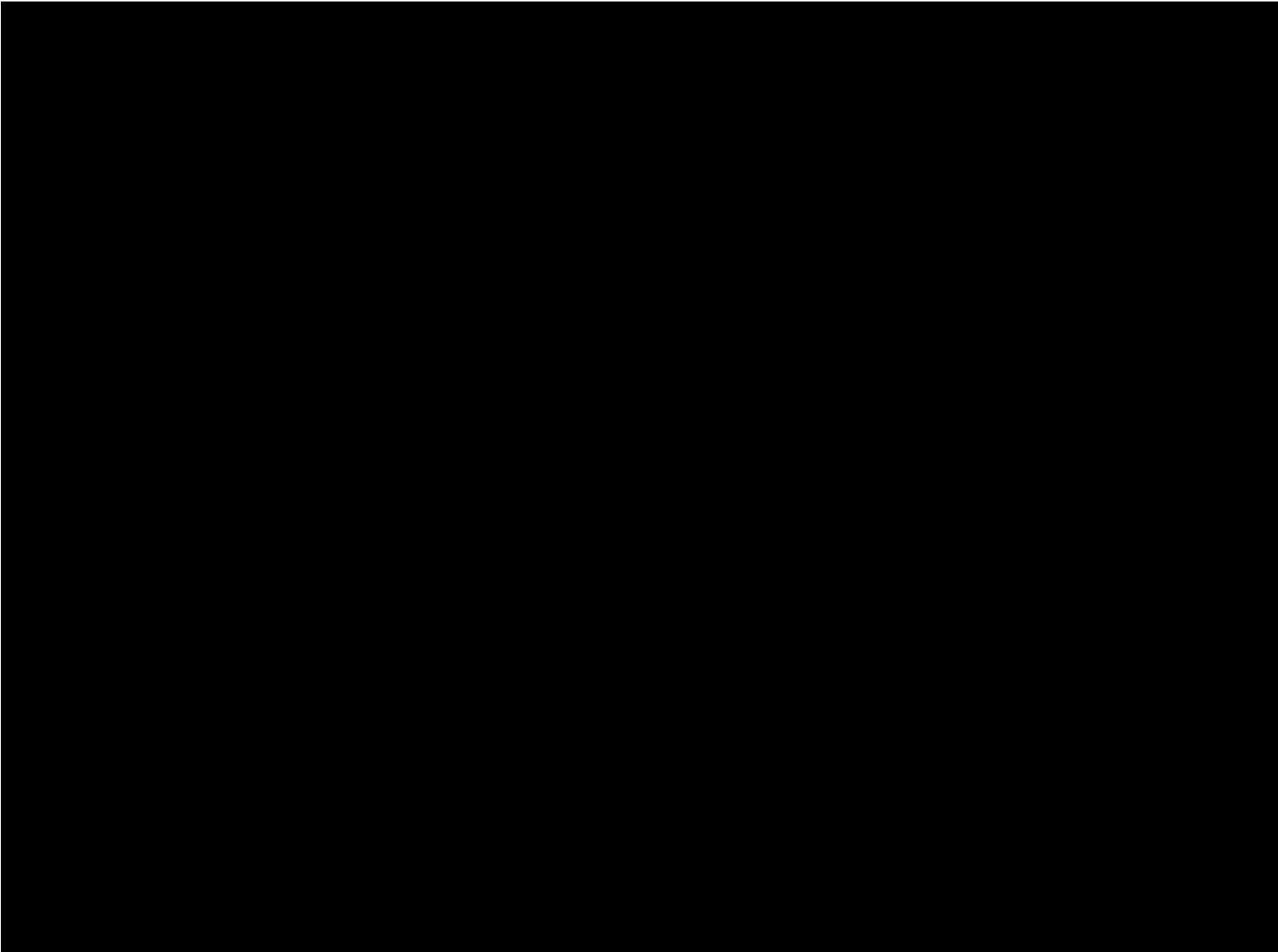
- Even Parfit says that the ego theory is common sense. The bundle theory, on the other hand, sounds crazy and distressing. (Why accept it then?)
- In particular, accepting the bundle theory might make us care less about the future.
  - Much of our concern about the future is focused on what will happen to *me* in the future. (E.g. we save for our own retirement, or worry about our own mortality.)
  - If the bundle theory is true, then there is no future person who is (literally) *me*, even an hour from now.

# Arguments against the ego theory

- Empiricism
  - There is no experience of any supposed “subject” of mental states.
- Physicalism
  - A human person is identical to their body, which is a process rather than an enduring substance.  
(Like a wave.)
- Split brain experiments

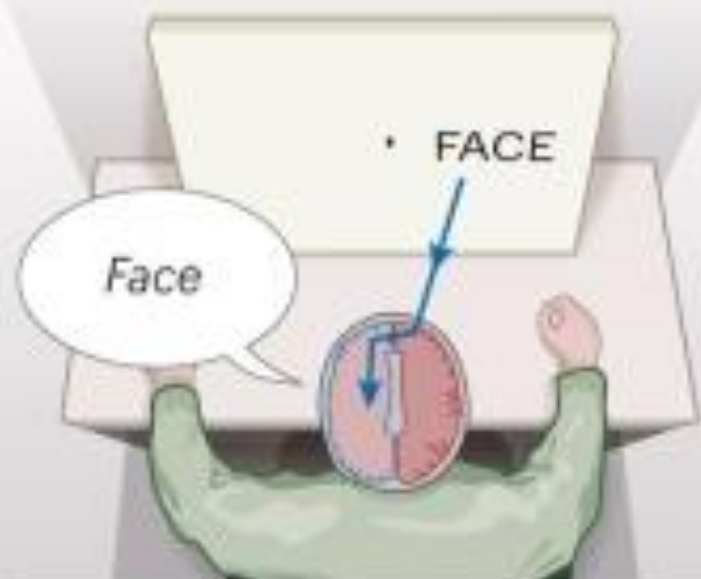
# Parfit: Split brain evidence

- To some extent, scientists can communicate separately with the two halves of a split brain.
  - How do we interpret this?
- Parfit: there are ‘two streams of consciousness’.
- Alternative: The sub-dominant hemisphere isn’t conscious
- Yair Pinto: There is only one mind in a split brain





A word is flashed briefly to the right field of view, and the patient is asked what he saw.



Because the left hemisphere is dominant for verbal processing, the patient's answer matches the word.

Now a word is flashed to the left field of view, and the patient is asked what he saw.



The right hemisphere cannot share information with the left, so the patient is unable to say what he saw, but he can draw it.

# Evidence for two conscious streams

“... if a person’s dominant hemisphere is **destroyed**, this person is able to react in the way in which, in the split-brain cases, the sub-dominant hemisphere reacts, and we do not believe that such a person is just an automaton, without consciousness. The sub-dominant hemisphere is, of course, much less developed in certain ways, typically having the linguistic abilities of a three-year-old. But three-year-olds are conscious.”

- (N.B. Parfit doesn’t infer that there are two *people* sharing the brain, as he doesn’t believe in people.)

# Against the ego theory

- Parfit argues against the ego theory, on the grounds that it cannot explain how there are *two* streams of consciousness in a split brain patient.

“What unifies the experiences, at any time, in each of this person’s two streams of consciousness? What unifies his awareness of seeing only red with his awareness of moving one hand? The answer cannot be that these experiences are being had by the same person. **This answer cannot explain the unity of each of this person’s two streams of consciousness**, since it ignores the disunity between these streams.”

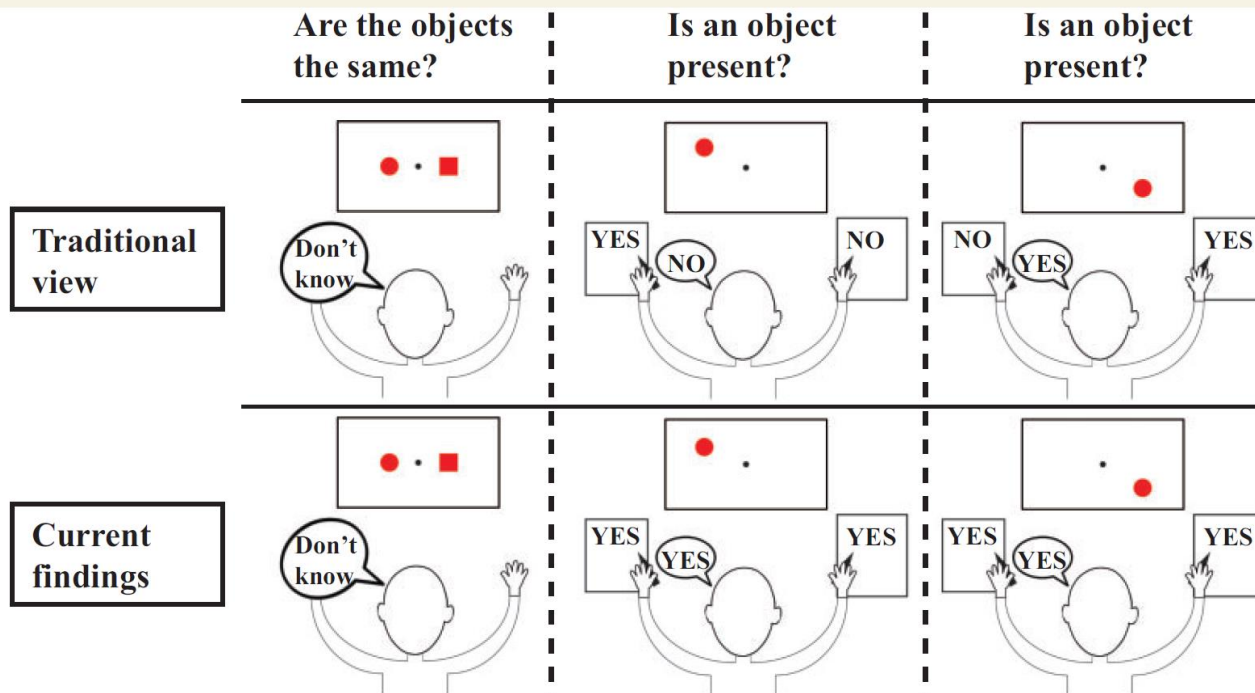
# Yair Pinto on split brain cases



Split-brain, split mind?

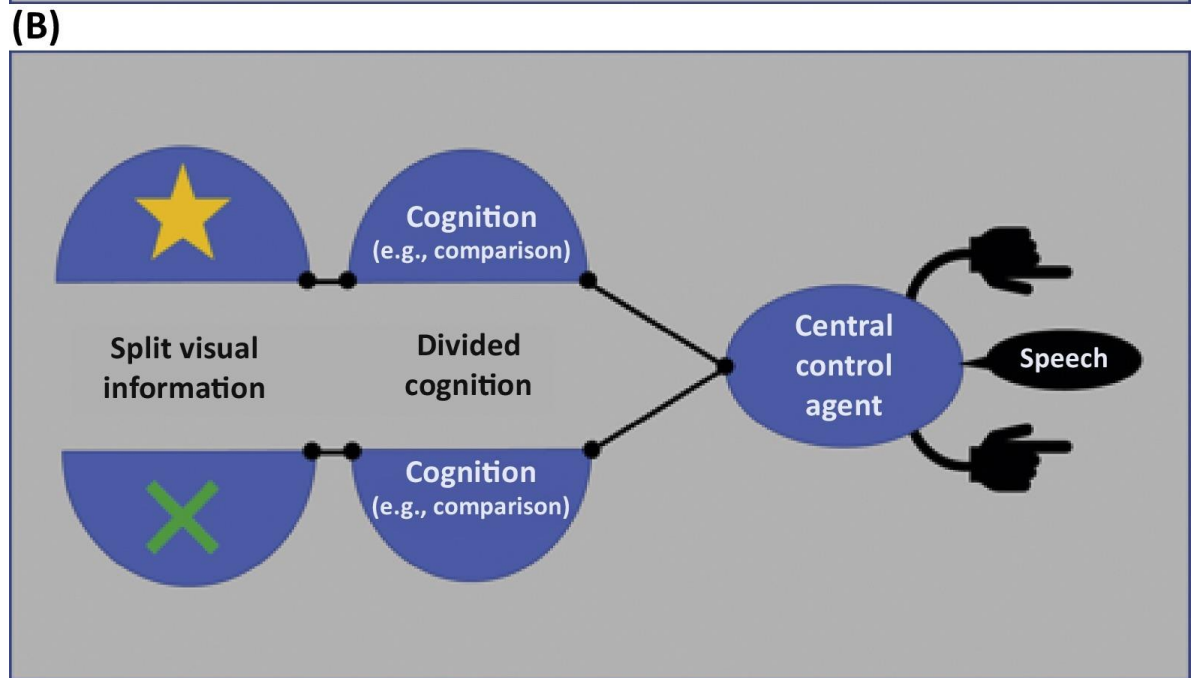
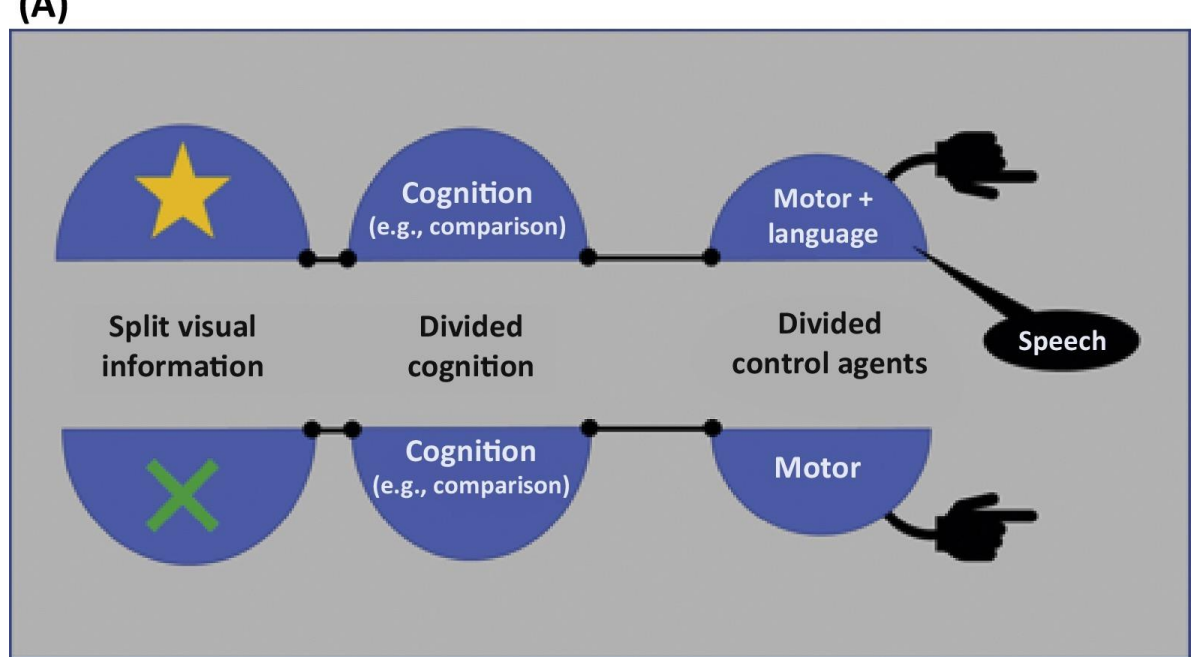
---





**Figure 1 A depiction of the traditional view of the split brain syndrome (top) versus what we actually found in two split-brain patients across a wide variety of tasks (bottom).** The canonical idea of split-brain patients is that they cannot compare stimuli across visual half-fields (*left*), because visual processing is not integrated across hemispheres. This is what we found as well. However, another key element of the traditional view is that split-brain patients can only respond accurately to stimuli in the left visual field with their left hand and to stimuli in the right visual field with their right hand and verbally. This is not what we found. Across a wide variety of tasks, we observed that split-brain patients could reliably indicate presence, location, orientation and identity of stimuli throughout the entire visual field regardless of how they responded.

(B) ... both visual information and cognition might indeed be split across the midline ... Yet each visual half field and cognition module can be monitored by a single central agent, which simultaneously controls both hands and speech (as in normal subjects)



# Egnor: “perceptual disabilities” only

“People after split brain surgery remain one person, with one consciousness, one intellect, and one will. **They have perceptual disabilities caused by the surgery** but those disabilities are subtle and not noticed in everyday life. **Their abstract intellect remains unified and the will that follows on that intellect remains unified.** Split-brain surgery doesn’t split logic or mathematics or abstract reasoning or moral decisions based on abstract reasoning.”

- Michael Egnor (neurosurgeon and Thomistic dualist)

# Unified response to bilateral hemispheric stimulation by a split-brain patient

[Justine Sergent](#)

[Nature](#) **305**, 800–802 (1983) | [Cite this article](#)

The two hemispheres of a callosotomized patient were simultaneously presented with information associated with conflicting responses, and the subject was requested to produce a single response. In all combinations of hemisphere stimulation and hand responding, the subject was capable of perfect accuracy, suggesting that **he could integrate and resolve the conflicting information** before the production of his response, and that his two disconnected hemispheres were **simultaneously aiming at the same goal**.



“... callosotomy leads to a **broad breakdown of functional integration** ranging from perception to attention. However, the breakdown is not absolute as **several processes, such as action control, seem to remain unified**. Disagreement exists about the responsible mechanisms for this remaining unity.”

“Does a split-brain harbor a split consciousness or is consciousness unified? The current consensus is that the body of evidence is insufficient to answer this question ...”

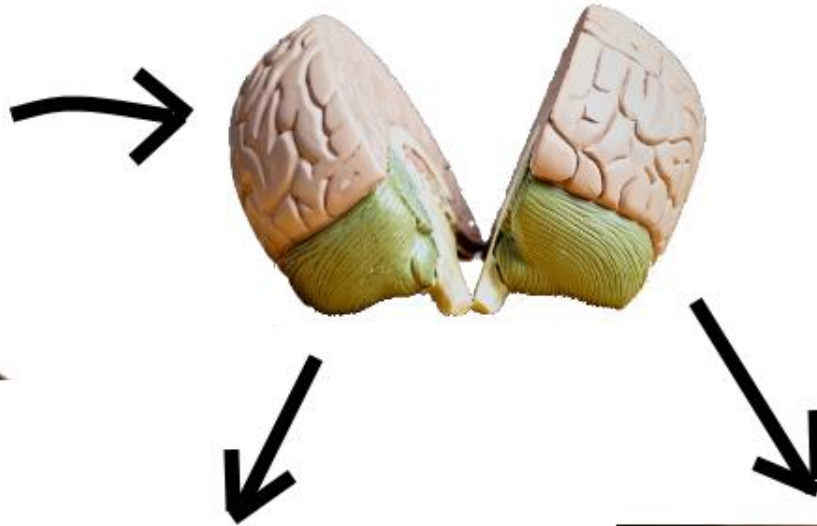
- Edward H. F. de Haan et al. (2020) “Split-Brain: What We Know Now and Why This is Important for Understanding Consciousness”, *Neuropsychology Review*, pages 224–233.

# Split brain *transplant* (thought experiment)

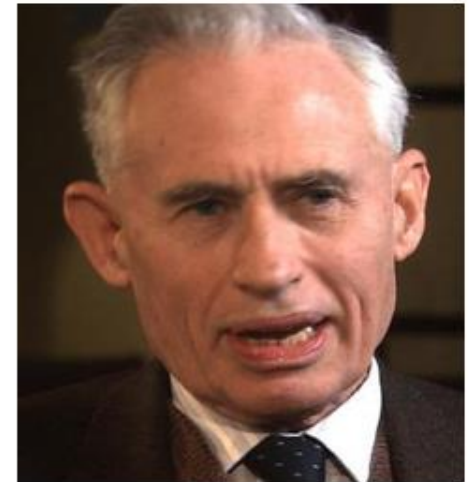
“In this imagined case a person’s brain is divided, and the two halves are transplanted into a pair of different bodies. The two resulting people live quite separate lives. This imagined case shows that **personal identity is not what matters**. If I was about to divide, I should conclude that neither of the resulting people will be me. I will have ceased to exist. But this way of ceasing to exist is about as good—or as bad—as ordinary survival.”

(Parfit)

# Split brain *transplant*



If both  
transplants  
“take”, then  
*neither* person  
can be Thomas  
Reid.



# Survival $\nRightarrow$ identity

- The mere *possibility* of multiple survival of a person seems to entail that persons don't exist. Suppose Fred is divided into 2 persons, A and B.
  1. A and B lead their own lives, so  $A \neq B$ .
  2. A and B each bear the same relation to Fred.
  3. The relation between Fred and (say) A is the same that each of us bears to our future 'selves'.

-----

$\therefore$  Fred is neither A nor B

$\therefore$  We are not identical to our future 'selves'

# Swinburne's Response

- Swinburne takes the split brain transplant case to show that continuity of brain matter, memories, etc. are mere *evidence* of identity, not *constitutive* of it.

“And of course the duplication objection that they allow for the two subsequent persons being the same person, which we brought against the brain and the memory theories, has no force against the simple theory. For **although there can be equally good evidence that each of two later persons is the same person as an earlier person, that evidence is fallible**; and since clearly only one person at one time can be strictly the same person as some person at an earlier time, it follows that **in one case the evidence is misleading – although we may not know in which case.**”

# Part 3

Richard Swinburne on “complex” theories

# Swinburne

- Swinburne accepts the ego theory. There is, he believes, an entity that is the subject of experiences, that persists through time. He concludes that:
- **A person is an indivisible, immaterial substance.**
- He starts by distinguishing between issues of
  1. *What makes  $P_1 = P_2$  and*
  2. *How we know  $P_1 = P_2$ .*
- Empiricist theories fail to distinguish between these two issues, because they follow the “verificationist dogma” that meaning is defined by conditions of verification.

# “Empiricist” theories

- Bodily continuity theory
- Brain theory
- Memory-character theory
- N.B. These are all “complex” theories, that give “criteria” for personal identity.
- Swinburne’s plan is to examine these theories and find them all unsatisfactory. Then in Section 2 he’ll argue for a dualist theory.



# Bodily continuity theory

- Same person  $\Leftrightarrow$  same body
- But to be the same body it's not enough just to have the same matter.
  - “If the matter of the oak tree is reduced to a heap of planks, the oak tree, lacking its essential properties, has ceased to exist.”
- A being consists of matter *and properties*.
- Properties are either *essential* or *accidental*.
- Sum of essential properties = “form”.

# Bodily identity defined

“What makes a substance the same substance as an earlier substance is that its **matter is the same**, or obtained from the matter of the former substance by **gradual replacement**, while **continuing to possess the essential properties which constitute its form.**”

- N.B. There's a difference though between living and non-living beings.
  - For non-living things (ships, tables, etc.) too much replacement destroys identity, even if gradual.

# Bodily identity defined

- If persons are their bodies, then their identity over time works the same way as for trees.
  - But what is the “form” (set of essential properties) for a human?

“for Aristotle, not merely shape and physiological properties, but a kind of way of behaving and a capacity for a mental life of thought and feeling.”

# Problem for the bodily theory

- Only one part of body seems to “count” for personal identity: the brain.
  - If you lose a limbs, etc. you’re still there.
  - With a whole-brain transplant, the person would go where their brain does.
  - (N.B. memories, character, attitudes, preferences, etc. would follow the brain.)

“Hence a philosopher seeking a **materialist** criterion of personal identity, will come to regard the brain, the core of the body, rather than the rest of the body as what matters for personal identity.”

# The brain theory

- Same person  $\Leftrightarrow$  same brain

# The memory-character theory

- Personal identity is constituted by continuity of memory and character.
  - (N.B. these seem to be matters of degree.)
- What do we mean by “memory”?
  1. First person (“personal”) memories, not factual memories.
  2. “Weak” (apparent) memories rather than “strong” (necessarily true) ones.

## E.g. John Locke

“So Locke’s theory can now be rephrased as follows:  $P_2$  at  $t_2$  is the same person as  $P_1$  at an earlier time  $t_1$ , if and only if  $P_2$  **apparently remembers** having done and experienced various things when those things were in fact done and experienced by  $P_1$ . **A person is who he thinks that he is.**”

- The failure of transitivity problem can be fixed by making identity a kind of ancestor (or descendent) relation.
  - I am linked to other stages of myself by chains of remembering.

# Major objection

- **“Duplication objection”**. (Similar to my obsessed historian case in Phil 1101.)
  - The man Charles who seems to remember being Guy Fawkes really is Guy Fawkes! But Robert also meets these criteria. Are they both Fawkes? They can’t be, as they are different from each other, and identity is transitive.

“So apparent memory cannot constitute personal identity, although it may be fallible evidence of it. ... the majority of contemporary writers ... favour a theory which makes some sort of bodily continuity central to personal identity.”



# Vagueness objection

- Continuity of memory and character are matters of degree, yet personal identity is all or nothing.
- We can set a threshold for how much memory and character have to be preserved, but this is arbitrary.

# Objections to the brain theory

- Swinburne appeals to Parfit's case of a person's two brain hemispheres being transplanted into different bodies.

“There are no logical difficulties in supposing that we could transplant one of  $P_1$ 's hemispheres into one skull from which a brain had been removed, and the other hemisphere into another such skull, and that both transplants should take, and it may well be practically possible to do so.”

- (Swinburne doesn't conclude that the bundle theory is true.)

# Objections to the brain theory

- The brain theory also faces the vagueness problem.
  - How much brain matter can you lose, and still be the same person?
- (Parfit also uses this argument.)

## Solution: the “simple view” (Ego theory)

“personal identity is distinct from, although evidenced by, similarity of memory and continuity of brain. This account ... can meet all the difficulties which have beset the other theories which we have discussed.”

- (Recall the distinction he began with, between what *makes*  $P_1 = P_2$ , and *how we know*  $P_1 = P_2$ )

# Solution: the “simple view”

- Complex views, which create ‘criteria’ of personal identity,

“suppose that mere logic could determine which of the experiences had by various persons, each of which was to some extent continuous with me in apparent memory and brain matter, would be mine. **There seemed to be a further truth**—that I would or would not have those experiences—beyond any truths about the extent of similarity in apparent memory and matter of future persons to myself”

# Part 4

Swinburne argues for dualism

# Section 2: The dualist theory

- Swinburne's *aim* in this section is show that continuity of brain and memory are totally dispensable to personal identity.
  - (Therefore, personal identity is identity of the *soul*.)
- **Method:** Use *a priori* judgements of logical possibility.
  - If an idea contains no contradiction (or seems to contain no contradiction) this is evidence that it's logically possible.

# *Are these cases logically possible?*

“Admittedly, there may be stories or beliefs which involve a hidden contradiction when initially they do not seem to do so. But the fact that there seems (and to so many people) to be no contradiction hidden in these stories is good reason for supposing that there is no contradiction hidden in them – until a contradiction is revealed. If this were not a good reason for believing there to be no contradiction, we would have no good reason for believing any sentence at all to be free of hidden contradiction.”



# Logical possibilities

1. “a person might acquire a totally new body (including a completely new brain)”
  2. “a person might become disembodied.”
  3. A person could “continue to exist without any apparent memory of his previous doings”
- In fact, Swinburne thinks it’s conceivable that a person could continue without their body *and* without their memories.

“Those who hope to survive their death, despite the destruction of their body, will not necessarily be disturbed if they come to believe that they will then have no memory of their past life on Earth; they may just want to survive and have no interest in continuing to recall life on Earth. Again, apparently, there seems to be no contradiction involved in their belief.”

- Do you agree?

# The “wider Aristotelian account”

- “two substances are the same if and only if they have the same form and there is continuity of the stuff of which they are made”
  - (and allow that there may be kinds of stuff other than matter)
- “there is a stuff of another kind, immaterial stuff, and that persons are made of both normal bodily matter and of this immaterial stuff but that it is the continuity of the latter which provides that continuity of stuff which is necessary for the identity of the person over time.”

# Classical dualism

- “The soul is the essential part of a person, and it is its continuing which constitutes the continuing of the person.”
- Soul stuff is *essentially* indivisible, since it is not *logically possible* that persons divide.
- “There is nothing odd about supposing that soul-stuff comes in essentially indivisible units”, because soul-stuff is not extended in space. Matter, being so extended, is **already** divided, in a sense.

# Cartesian-style argument

1. It is logically possible for me to continue existing without my body.
2. It is logically impossible, if I am simply a body, to continue existing without my body.

-----

∴ I am not simply a body. (I must have another part that can continue after my body is destroyed.)

# Swinburne's statement

“For there is not even a logical possibility, that if I now consist of nothing but matter and the matter is destroyed, I should nevertheless continue to exist. From the mere logical possibility of my continued existence there follows the actual fact that there is now more to me than my body; and that more is the essential part of myself. A person's being conscious is thus to be analysed as an immaterial core of himself, his soul being conscious.”

- Looks valid. Is it?
- (I don't think so)
  - Even if it's invalid, I think it's still true that the ego theory is inconsistent with physicalism, and may well require substance dualism.

- E.g. suppose someone tells you about their oak table, that is 30 inches high, 7 feet long, and weighs 150 lb. Is it logically possible that it has 8 sides?

$$K = A \wedge B \wedge C$$

- Yes, of course. (It's *epistemically* possible, i.e. logically consistent with the information we have.)
1. It is logically possible for the table to have 8 sides.
  - \* 2. It is logically impossible, if the table is rectangular, for it to have 8 sides.

-----

∴ The table is not rectangular



# What's wrong with the argument?

- There are various ways to look at it. If it seems valid at first sight, it may be because we're reading it like this:

$\Diamond$  possibly

1.  $\Diamond D$

D = "I can exist disembodied"

2.  $B \rightarrow \neg \Diamond D$

B = "I am simply a body"

-----

$\therefore \neg B$

MT: 1, 2 -

1.  $\Diamond D$

2.  $B \rightarrow \neg \Diamond D$  ?

-----

$\therefore \neg B$

The second premise is incorrectly stated here. It should be written  $\neg \Diamond (B \ \& \ D)$ , or  $\Box (B \rightarrow \neg D)$ .

$\Box$  It is necessary that

- $\Diamond D$  means that  $D$  is consistent with the knowledge  $K$  that I have right now. (We could write it  $\Diamond_K D$ .)
- $(B \rightarrow \neg \Diamond D)$  is false, because  $B$  being true doesn't make it part of my knowledge  $K$ . Even if  $B$  is true,  $D$  remains consistent with  $K$ .

1.	$\Diamond D$	$\sim$	$D$ is true in $\geq 1$ poss. world.
2.	$\Box(B \rightarrow \neg D)$		
3.	$\underline{B}$		
4.	$B \rightarrow \neg D$	$\Box$ Elim:	2
5.	$\neg D$	$\sim$	$D$ is false in @.
	$\underline{1}$		Can't get.
	$\neg B$	$\neg$ Intro	(1?)