

# Review Sheet

Philosophy 610QB  
Spring 2013  
Professor Strawson

*By Nikhil Garg and Will Hoenig*

Note: This document is viewable by all, but edit permissions are restricted. This is not in any way done to discourage collaboration/suggestions/contributions to this document, but rather an attempt at quality control (accuracy, maintaining consistent structure, avoiding repetition). If you would like to contribute to the document, please contact us. In addition, if the TA session review documents ever take off and start getting filled, we will be using its material at our discretion, citing it where appropriate. Our goal is to make an effective study aid without unnecessary clutter.

Secondary Note: For the most part, the content is written in our own words, but we link to either class material or other sites from which we drew material. In some cases, we copy useful blocks of paragraphs from external sites, citing where appropriate. We do so only when we feel that it is relevant/extremely clear/well said. Regardless, these notes are not just copied class/reading notes, but are written with clarity and succinctness as the goal. As such, there is not a one-to-one mapping to what was discussed in class.

Tertiary Note: Though we have done our best to be comprehensive and accurate, we may have overlooked something. Furthermore, some of the material is subject to interpretation, and our interpretation may be different than those of others. In other cases, we may just be wrong. For this reason, we encourage you to do the readings yourself and to develop your own understanding, in addition to using this document.

If you find any mistakes, please contact us.

# TABLE OF CONTENTS

[Color](#)

[Problems of Philosophy of Mind](#)

[What is the Philosophy of Mind?](#)

[Relevant problems](#)

[General Terms and Definitions](#)

[Thought Experiments](#)

[What do we know Consciousness/Mental States?](#)

[Behaviorism#](#)

[Wittgenstein-ein theorists](#)

[Objections to Behaviorism](#)

[Functionalism](#)

[Behaviorism and Functionalism: Epistemological or Metaphysical?](#)

[Real Realism ®](#)

[Problem of Other Minds](#)

[A Problem, or is it One?](#)

[Argument from Analogy](#)

[Criticism of Argument from Analogy](#)

[Famous Philosophers' Views on the problem](#)

[Mind-Body Problem](#)

[Dualism](#)

[Descartes](#)

[Antoine' Arnauld's 4th Set of Objections to Descartes' arguments](#)

[Interaction Problem](#)

[Descartes on Interaction \(Correspondence with Princess Elizabeth\)](#)

[Epiphenomenalism ®](#)

[No Interaction \(Parallelism\)](#)

[Miscellaneous](#)

[Monism](#)

[Neutral](#)

[Objection to Neutral Monism](#)

[Idealism/mentalism/Immaterialism](#)

[Mentalism](#)

[Idealism](#)

[Physicalism/Materialism/Eliminativism](#)

[Physicalism](#)

[Professor Strawson's Criticism](#)

[Panpsychism](#)

[Realism, Monism and Materialism](#)

[Dual Aspect Theory](#)

[Self-Consciousness/Awareness of Awareness](#)

[Personal Identity](#)

[Bodily Continuity](#)

[Bodily Replacement Objection](#)

[Person/Body Separation Objection](#)

[Psychological Continuity \(Continuity of Memories\)](#)

[Direct Continuity](#)

[Indirect Continuity](#)

[Amnesia Concern](#)

[John Locke's Views on Personal Identity](#)

[Issues with Locke's View](#)

[Locke's Definition of a Person](#)

[David Hume's Views on Experience and the Mind](#)

[Most Philosophers' Interpretation](#)

[Professor Strawson's Interpretation](#)

[Hume and E-S-O Moves](#)

## Color

The class began with a discussion about the nature of color. It served as a general introduction to the philosophy of mind, and it has significance with many questions: those of the existence of other minds, the existence of mental states, the nature of qualia, and basically everything else in the class.

In the first week, Professor Strawson got the class to accept a functionalist definition of color - that everyone 'knows' the **same** color red even under the inverted spectrum conditions, in which one may experience a qualitatively different color. This acceptance broke down when he moved from inverted spectrum to sound-as-color, though they are logically the same. The class also took issue with deeper brands of functionalism that reduce consciousness down to causal, mechanistic reaction systems.

John Locke seems to advocate a definition of color in which the important thing is whether the material differentiates between things. The qualitative feeling is right for each individual person, regardless of whether the feelings are the same between people. The important aspect is the relation between colors and the objects from which they seem to come, and how color distinguishes those objects.

## Problems of Philosophy of Mind

### What is the Philosophy of Mind?

**Philosophy of mind** is a branch of [philosophy](#) that studies the nature of the [mind](#), [mental events](#), [mental functions](#), [mental properties](#), [consciousness](#), and their relationship to the physical body, particularly the brain.<sup>1</sup>

### Relevant problems

How do we define mental states and consciousness? What is the criteria for consciousness (ie, how do we determine whether someone - or something - is conscious?)

Problem of Other Minds

The Mind/Body Problem

---

<sup>1</sup> [http://en.wikipedia.org/wiki/Philosophy\\_of\\_mind](http://en.wikipedia.org/wiki/Philosophy_of_mind),

## General Terms and Definitions

1. Abduction<sup>2</sup>: Also called Inference to the Best Explanation, is a type of inference is which claims that have the best explanation are preferred. Even if the conclusion does not logically follow from the premises, it can be inferred to be true, because it is the best explanation.
2. Determinable: Something that is implemented by a determinate.
3. Determinate: An implementation of a determinable. For instance, if color is a determinable, red is a determinate. Something can be both a determinate in some circumstances and a determinable in others. For instance, red can be the determinate of color, the determinable, or red can be a determinable, implemented by the determinate crimson.
4. Emergence, Normal: Properties that are not found in individual particles can be found in collections of those particles. For instance, liquidity cannot be found in one water molecule but can be found in a group of water molecules.
5. Emergence, Radical: A more extreme version of the above. When particles put together have radically and inexplicably different quantities than the individuals. Professor Strawson says that consciousness out of inexperiential matter would be such a case. He also says that radical emergence does not exist.
6. E-S-O Move: Epistemological claim -> semantic claim -> ontological claim. Works something like this:
  - a. Epistemological claim (E): All we can know of the world around us is our perceptions of it.
  - b. Semantic claim (S): All that we can *genuinely mean* about the world around is our perceptions.
  - c. Ontological claim (O): Perceptions are all that the world around us really is.This is a rather questionable logical move. Many philosophers have accused David Hume of employing such moves, but Professor Strawson thinks that Hume tended to employ E-S-OE moves instead. E-S-O moves make fairly broad claims about the universe that hinge upon the ability of humans to understand things. They are widely criticized because of that
5. E-S-OE Move: Epistemological claim -> semantic claim -> ontological/epistemological claim. Like an E-S-O move, but with an epistemological qualification on the ontological part. Works something like this:
  - a. Epistemological claim (E): All we can know of the world around us is our perceptions of it.
  - b. Semantic claim (S): All that we can *genuinely mean* about the world around is our perceptions.
  - c. Ontological/epistemological claim (OE): Perceptions are all that the world around us really is *so far as we have any idea/notion of it*.The epistemological qualification to the ontological claim removes main criticism of E-S-O moves (described above).

---

<sup>2</sup> <http://plato.stanford.edu/entries/abduction/>

6. Induction, Weak - when  $P(n)$  true, then  $P(n+1)$  is true
7. Induction, Strong -  $P(k)$ ,  $P(k+1)$ ... $P(n)$  are true, then  $P(n+1)$  is true
8. Instrumentalism
9. Mirror Neurons:
10. Operationalism
11. Phantom Limb Pain:
12. Phenomenalism
13. Positivism
14. Q-memory<sup>3</sup> - One is 'q-remembering' if one thinks about a past experience in a way that seems like a memory and that thought is causally dependent on someone having that experience. However, there is no requirement for person having the experience is the same as the person with the memory
15. Rubber hand Illusion:
16. Reductionism: a general activity of explaining complex ideas in terms of more fundamental parts. For instance, one may employ a reductionist mindset to reduce chemistry down to just an application of physics. In philosophy of mind, reduction is often associated with physicalism.
17. Regularity theory of causation
18. Solipsism: The belief that only one's mind is sure to exist. This comes in many varieties, which can be broadly categorized into metaphysical solipsism and epistemological solipsism. Metaphysical solipsism is the belief that one's own mind is all that exists, while epistemological solipsism is all that can be known to exist. A flavor of solipsism that was discussed in class is present-moment solipsism, or the belief that all that can be known that exists, or all that exists (depending on if you are implementing it metaphysically or epistemologically), is one's own mind in the present moment.
19. Turing Test - A criterion for consciousness devised by British computer scientist Alan Turing in 1950. It states that a computer can 'think' if a person, upon communicating with both the computer and another person, is unable to distinguish between the two.
20. Verificationism
21. Zombies, Australian - A perfect physical duplicate of a human being but not conscious. The answer to the question "Are Australian Zombies possible?" has great impact on the theories of consciousness and the mind/body problem.
22. Zombies, Classical - A being that behaves indistinguishably from a conscious human being, but which may be internally composed of different materials and does not experience qualia.<sup>4</sup>

## Thought Experiments

1. Brain in a vat
  - a. Experiment: Your brain has been removed from your body and been attached to a computer which simulates the world.

---

<sup>3</sup> <http://www.uark.edu/campus-resources/rlee/intrsp03/oh/pft-q-mem.html>

<sup>4</sup> [http://en.wikipedia.org/wiki/Philosophical\\_zombie](http://en.wikipedia.org/wiki/Philosophical_zombie)

- b. Importance: The experiment raises questions about the existence of external reality and the problem of other minds.
  - c. Other: Effectively identical to Descartes' Evil Demon.
2. Chinese Room
- a. Experiment: There is a room in which there sits a man who knows no Chinese. From an open slit on the door, he is given sheets of paper with Chinese on it. Using a giant book, he looks up the symbols and writes down a reply, also in Chinese, as specified in the book. He then sends his reply out the door.
  - b. Takeaway: There is a strong intuition that the person in the room, and the room itself, does not know Chinese. However, it behaves indistinguishably from a person who does know Chinese, and so, according to behaviorist theories of thought, the man knows Chinese.
  - c. Importance: The experiment is a strong argument against the consciousness of artificial intelligence, and behaviorism in general.
  - d. Criticisms of the experiment:
3. Chinese Nation
- a. Experiment: The entire country of China, and all its people, are being used to simulate the function of a human brain. Each person acts as a neuron in the brain, and different parts of the country make up different parts of the brain, reacting to different inputs from the 'senses'.
  - b. Takeaway: The 'mind' reacts to inputs as a brain would, even simulating pain. However, there is a strong intuition that the entire system does not feel pain, that the simulated neural actions do not make consciousness.
  - c. The Chinese Nation experiment is an argument against functionalism, which argues that the same mental states can be realized in different ways. If that were true, the realization of the mind on top of a 'brain' composed of a network of people would be the same as the realization of the mind on a human brain. Some functionalists, like Daniel Dennett, accept that conclusion.<sup>5</sup>
4. Inverted Spectrum
5. Leibniz's Mill<sup>6</sup>
- a. Experiment: Suppose there was a machine that behaves exactly like a conscious being. If it were increased to the size of a mill so that one could walk inside, one would not find anything that explains consciousness, perception or qualia.
6. Mars Experiments<sup>7</sup>
- a. Experiment one: You are transported to Mars via teletransporter. The teletransporter creates a digital description of your matter down to the atom. It destroys your body on Earth, and sends that digital description to Mars, where an exact copy of your body on Earth is produced using the description.
    - i. One's instinct in this case would probably be to conclude that one transported

---

<sup>5</sup> [http://en.wikipedia.org/wiki/China\\_brain](http://en.wikipedia.org/wiki/China_brain)

<sup>6</sup> <http://plato.stanford.edu/entries/chinese-room/#1>

<sup>7</sup> Derek Parfit reading

via teletransporter is the same person on both ends of the voyage, and that the teletransporter can thus be considered a mode of transport. The Martian copy would be psychologically continuous with the original (as is described in Parfit's story), there would be a causal connection between person on Earth and their clone on Mars, memories would transfer from Earth to Mars, et cetera.

- b. Experiment two: The same process happens as in experiment one, except for the fact that your body on earth is not successfully destroyed. Thus, there are two copies of you: one on Earth and one on Mars.
    - i. Things become a little more difficult here. Which person has the correct claim of being the real "you"? If the Earthly you has that claim, that presents a problem, because the Martian you has done nothing to deserve being considered less "you" than it did in experiment one.
  - c. Experiment three: The same as experiment two, but, in this case, the teletransporter causes massive cardiac damage to your Earthly body that is going to cause your Earthly body to die from cardiac arrest in a matter of days.
    - i. This presents the same issues as experiment two, but with a twist. When Earthly you dies, will it make any difference to the external world, since there is another you on Mars?
  - d. To Parfit, these experiments are useful for "discover[ing] what we believe to be involved in our own continued existence, or what it is that makes us now and ourselves next year the same people."
7. The Prince and the Cobbler<sup>8</sup>
- a. Experiment: A cobbler wakes up in the body of a prince, and that prince wakes up in the body of that cobbler. That is the basic premise, and many "what-if" extensions can be added. For instance: if the cobbler stole from the castle, and is currently being searched for to be put on trial, who should bear the responsibility for the crime? Who is the prince, and who is the cobbler?
  - b. John Locke devised this experiment to show his belief in the distinction between "man and person". To him, it is (at least logically) possible for personal identity to separate from bodily identity.
8. Ship of Theseus<sup>9</sup>
- a. Experiment: The Ship of Theseus is an ancient ship based in Athens. Over the years, as its part age, its parts are replaced with new parts, to the point where, eventually, the ship has none of its original parts. Is it the ship still the Ship of Theseus?
  - b. Extension: These parts are replaced while they still have some life in them. For whatever reason, as parts of the ship are replaced, someone hoards all of the ship's old parts. Eventually, that person rebuilds a Ship of Theseus using the original parts. Now which ship is the Ship of Theseus?
  - c. This thought experiment makes one think about the relationship between material

---

<sup>8</sup> <http://plato.stanford.edu/entries/locke/supplement.html>

<sup>9</sup> <http://plato.stanford.edu/entries/material-constitution/>



composition and identity. There are several ideas to solve it. For instance, one could argue that, as a part of the Ship of Theseus is removed, it ceases to be a component of the Ship of Theseus, and its replacement becomes a part of the Ship of Theseus instead. It could also be argued that the definition of the Ship of Theseus is that the Ship is simply the object which continuously holds that identity to those invested in its identity. The extent to which those definitions are actually different, however, is questionable. One could also argue that language simply cannot resolve this conflict.

9. Swampman<sup>10</sup>

- a. Experiment: You go on a hike in a swamp and are struck and killed by lightning. Simultaneously, another strike hits the swamp and rearranges molecules such that they take on the exact form of you, including an exact copy of your brain. This swampman behaves exactly as you do (goes to your work, hangs out with your friends, etc).
- b. Davidson, who proposed the thought experiment, argues that the swampman and his actions/words have no meaning because the swampman has no causal history. The swampman experiment, articulated this way, is an argument against functionalism/behaviorism, under which the swampman would be literally the same person as you (due to behavior), even though there was no causal connection between you and swampman.

10. Twin Earth

11.

## What do we know Consciousness/Mental States?

The theories that follow answer the epistemological question about what we can know and say about consciousness, if we can know/say anything at all. Some theories want to reduce consciousness down to observable events, while others argue that we have direct access to an undeniable consciousness.

### Behaviorism<sup>11</sup>

We are what we do. Human beings can be studied and described completely by their inputs and outputs. It rejects the use of internal, unobservable, subjective experiences and thought processes in favor of events that are scientifically observed.

Under this view, a computer that can pass the Turing Test is considered conscious. Under behaviorism, Classical Zombies are conscious, and therefore not 'zombies' at all. Australian Zombies would meet a Behaviorist definition of consciousness as well, and in that sense, would also not be 'zombies'. Both types of Zombies would pass the Turing Test, and be outwardly functionally identical to human beings; Australian Zombies would just be composed of normal biological matter while Classical Zombies could be made of anything. It is worth noting that Behaviorists are quite often

---

<sup>10</sup> <http://en.wikipedia.org/wiki/Swampman>

<sup>11</sup> <http://www.iep.utm.edu/behavior/#H1>, <http://plato.stanford.edu/entries/functionalism/#WhaFun>,

Physicalists as well, of the variety that would believe Australian Zombies to be just the same as human beings in all possible ways.

### **Wittgenstein-ein theorists**

Wittgenstein, a philosophical behaviorist, approaches the problem by analyzing the structure, content, and ability of language.

Wittgenstein-ein theorists argue that we are only aware of the word 'pain' by learning it from others who displayed outwardly observable signs of pain. We cannot say anything about pain except by referring to these outward signs, since that is how we learned the word in the first place. Moreover, we cannot discuss the essential nature of pain except by including outward appearances.

More specifically, from Professor Strawson<sup>12</sup>, the claim is:

reference to publicly observable circumstances will enter essentially into any satisfactory account of the correct circumstances of use or application of the word  
→ it will therefore enter essentially into any satisfactory account of the meaning of the word  
→ that it will enter essentially into any satisfactory account of the essential nature of what the word is a word for.

Note: Wittgenstein does not want to say that pain is just pain-behavior. He admits that pain exists ("Not a Something, but not a Nothing Either"<sup>13</sup>), just that we cannot say anything about it except through pain-behavior. We saw (and accepted) a similar view concerning color in the first days of class. We accepted that we couldn't say anything about color except through our use of color.

Note: In this subsection, we use the word 'pain' for historical and conceptual reasons. The same argument can be extended to all qualia, and consciousness in general.

### **Objections to Behaviorism**

For objections to this view by Professor Strawson, read *Mental Reality* pg 215-26 (linked on Blackboard)

### **Functionalism**

Functionalism is very similar to Behaviorism; the distinction between the two is subtle. Broadly speaking, Functionalists believe that mental states exist simply to process outside stimuli, and to provide responses to them. However, Functionalists do admit that the mental states exist, while Behaviorists may not.

---

<sup>12</sup> Strawson, *Mental Reality*, pg 215-26

<sup>13</sup> *Philosophical Investigations* Sect. 304

## **Behaviorism and Functionalism: Epistemological or Metaphysical?**

It can be difficult to exactly define Behaviorism and Functionalism because they can be used either as epistemological arguments or as metaphysical arguments, depending on the flavor.

An epistemological Behaviorist may say: “all we can know about the mind is that it behaves in certain ways in reaction to outside stimuli.”

An epistemological Functionalist may say: “all we can know about the mind is that it behaves in certain ways in reaction to outside stimuli, using some sort of real mental state.”

A metaphysical Behaviorist may say: “all that exists in the mind is a process for reacting to external stimuli.”

A metaphysical Functionalist may say: “all that exists in the mind is a process for reacting to external stimuli, which expresses itself internally with some sort of real mental state.”

In effect, the metaphysical versions of Behaviorism and Functionalism are E-S-O moves from their epistemological counterparts.

So, are Behaviorism and Functionalism epistemological theories or metaphysical theories? According to Professor Strawson, they started out as epistemological theories, but have been adopted as metaphysical theories more and more.

Epistemologically, Behaviorism and Functionalism are often used to examine whether something is conscious or not. An epistemological Behaviorist, for instance, would likely claim that a computer with behaviors so human-like that it appears conscious could not to be determined to be *not* conscious. A metaphysical Behaviorist in that instance would claim that such a computer would indeed *be* conscious.

## **Real Realism<sup>®</sup>**

Strawson’s term for those who (like him) do not try to explain away (or looking-glass) the strong intuition that consciousness, as those who do not do philosophy understand it, exists. Real Realists recognize that qualia exist: there is more to life than behavior, that the feeling of pain/color/perception is real. Real Realist theories can be dualistic or monist, as long as they recognize that mental states as we understand them exist.

For Real Realists, there is no denying a feeling, because seeming to perceive something is the same thing as perceiving something. If I think I am in pain, and feel pain, then I am in pain. No one can trick me into thinking I am in pain when I am not, because if I think I am, I am.

Whereas behaviorists and functionalists try to reduce consciousness down to its observable relation to physical substances, Real Realists recognize that conscious states are more than that.

Note: Professor Strawson uses the term 'Real Realists,' rather than 'Realists,' because the latter has been co-opted by functionalists who claim that they take into account consciousness.

Secondary Note: In this document, Real Realist theories will be marked as ®

## Problem of Other Minds

How do one know that other people are conscious? How does one know that other people have the same qualitative perceptions/feelings (of pain, color, etc) that s/he has?

The general question:

1. Can you know other humans exist?
2. Can you know they have experiences?
3. Can you know they are having particular experiences at particular times?
4. Can you know what those experiences are like for them?

The problem of other minds deals primarily with [2], [3], and [4], but [1] is required for us to even formulate the question (If others don't exist, how can they be conscious?)

### A Problem, or is it One?

There are 2 schools of thought concerning the importance of finding a proof for the existence of other minds

1. The Kantian view, that we must find such a proof
2. The Heidegger view, that it is obvious that other minds exist, and searching for a proof misses the point.
  - a. Professor Strawson agrees, but for different reasons. He thinks that it can't be proven, that it is a philosophical mistake to pursue it, and this skepticism is often given too much weight.
    - i. How likely is it that I am the first case of consciousness? We are fundamentally similar to other humans in so many ways (though our observations are necessarily physical rather than mental), why would consciousness be any different? The burden of proof should be on the other side, which maintains that other minds do not exist.

The primary proof for the existence of other minds comes from the Argument From Analogy.

## Argument from Analogy

Also referred to as the inductive argument, the argument is that I know other people are conscious because they act like I do, and I know that I am conscious.

More Specifically:

Other's have feelings because

1. They have bodies (the antecedent to feelings)
2. They exhibit outward signs that I know to be caused by feelings
3. In other words: I have evidence of 1st and last links (they have body → ? → they have experiences) but not have middle of feelings

Thus: Since the first and last links are directly observable, I can assume that they have a mind/consciousness like I do to fill in the gap.

## Criticism of Argument from Analogy

1. Standard arguments against induction, especially since this argument is a form of weak rather than strong induction. Because we only know 1 case of consciousness (our own), induction to the general case may be fallacious.
2. Wittgenstein's answer: The analysis is meaningless, since I only know the middle link in reference to how it feels to me.

Wittgenstein's famous rejoinder "But if I suppose that someone has a pain, then I am simply supposing that he has just the same as I have so often had." That gets us no further. It as if I were to say: "You surely know what 'It is 5 o'clock here' means; so you also know what 'It's 5 o'clock on the sun means. It means simply that it is the same time there as it is here when it is 5 o'clock."  
(Wittgenstein 1953, 350)

Another Wittgenstein example: The beetle in the box thing as discussed in class.

## Famous Philosophers' Views on the problem

1. Kant - We must find a proof that other minds exist.
2. Heidegger<sup>14</sup> - Problem is not epistemological but rather ontological. That is, we shouldn't be looking for a proof that other minds exist, because we exist with others as part of our own existence. We directly experience others, and a "radical asymmetry" between our experiences

---

<sup>14</sup> <http://plato.stanford.edu/entries/other-minds/>

and those of others does not make sense.

3. Nagel - Problem is not epistemological but conceptual. How do we *understand* that others have their own mental states? My own thought processes and understanding are so colored with my own perspective that I cannot understand the mental states of others. This conceptual problem is considered harder than the epistemological problem.
4. John Stuart Mill - Argument from analogy view
5. Wittgenstein - Similar to Nagel's view. One can't really think that other people have experience because one grounds one's conception of what experience is on one's own case. One's experience is stamped 'mine' in a solipsistic that makes it contradictory or unintelligible to think that other people might have it. See criticisms of Argument from Analogy for more.
  - a. Note: Professor Strawson doesn't buy this argument. "One could think that things might be for others something like what it is for us."

## Mind-Body Problem

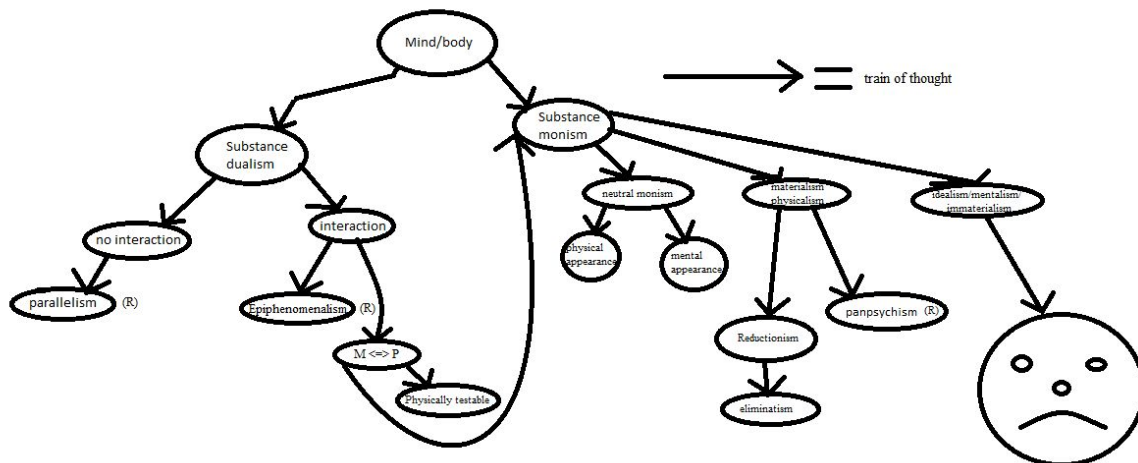
How are the mind and the body related? Are they two separate substances with different essential properties, or are they the same substance? If they are different substances, how do they interact and causally affect one another?

The theories that follow answer the metaphysical questions concerning what substances exist and what essential properties they hold. While they are connected to the epistemological problem of *knowing* what the mind is like, there is not a strict mapping from one to the other. That is, two people may hold the same view on the mind-body problem but different ones concerning what we can know about consciousness (and vice versa).

Will's diagram of different positions (a slightly modified version of what Professor Strawson created on the board):<sup>15</sup>

---

<sup>15</sup> For a close up view, look here <http://img825.imageshack.us/img825/8466/20130205.png>. Also, Will apologizes (but doesn't really apologize) for the frowny face in the bottom-right corner.



## Dualism

Dualism is the view that there are two substances, the mental and the physical. These substances have their own, separate, essential properties and may/may not interact, depending on the type of Dualism. Descartes has the most famous initial formulation of the theory and was the first to tackle the Mind-Body Problem as it exists today.

## Descartes

In the *Meditations*, Descartes searches for the meaning of the 'I', beginning with his famous Cartesian doubt and then progressing from there, eventually concluding in the 6th meditation that the mind/body are separate substances that are intimately connected.

### 2nd Meditation

1. 'I' necessarily exists
  - a. I cannot be deceived of this fact
2. I know the essential properties of the body and the mind
  - a. 'I' is a think that thinks. "A thing that doubts, understands, affirms, denies, is willing, is unwilling, and also imagines and has sensory perceptions"<sup>16</sup>
  - b. Body occupies space (essential characteristic of extension), has definable location, can be perceived by senses, moves by external force (is not a self-moving things)

### 6th Meditation

1. Arguments that the mind/body are separate
  - a. I have a clear and distinct idea of myself as a thinking, non-extended thing, and a separate distinct idea of the body as an extended, non-thinking thing.

<sup>16</sup> Descartes 2nd meditation

- b. Mind is indivisible, Body is divisible
  - c. I have a complete understanding the mind (remember the list above), and the body is not part of that understanding
    - i. Note: this argument is strongly attacked by Antoine Arnauld in The 4th set of objections, posted on Blackboard.
2. There is an intimate union between the mind and the body - "I and the body form a unit"<sup>17</sup>

*Antoine' Arnauld's 4th Set of Objections to Descartes' arguments*

1. It does not follow from "I am not aware of Something x belonging to my essence" that "x does not belong to my essence." It's possible that I am simply unaware of that something.
  - a. Example: Triangle Example by Arnauld. I can clearly and distinctly understand a right triangle. However, at the same time, I can simply not be aware of a resulting property of the right angled triangle. According to Descartes, it would follow that the right triangle does not have that property.
  - b. Descartes tries to dismiss this argument by creating a dichotomy (one that is most likely false, but at the least confused) between complete understanding and adequate understanding. We may not have a wholly adequate understanding of something (only an omniscient God can), but we can have an adequate enough understanding to be called complete. Then, if I have a complete (adequate enough) understanding of my essence, and the body is not part of that understanding, then the body is not part of my essence.
2. Circularity argument: Descartes established the Cogito argument by regarding himself as a thinking thing (as long as an I thinks, then I know that I exist). However, if we validate our own existence in such a way, then it would be circular to think of myself as only a thinking thing, *because that idea itself came from my thoughts*. Evidence solely from one's thoughts is not evidence, because that is already presupposing the nature of myself.
3. Arnauld also presents an alternative: It may be possible to theoretically abstract the thinking thing away from the body, as Descartes does, but it's simple that: a theoretical abstraction. Physical (bodily) attributes may still necessarily belong to the thinking thing.
  - a. He uses the example of a line: Any line can be abstracted to only have a length (we only consider 1 dimension), but it necessarily has 2 dimensions (a width). Just because in math we abstract away the width (regarding it as infinitesimal) does not mean that it's not necessary.

**Interaction Problem**

If the mind and the body are two separate substances with different essential properties, how do they communicate and causally affect one another?

---

<sup>17</sup> Descartes 6th meditation



*Descartes on Interaction (Correspondence with Princess Elizabeth)*

Princess Elizabeth of Bohemia and Descartes corresponded through letters, in which she tried to get Descartes to answer the interaction problem. He couldn't, and Elizabeth concluded that it is possible that the soul has essential components that are extended.

Here, we give a summary of the correspondence and Descartes' multiple (unsatisfactory) attempts to neutralize the interaction problem.

Elizabeth: Causing the body to move requires either extension or physical contact, neither of which is a property of the mind. Then, how do mental substances cause the body to move?

Descartes': He tries to avoid the question, saying it basically just happens. - It doesn't make sense to ask how that happens, because the union is a primitive notion that cannot be described in terms of other primitive notions (the notions of the essential properties of mind/body).

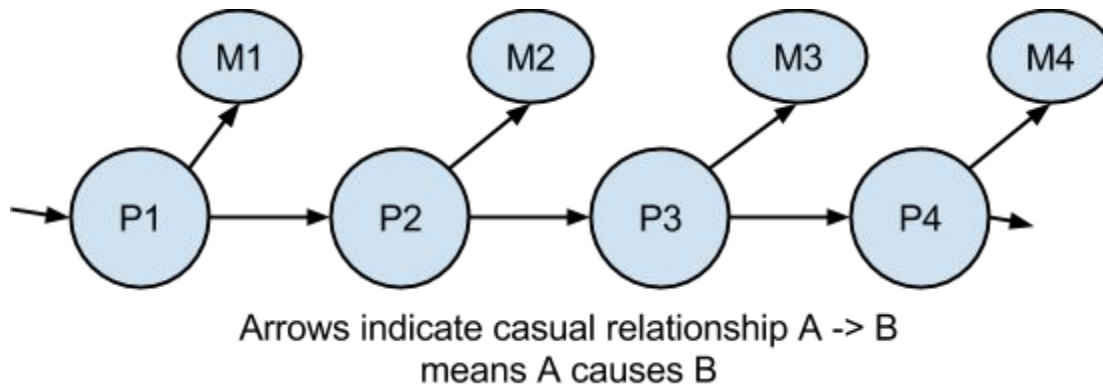
Elizabeth: I don't understand. Given what Descartes has said about the essential properties of mind/body, interaction is hard to understand.

Descartes: ... repeats other reply

Elizabeth: You haven't convinced me. It's entirely possible that the soul has essential extended components (more possible than a non-extended thing moving an extended thing).

*Epiphenomenalism* ®

"the theory in [philosophy of mind](#) that mental phenomena are caused by physical processes in the brain or that both are effects of a common cause, as opposed to mental phenomena driving the physical mechanics of the brain"<sup>18</sup>.



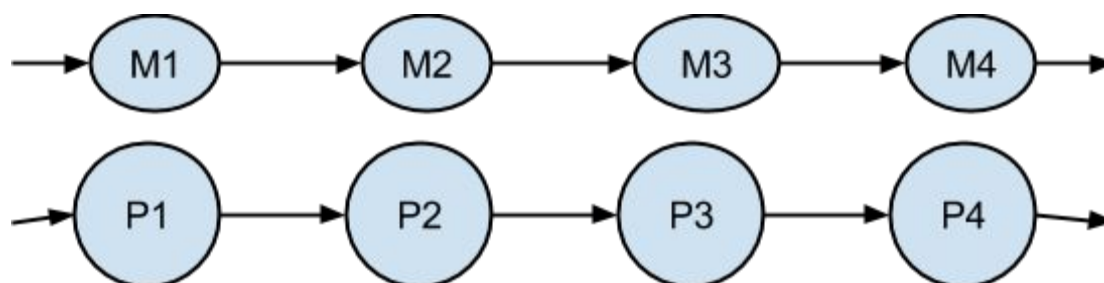
<sup>18</sup> <http://en.wikipedia.org/wiki/Epiphenomenalism>

If mental processes are causally after or simultaneous to the physical processes to which they are related, then they don't do anything (can't affect the physical processes)

Professor Strawson's criticism:

When you want something and act to do it, then according to the theory the mental state of wanting had no impact on the action. This is highly unintuitive.

*No Interaction (Parallelism)*



This view simply denies the interaction problem by saying that the 2 substances don't interact. Physical causation is independently complete, as is mental causation. (No interaction, no causation). A rather ridiculous view, would suggest that someone wanting to do something, and then physically doing it, have nothing to do with each other.

## Miscellaneous

*Dualism: Empirically Testable?*

It has been theorized that Dualism is empirically testable. If, the argument goes, it is possible to build a machine that can scan the brain and determine how every single particle in it is behaving, then proving Dualism would be simple. If the machine can find a place where matter is impacted by a force that cannot be explained by interaction with other matter, then Dualism is proven, because physical causation would have been proved to not be complete.

## Monism

Monistic theories hold that there is only 1 fundamental substance. This one substance makes up the universe, consciousness, observable phenomena, and anything else that may/may not exist. Different Monist theories disagree concerning the nature of the singular substance.

## Neutral

Neutral monism, as its name suggests, is neutral between the mental and the physical. More specifically, the claim is that the 1 substance is neither physical nor mental, but rather, physical and mental states are composed of, and can be reduced to, the neutral substance. The neutral substance

acts as either physical or mental “on which other groups it happens to be connected with.”<sup>19</sup> Physical and mental states are then state appearances derived from this neutral substance.

### *Objection to Neutral Monism*

The same objection that applies to all theories that try to reduce consciousness/mental states to some other type of state applies here. “To seem to be means there just is.”<sup>20</sup> That is, we know that pain exists, because we have pains. Appearance of pain is pain. We cannot ‘appear’ to have a mental state without actually having it. Neutral Monism is not real realist.

## **Idealism/mentalism/Immaterialism**

This is a rather broad set of ideas that all focus, to some extent, on the belief that the essential nature of the universe lies outside of the physical. An important philosopher from last semester to remember in association with these ideas is Bishop Berkeley.

There is a subtle but important distinction between Mentalism and Idealism.

### *Mentalism*

Mentalism is the belief that the world is only composed of mental substances and ideas. This is probably where Berkeley could best be described as falling. It admits that the mind of the perceiver exists (in other words, it doesn’t deny Descartes’s Cogito), but it denies the existence of a world outside of the mind.

### *Idealism*

Idealism is more extreme than Mentalism. Idealism is the belief that everything in the universe is ideas. It goes as far as to claim that “mental substances” do not exist. In effect, it denies the Cogito. If this seems hard to imagine, then you are not the only one. It is difficult to imagine that anyone would be a true Idealist.

## **Physicalism/Materialism/Eliminativism**

Broadly speaking, philosophers in this category seek to explain all mental phenomena in purely physical terms. There are multiple categories of this, why are very broadly clumped into two.

### *Physicalism*

Physicalism (note the bolded s) is a name informally given by Professor Strawson to a belief that he deeply disagrees with. It is strongly associated with Reductionism and Eliminativism, although it probably would not be appropriate to refer to this belief as Reductionism or Eliminativism (which are ways of approaching philosophy and the world, rather than belief systems).

Physicalism contends that all mental processes can be reduced down to physical properties, **and** that

---

<sup>19</sup> <http://plato.stanford.edu/entries/neutral-monism/>

<sup>20</sup> Professor Strawson, in class.

those physical properties can all conceivably be described scientifically. In other words, what we think of as the mental is simply the interaction of physical processes, and those physical processes (and their components) are within the reach of description by science. They have no inherently mental qualities.

**NOTE:** Physicalism is not a formal, technical name for this belief. Rather, there are two branches of Physicalism (without the middle s) discussed in class. One is Panpsychism, and one is this. This one was not given an official name, but was sometimes referred to as Physicsalism. That is because Physicsalism entails the belief that the physical has the potential to be explained entirely by physics.

Most of the world would consider what Professor Strawson calls Physicsalists to be simply Physicalists. However, as Professor Strawson categorizes Panpsychism as another branch of Physicalism, it is necessary to categorize what is normally considered Physicalism (i.e., Physicsalism) as a branch of Physicalism. As Panpsychism is a type of Physicalism, it is definitely not wise to take Physicsalism as representative of all of Physicalism.

Most of the world would consider what Professor Strawson calls Physicsalists to be simply Physicalists. However, as Professor Strawson categorizes Panpsychism as another branch of Physicalism, it is necessary to categorize what is normally considered Physicalism (i.e., Physicsalism) as a branch of Physicalism. As Panpsychism is a type of Physicalism, it is definitely not wise to take Physicsalism as representative of all of Physicalism.

#### Professor Strawson's Criticism

Professor Strawson ties Physicsalism strongly with Functionalism and Behaviorism. While Functionalism and Behaviorism are epistemological theories, and Physicsalism is a metaphysical theory, Professor Strawson criticized Physicsalism in ways similar to those in which he criticized Functionalism and Behaviorism.

His argument is roughly as follows. While Physicalism does not absolutely deny the existence of consciousness, the type of consciousness that it allows is sufficiently devoid of meaning to be considered "looking-glassed". Physicsalism allows for a version of consciousness that aids in reaction to physical sensations, but it does not provide any mechanism by which physical sensations are turned into qualia, or, indeed, for those qualia to exist. That, in Professor Strawson's view, is not real consciousness.

#### *Panpsychism*

Panpsychism is Professor Strawson's belief. It is a branch of Physicalism which says that all mental activities/properties can be reduced to physical ones, **but** that the physical world which underlies the mental world cannot conceivably be understood with physics, or any sort of science. Instead, in this view, the physical is fundamentally *mental* if broken down. Furthermore, all physical objects are subjects of experience. This is monism, not dualism, because the only sort of substance that it fundamentally allows for is the mental.

The argument in favor of Panpsychism, in its basic form, is as follows. Assume that the universe is made of only one substance and that substance is physical (one can concoct a fairly convincing argument for this). Also assume real realism about experience (i.e. that experience of qualia exists and is what we instinctively think that it is). If we assume those two things, it follows that experience is a wholly physical phenomenon. Now assume that radical emergence does not exist (see the definition of radical emergence towards the beginning of this document). If experience is physical, and experience cannot emerge out of parts that are wholly non-experiential, then physical matter must be, in some way, fundamentally experiential.

That argument shows that matter is fundamentally experiential, or mental. There is a more detailed set of arguments that get from that to panpsychism, although the above proof that physical matter is mental is the most important. Below is an argument to close the few remaining gaps.

If the physical is fundamentally experiential, and physical reality can be broken down into some sort of smallest component, it follows that the smallest component of matter is necessarily experiential. If the smallest component of matter necessarily experiential, and all experience requires a subject of experience, it follows that the smallest, **most fundamental building blocks of matter are subjects of experience. That is the strictest definition of Panpsychism.**

#### Realism, Monism and Materialism

Panpsychism is a view that is realist (recognizes consciousness as real), materialist (physical and in accordance with physics) and monistic (there is only 1 substance), according to Professor Strawson. A consequence of such a view is that experiential phenomena are real and somehow include the consciousness that we are all so intimately familiar with. For Professor Strawson, materialists who are not realist materialist haven't "got to the starting line" in thinking about their materialism, because they have not "faced the thought that consciousness is a wholly physical phenomenon."

Note: physics here is consistent with the everyday use of the world physics, but adds subjective aspects to the normal characteristics.

## Dual Aspect Theory

To be added later.

## Self-Consciousness/Awareness of Awareness

There are competing definitions for self-consciousness<sup>21</sup>. They include:

- In philosophy, many link self-consciousness to the ability to think "I" thoughts. In this view, one must be able to conceptualize oneself as oneself to be considered self-conscious.
- A related view in philosophy is that, for one to be self-conscious, one must be capable of

---

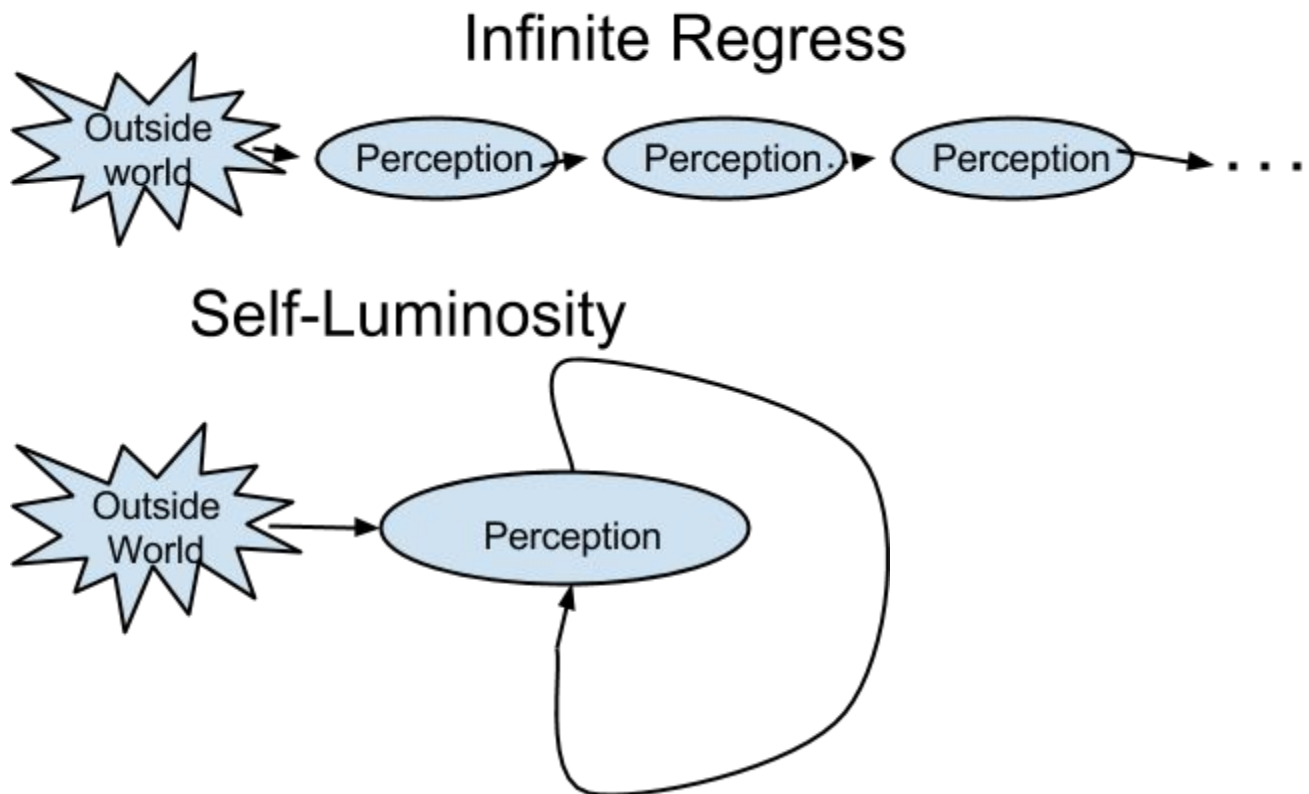
<sup>21</sup> Page 48 of the Gallagher Zahavi reading

thinking that self-ascribed experiences belong only to itself.

- In social psychology, a view is that self-consciousness is the ability to adopt the perspective of another towards oneself.
- In developmental psychology, the mirror test (look in a mirror and see if one can figure out that the image is the self, not another) is often employed.
- Narrative theorists sometimes describe self-consciousness as the ability to develop self-narratives.

The Gallanger Zahavi reading distinguishes between two different forms of self-consciousness: a strong form, which entails the ability to think “I” thoughts, and a weak form, which entails simply having a perspective of being oneself. The strong form can roughly be considered the having of the weak form and the ability to embody it linguistically.

From class and the Caston/Aristotle reading: it is generally acknowledged that consciousness requires a degree of self-consciousness. In other words, to perceive the outside world requires perception that one is perceiving the outside world. It is tempting to say that perception of the outside world requires perception of that perception which requires perception of that perception which requires perception of that perception... ad infinitum. That infinite regress seems rather silly, so Strawson prefers another view: the self-luminosity, or self-intimation, view. In that view, perception perceives the outside world, and is also in charge of illuminating (causing the perception of) itself, in a closed loop.



## Personal Identity

The issue here is coming up with a definition for what makes a person at one point in time the same as a person at another point in time. There are different definitions to consider.

### Bodily Continuity

It may be claimed that personal identity is maintained by bodily continuity - i.e., that a person remains the same as long as their body remains the same over time. There are two major objections to this: one of bodily replacement and one of person/body separation. They can be demonstrated with two different [thought experiments](#): the Ship of Theseus, and the Prince and the Cobbler.

#### Bodily Replacement Objection

This objection lines up with the Ship of Theseus thought experiment. In popular understanding, it is generally thought that every cell in the human body is replaced within seven years. If that is the case, can it truly be claimed that a person has the same body through their entire life?

This objection also ties into the Mars teletransporter thought experiments. If we are to deny that a person teletransported to Mars is not the same as their original on Earth on the grounds that the Mars person does not have the same matter as the Earth person, is that valid if a person on Earth does not even have the same matter through his/her lifespan?

Perhaps the fact that cell replacement on Earth happens gradually, while matter replacement in the teletransporter is more or less instantaneous (and there is never any overlap between new matter and old matter within the same body), is a distinction that can break out of this issue.

Or perhaps we can employ one of the potential resolutions to the Ship of Theseus problem to claim that one does carry the same body through one's life. In other words, we could potentially claim that, when the body sheds one cell and gains another, the old cell ceases to be a part of the body and the new cell begins to be a part of the body.

Finally, we could define the person through the information content (its organization; basically, that which those invested in its definition continually associate the person with).

#### Person/Body Separation Objection

This objection lines up with the thought experiment of the Prince and the Cobbler. If the "soul" of a prince wakes up in the body of a cobbler, and vice-versa, it seems logical that the soul of the prince in the body of the cobbler should be considered the real prince, and vice-versa. This objection contends that the body and the essence of a person are separate entities, that it is logically conceivable for them to be separated, and that it is unwise to link personhood to a body, as personality, memory, experience, and all other traits go with the "soul" instead of with the body. Whether the "soul" and

the body can actually be separated, of course, is a rather difficult matter.

## **Psychological Continuity (Continuity of Memories)**

Note: According to Strawson, this is NOT how Locke defined personal identity. However, many philosophers characterize it as such. Locke's views, in Strawson's eyes, can be found in the next section.

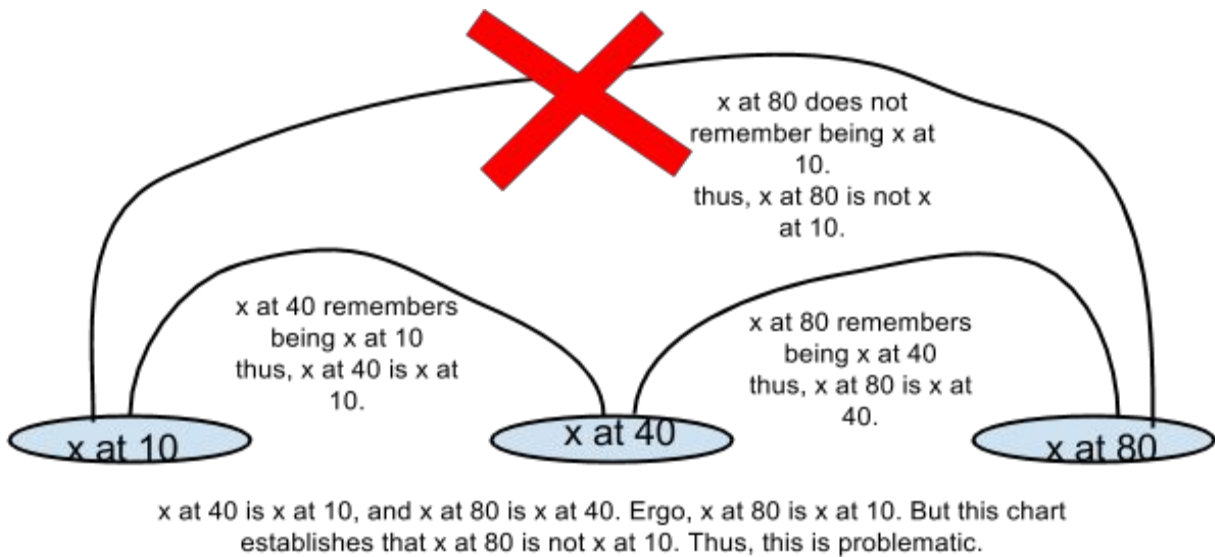
Some contend that personal identity should be defined in terms of continuity of memories. In other words, perhaps a necessary and sufficient condition for personal identity with a person in the past is some element of sharing memories, or remembering being that person. There are two versions of this definition: one requires complete continuity, and the other requires only incomplete continuity.

### **Direct Continuity**

Perhaps the definition of continuity of personhood should rely on continuity of memory. In other words: if you can remember being someone in the past, then you are that someone. This definition causes both intuitive and logical concerns. Amnesia, of course, is an intuitive issue, but that plagues all definitions of identity as continuity of memories, and will be addressed in the 'remaining concerns' section.

Another intuitive concern is this: if an elderly person cannot remember being a young child, should we not consider them the same person as that young child? Logically, that brings a related issue. Let's say that person  $x$  is eighty years old.  $x$  cannot remember being ten years old, but  $x$  can remember being forty years old. Furthermore, when  $x$  was forty years old,  $x$  could remember being ten years old. Under this definition, eighty-year-old  $x$  is the same person as forty-year-old  $x$ , and forty-year-old  $x$  is the same person as ten-year-old  $x$ . Logically, it should follow that eighty-year-old  $x$  is the same person as ten-year-old  $x$ . However, as eighty-year-old  $x$  does not remember being ten-year-old  $x$ , under this definition, they cannot be the same person. This definition struggles with that logical issue.

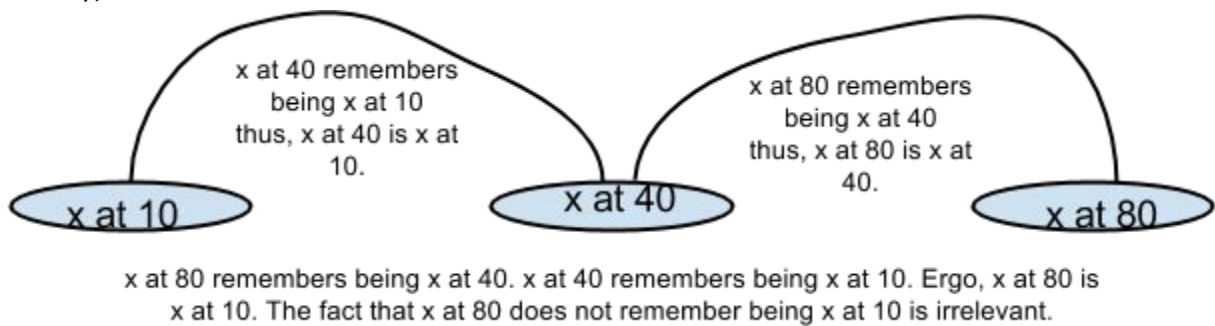




### Indirect Continuity

Indirect continuity avoids the problem of complete continuity described above. In indirect continuity, one shares memory continuity with someone if one can remember being that someone, or if someone whom one can remember being shares memory continuity with that someone. In other words, eighty-year-old x would be the same person as ten-year-old x through forty-year-old x.

It is not a requirement that eighty-year-old x remember being ten-year-old x to be considered the same person as ten-year-old x, as long as eighty-year-old x can remember being someone who remembered being eighty-year-old x (or even remembered being someone who remembered being someone who remembered being ten-year-old x... and as many intermediate steps can be added as is necessary).



### Amnesia Concern

A major remaining concern is that of amnesia. Everyone faces this to some extent, because nobody can remember being a baby. Requiring continuity of memory for personal identity would seem to lead to the conclusion that everyone becomes a different person when they cease to be a baby. This is certainly the case if direct continuity of memory is required for personal identity continuity. It may also be the case if indirect continuity of memory is all that is required for personal identity continuity,

depending on whether or not there are times that the indirect memory condition just cannot reach, no matter how indirectly.

Amnesia presents another major concern. If one is traumatized, faces *radical* amnesia to the point where he/she cannot at all remember anything from before the trauma, then the chain of memory continuity is severed regardless of whether the direct definition or the indirect definition is used. Is it reasonable to say that one becomes a new person upon losing all memory of the past? A belief in defining personal identity with a requirement for continuity of memory would say yes.

## **John Locke's Views on Personal Identity**

Many philosophers characterize John Locke as advocating a definition for personal identity based off of continuity of memories (i.e., psychological continuity). Professor Strawson believes that this is an incorrect reading of Locke. Locke defines *person* as a forensic term, or a term useful in investigation. In other words, Locke (according to Professor Strawson) really defines a theory of action appropriation, instead of one of metaphysical identity.

Note: the rest of this section shall follow Professor Strawson's interpretation John Locke. If it says "Locke says", it means that "Professor Strawson believes that Locke says", and the like.

John Locke advocated a view of personal responsibility based on consciousness of events. If a person is *conscious* of doing something, then they did it, and they deserve to reap the benefits or punishments (to Locke).

How should consciousness of an event be defined? According to Professor Strawson, it's a relationship stronger than memory. To be an event, one must remember it in such an intimate, immediate, and emotional way. One is always conscious of present events. One is not always conscious of past events, however. To be conscious of a past event, one must both remember that event and have some sort of an immediate and intimate connection with it.

### **Issues with Locke's View**

Many people attack Locke's view on personal identity as if it were a metaphysical identity; however, according to Professor Strawson, such criticisms are not valid ones to make. There are other possible criticisms of it, though.

The overarching theme of criticisms of Locke's views is consciousness not always existing in the right places. What if one is conscious of something which they did not actually do, or one is conscious of something which they did do and should be punished/praised for?

Locke answered these criticisms by saying that "the secrets of all hearts shall be laid open" on Judgment Day.

## Locke's Definition of a Person

A person is one with the following qualities:

- Thinking
- Intelligence
- Reason
- Reflection
- The ability to consider oneself as oneself
- Consciousness
- Self-concern
- Capacity to understand happiness, misery, and the law.

This is a functional definition of a person. If a non-human animal were to develop such qualities, that animal may be considered a person. Likewise, some humans may not be considered people. Sociopaths, for instance, would not seem to meet the last criterion. This would require that the question "should sociopaths be considered personally identified with their crimes" be answered with a "no", because one cannot be personally identified with something if one is not a person.

Note: 'Functional,' as stated here, is not identical to the functional of functionalism, though they are related. The definition here contains consciousness and subjective feeling of qualia, through the first part (thinking component) above. Functionalism, according to Professor Strawson, does not make room for such qualia. In both, 'functional' references the definition's focus on characteristics and actions that could theoretically exist on top of any type of creature.

## David Hume's Views on Experience and the Mind

The cornerstone of David Hume's beliefs is that all ideas must come from perceptions to be clear, distinct, and useful. Additionally, Hume says that some philosophers say that we have perfect perceptions of ourselves as nice, individual, indivisible units that perceive.

Putting those pieces together, to Hume, such views of ourselves must come from perceptions. To have a Humean idea of ourselves as continuous perceivers, we must be able to directly perceive some continuity of self every time we introspect. Is this the case? To Hume, the answer is: of course not. Thus, to Hume, we are incapable of having clear and distinct ideas of ourselves as continuous perceivers. We cannot see a constant stream of self when we introspect - all we can perceive is ourselves perceiving at the present moment.

### Most Philosophers' Interpretation

An interpretation taken by many philosophers is that, in making this claim, Hume denies that we can consider ourselves to be subjects of experience. They interpret the claim that we cannot perceive ourselves as continuous perceivers to mean (according to Professor Strawson) that we

cannot perceive ourselves as perceivers. Combined with Hume's tendency to use language that suggests E-S-O moves (see the Hume and E-S-O Moves section), this could lead to an interpretation of Hume suggesting an ontological belief that we are not actually subjects of experience.

Furthermore, many philosophers contend that Hume holds a "bundle theory" of the mind. The argument follows from the Humean concept of all ideas needing to be backed by perceptions. Hume says (according to the argument) that we can perceive of the mind are its behavioral characteristics, and then makes an E-S-O move to say that that is all that the mind *is*. Strawson disagrees with this interpretation.

### Professor Strawson's Interpretation

Professor Strawson defends David Hume by attacking the above interpretation of Hume. Professor Strawson believes that Hume's issue lies with continuity, not perception. In effect, Strawson believes that Hume believes that one can perceive oneself as a perceiver in the present moment, but not as a *continuous* perceiver. In other words, Hume believes that we can perceive ourselves perceiving, but we cannot perceive a steady stream of perception. Rather, we perceive ourselves perceiving in discrete glances.

This interpretation makes Hume compatible with Descartes's Cogito. According to Professor Strawson, Hume never attacks the notion that perception requires a perceiver, which guarantees that humans are indeed perceivers.

The way in which Hume (according to Strawson) outlines humans perceiving themselves works roughly as follows. We must first ask ourselves what we experience when we have experiences. Do we experience the actual content of experience? No, we just experience its properties. Do we experience ourselves *independently*? No. Do we experience the perceptions? Yes. In doing so, are we indirectly aware of ourselves doing so? Yes.

So we experience ourselves in the sense that experience can be defined as the self *having experiences*, and we are aware of ourselves in the process of having experiences.

Strawson also denies that Hume is making any sort of "bundle theory" about the mind. Strawson contends that the E-S-O movement involved there is actually an E-S-OE movement, and that Hume is saying that all we can *know* about the mind comes from its outward behaviors. This also ties into the fact that Strawson claims that Hume believes strongly in the premise that experience needs an experiencer, and by experiencing things we are indirectly conscious of our own minds in a sense extending beyond their outward behavioral characteristics.

### Hume and E-S-O Moves

David Hume has been accused of making ontological claims from E-S-O claims (see the

definitions section). He has a tendency to say things like “the mind is  $x$ .” According to Professor Strawson, however, Hume was enough of a skeptic to not favor the strong and questionable ontological claims that come out of E-S-O moves. Professor Strawson instead believes that his ontological claims are often shorthand for ontological claims with epistemological qualification (for example “the mind is  $x$  are far as we are capable of observing it.”) According to Professor Strawson, Hume was rather excitable, which contributed to his usage of this shorthand.