

IDENTITY & SURVIVAL

the survival
question

conceivability
arguments

are there
immaterial
souls?

the
interaction
argument

Henri Rousseau

Today we turn to our third big question. This question can be introduced by thinking about a simple example.

Suppose that in the year 2070 the surviving members of this Introduction to Philosophy class decided to have an Intro to Philosophy reunion, and all gathered in this room. Suppose that they decided to get a group picture taken.

Now imagine that, via some sort of time travel device, I now have that photo, and show it to you. You might ask: Am I one of those people? Which one am I?

It is very natural to assume that these questions must have determinate answers. There must be some fact of the matter about whether one of the people in the photo is you. And, if one is you, there must be some fact of the matter about which one is you.

Let's suppose that this is true: there must be a fact about whether you survive to be in this picture, and must be a fact about which of the survivors you are.

Let's suppose that this is true: there must be a fact about whether you survive to be in this picture, and must be a fact about which of the survivors you are.

Then we can ask a question about these facts:

The survival question: What does it take for for some person at some other time to be you?

This might seem like kind of a weird question. It also might seem to be a really easy question; you might think that it would just be the person who looks like you, or who has a driver's license with your name on it.

It turns out that this is not such an easy question. One way to see this is by thinking about some harder cases where this question arises.

The survival question: What does it take for for some person at some other time to be you?

It turns out that this is not such an easy question. One way to see this is by thinking about some harder cases where this question arises.

Many people believe in the possibility of life after death. To believe in life after death is to believe that in the afterlife, some time after your death, some person will be you. But what would it take for some person in heaven (say) to be you?

Surely you are not confident that people in heaven will look like people on earth, or carry driver's licenses. So our seemingly easy answers to the survival question don't help us here.

If we want to know whether life after death is possible, it looks like we need a better answer to the survival question.

The survival question: What does it take for for some person at some other time to be you?

Here is another question about the future.

Given advances in computing, it may well be possible in your life time for you to enhance your cognitive powers by replacing parts of your brain with computing devices. It may even be possible for your cognitive apparatus to be, in some sense, uploaded to a computer.

The resulting thing would be, wholly or in part, a synthetic device. Would that thing still be you?

Again, the easy answers don't help. It looks like we need an answer to the survival question.

**The survival
question:** What does it
take for for some
person at some other
time to be you?

A different question concerns the past.

At some time roughly 20 years ago, there was an embryo in some woman's uterus from which you grew. Was that embryo you?

Again, the easy answers are no help. But the question seems to matter; it seems relevant to the question of whether, and when, abortion is morally permissible.

The survival question: What does it take for for some person at some other time to be you?

To introduce our main answers to the survival question, it will be useful to think about a simple, uncontroversial example of survival.

All of you believe that you will wake up tomorrow in your bed. To put the same point another way, all of you believe that the person now sitting in your seat is the same person as — identical to — the person who will wake up in your bed tomorrow morning.

What do we mean when we say that you are identical to that person?

All of you believe that you will wake up tomorrow in your bed. To put the same point another way, all of you believe that the person now sitting in your seat is the same person as — identical to — the person who will wake up in your bed tomorrow morning.

What do we mean when we say that you are identical to that person?

Here it is important to get clear at the outset on one distinction which, if not attended to, can make these questions more confusing than they have to be.

This is the distinction between **numerical** and **qualitative** identity.

To say that x and y are numerically identical is to say that they are literally the same thing — they are one, not two.

To say that x and y things are qualitatively identical is to say that they are exactly resembling — they have just the same properties.

To say that x and y are numerically identical is to say that they are literally the same thing — they are one, not two.

To say that x and y things are qualitatively identical is to say that they are exactly resembling — they have just the same properties.

Here are some examples to help you see the distinction.

Suppose that I have a pair of golf balls that are just the same in every respect — they have the same things printed on them, and they are the same shape and color. They are therefore qualitatively identical. But are they numerically identical? No. They are two, not one.

Now consider a different golf ball. Suppose that tomorrow you paint the golf ball green. Now think about the golf ball today, and the golf ball tomorrow. Are they qualitatively identical? No — one is white, and the other is green. But are they numerically identical? It seems like they are — it is one and the same golf ball that was white today, and is green tomorrow.

When we say that you are identical to the person who will get out of your bed tomorrow morning, we are not of course saying that you are qualitatively identical to that person. Their hair will be messed up, and they will be wearing different clothes.

Suppose that I have a pair of golf balls that are just the same in every respect — they have the same things printed on them, and they are the same shape and color. They are therefore qualitatively identical. But are they numerically identical? No. They are two, not one.

Now consider a different golf ball. Suppose that tomorrow you paint the golf ball green. Now think about the golf ball today, and the golf ball tomorrow. Are they qualitatively identical? No — one is white, and the other is green. But are they numerically identical? It seems like they are — it is one and the same golf ball that was white today, and is green tomorrow.

When we say that you are identical to the person who will get out of your bed tomorrow morning, we are not of course saying that you are qualitatively identical to that person. Their hair will be messed up, and they will be wearing different clothes.

Rather, we mean that you are numerically identical to that person: there is just **one** person who is today in this class, and is tomorrow morning in that bed.

Rather, we mean that you are numerically identical to that person: there is just **one** person who is today in this class, and is tomorrow morning in that bed.


But then we can ask the survival question: in virtue of what are you numerically the same person as the person who will wake up in that bed?

This focus on numerical identity is not just an arbitrary choice. Intuitively, this is the question we care about. When we ask about whether life after death is possible, we are not asking whether after your death someone will exist who has the same properties as you. We are asking whether **you** — this very individual — will exist. And to ask this is to ask whether someone numerically identical to you could then exist.

There are three main answers to that question (though, we will see, they can also be combined in interesting ways).

But then we can ask the survival question: in virtue of what are you numerically the same person as the person who will wake up in that bed?

There are three main answers to that question (though, we will see, they can also be combined in interesting ways).



materialist
survival

Our first answer is also the simplest one. This says that you are the same person as the person who will wake up in that bed tomorrow because you are the same material thing as that person.

On one natural version of this view, you are an organism — a member of the species *homo sapiens*. The explanation of the fact that you are identical to the person waking up in your bed tomorrow is fundamentally the same as the explanation of the fact that my dog is the same dog as the one who will wake up in his crate tomorrow morning.

But then we can ask the survival question: in virtue of what are you numerical the same person as the person who will wake up in that bed?

materialist
survival

But this is not the only answer to our question. It is a widely held view that we are not simply material beings, but also have immaterial souls.

If one believes in immaterial souls, then it is natural to think that one's survival is closely linked to the continued existence of one's immaterial soul.

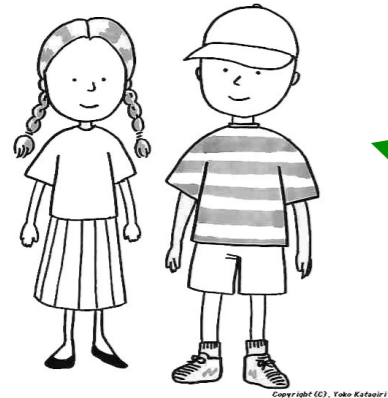
soul
survival

This suggests an alternative theory of survival. On this view, the person who wakes up in that bed tomorrow is you because they are, or have, the same immaterial soul.

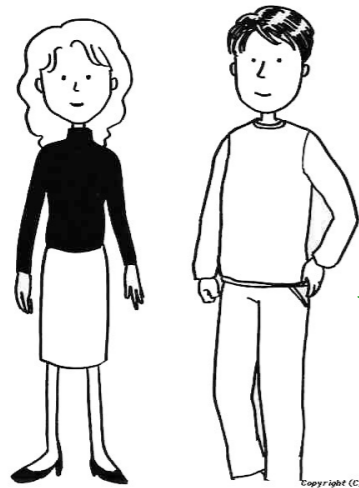
psycho-
logical
survival

But these are not the only possible views here. One might also hold that you are that person not because of any physical connections, and not because of anything involving immaterial souls, but rather because of **psychological connections** between you and that person.

psycho-
logical
survival



psycho-
logical
connection

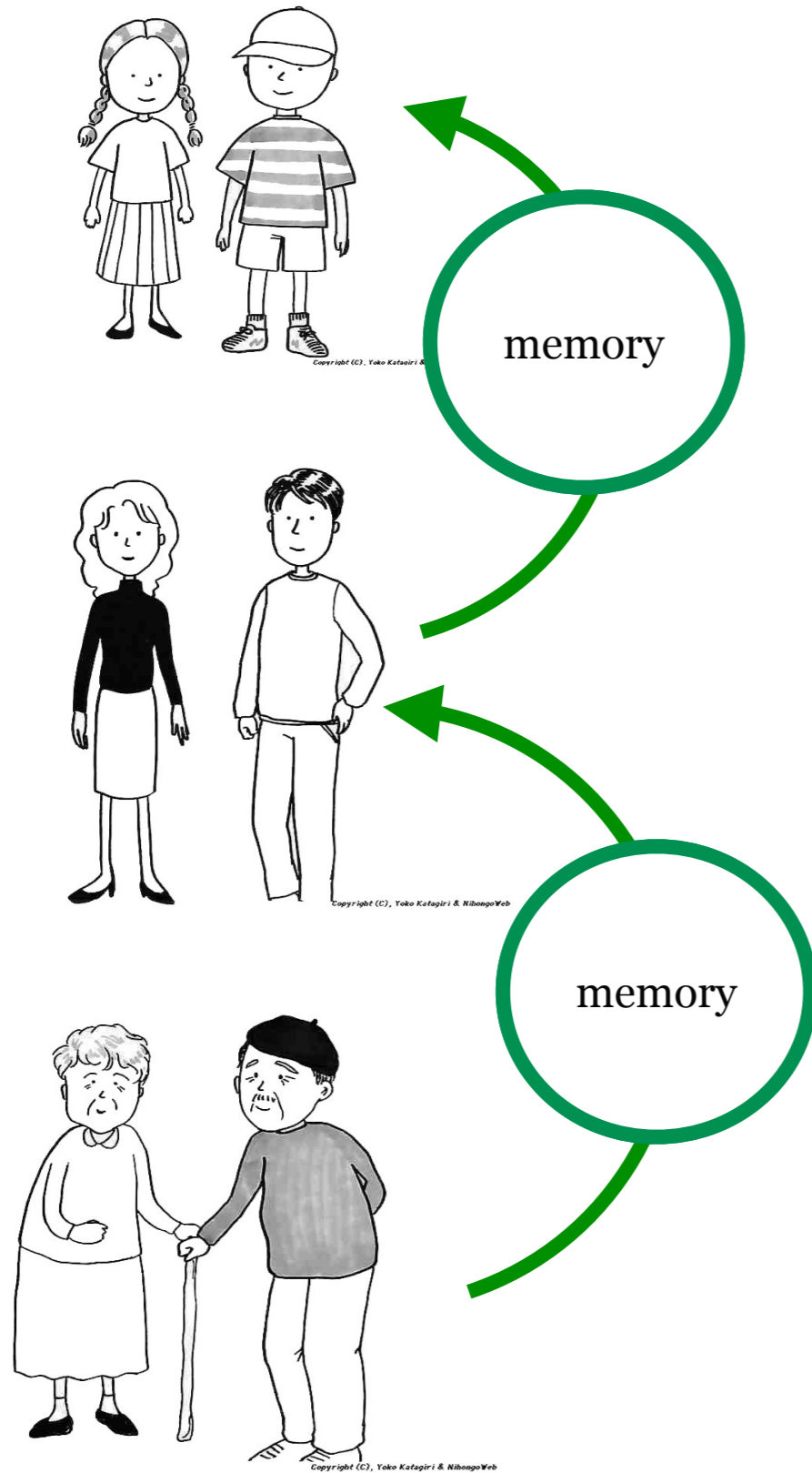


psycho-
logical
connection



This was John Locke's theory. His view of personhood can be illustrated by considering a few different stages in the lives of some people.

What makes the child, the adult, and the elderly person stages of the same person? The materialist says: because they are the same material thing. Locke thought: it is because of **psychological connections** between the individuals.




But what are the relevant psychological relations?

Locke's answer was: relations of **memory**.

Of course there are plenty of other psychological connections between people at one time and those people at a later time. Different versions of the psychological theory of survival focus on different sorts of psychological connections.



soul
survival



psycho-
logical
survival

It is worth pausing to think about the differences between the soul theory and the psychological theory of survival. At a first glance, they can look like the same thing. But they are not.


In ordinary language, 'soul' is sometimes used as a synonym for 'mind.' But remember that in the present context a soul is an immaterial thing which is closely connected to your identity.

What does it mean to be an immaterial thing? A standard view is that material things are by definition things which occupy space. So an immaterial soul would, it seems, be something which exists outside of space. It would not be composed of quarks and electrons, as the things in this room are.

Given that description, it is perhaps not surprising that it is very controversial whether there are such things as immaterial souls.



soul
survival



psycho-
logical
survival

What does it mean to be an immaterial thing? A standard view is that material things are by definition things which occupy space. So an immaterial soul would, it seems, be something which exists outside of space. It would not be composed of quarks and electrons, as the things in this room are.

Given that description, it is perhaps not surprising that it is very controversial whether there are such things as immaterial souls.

Suppose that you do not believe in immaterial souls. You would still of course believe that there are such things as memories, beliefs, and personalities. So you could still endorse the psychological theory of survival.

Similarly, we'll encounter cases in which the soul theory says that an individual survives but in which the psychological theory says they do not.

For now, the main thing to see is that the two theories say very different things about what is required for survival.

materialist
survival

Once we have these three 'pure' theories on the table, it is clear that we can combine them in various ways. For example, one common view is that we are not just a material thing and not just an immaterial soul, but a kind of combination of the two.

soul
survival

Someone with this view might think that, because we are a combination of a body and soul, our continued existence requires the continued existence of both parts of that combination. We might call this 'M+S', since it is the view that survival requires **both** materialist survival and soul survival.

psycho-
logical
survival

M
+
S

Similarly, we could require both material and psychological continuity — M+P — or require both psychological connections and the continued existence of a soul — P + S.

M
+
P

P
+
S

materialist
survival

soul
survival

psycho-
logical
survival

M
+
P

M
+
S

P
+
S

For completeness, we can also consider the maximally demanding view that survival requires all three of material, psychological, and soul continuity:

M + P + S

Our aim for the next two classes will be to try to figure out which of these views is most plausible. We'll then turn to questions about what these views might tell us about the possibility of life after death.

materialist
survival

soul
survival

psycho-
logical
survival

M
+
P

M + P + S

M
+
S

P
+
S

You will be glad to know that we are not going to go through these seven theories one-by-one. Instead, we will look at a series of arguments and thought experiments which seem to make trouble for a number of these theories.

For example, we will look at arguments against the existence of immaterial souls. If successful, such an argument would rule out 4 of our 7 theories.

First, though, I want to look at some arguments which seem to count against the idea that survival requires sameness of material thing. That too would rule out 4 of our 7 theories.

materialist
survival

First, though, I want to look at some arguments which seem to count against the idea that survival requires sameness of material thing. That too would rule out 4 of our 7 theories.

M
+
P

We know that materialist theories require the survival of some material thing. But which one?

M
+
S

As we saw above, it is very natural for the materialist to say that I am an organism. So, it would be natural for the materialist to say that my survival requires the survival of the human organism which I am.

M + P + S

materialist
survival

M
+
P

M
+
S

M + P + S

As we saw above, it is very natural for the materialist to say that I am an organism. So, it would be natural for the materialist to say that my survival requires the survival of the human organism which I am.

Here's an example, due to Derek Parfit, which seems to make trouble for this view:

“Suppose first that I am one of a pair of identical twins, and that both my body and my twin’s brain have been fatally injured. Because of advances in neuro-surgery, it is not inevitable that these injuries will cause us both to die. We have between us one healthy brain and one healthy body. Surgeons can put these together.

If all of my brain continues both to exist and to be the brain of one living person, who is psychologically continuous with me, I continue to exist. This is true whatever happens to the rest of my body. ...”

materialist
survival

M
+
P

M
+
S

M + P + S

In this example — which we will call **Brain Transplant** — your brain survives a car crash undamaged, while your body is destroyed, and your brain is transplanted into the healthy body of a passenger, whose brain was destroyed in the crash.

Parfit's view is that the person who would survive this surgery is you. Is he right about that?

The problem is that the human organism which you were does not survive — only one of its organs does. So we appear to have a case in which I survive even though the organism with which I was associated does not. And that seems to rule out the idea that my survival requires the survival of an organism.

Does that rule out all of the theories at left?

materialist
survival

M
+
P

M
+
S

M + P + S

Does that rule out all of the theories at left?

Not quite. Someone who thinks that survival requires material continuity could simply say that the material thing which must survive is not the organism, but the brain. Call this **the brain view**.

This could be fit into any of the views at left. For example, if you think that survival requires the survival of a material thing and an immaterial soul, you could say that the relevant material thing is the brain.

The brain view is not challenged by the case of Brain Transplant. But it does have some somewhat odd consequences. For example, if you adopt the brain view of survival, it is at least somewhat tempting to adopt the view that I am a brain (rather than an organism). But then it looks like it should be true for me to say things like 'I weigh less than 5 pounds' - after all, my brain does!

But rather than pursue this line of thought, I want to look at what is often taken to be the strongest argument against materialist views of survival: the **conceivability argument**.

This argument is due to René Descartes.

Descartes was one of the most important philosophers who ever lived — a distinction which is especially impressive given that he devoted most of his energies to mathematics (in which he developed what is now analytic geometry) and natural science.

In 1649 Descartes moved to Sweden to join the court of Queen Christina of Sweden. After complaining that “men’s thoughts are frozen here, like the water,” Descartes died in February of 1650, during his first winter in Sweden.





Descartes' argument begins with his thought that all of our beliefs about the existence of material things can be called into doubt:

“Every sensory experience I have ever thought I was having while awake I can also think of myself as sometimes having while asleep. Since I do not believe that what I seem to perceive in sleep comes from things located outside me, I did not see why I should be any more inclined to believe this of what I think I perceive while awake.”

Descartes is saying that we can imagine any sensory experience we have occurring in sleep rather than waking life. But in sleep our seeming sensory experiences do not reflect the reality of the material world around us; so, we can imagine all of the sensory experiences we have failing to reflect the world around us. That is, we can coherently imagine a scenario in which there are no tissue boxes, cats, planets, or other material things, even though in our experience it seems to us that there are such things.

Now let's ask another question: when we conceive of the possibility that there are no material things, are we conceiving a situation in which **nothing at all** exists?

Now let's ask another question: when we conceive of the possibility that there are no material things, are we conceiving a situation in which **nothing at all** exists?

“I have convinced myself that there is absolutely nothing in the world, no sky, no earth, no minds, no bodies. Does it now follow that I too do not exist?

No: if I convinced myself of something then I certainly existed. ... This proposition - *I am, I exist* - is necessarily true whenever it is put forward by me or conceived in my mind.”

Descartes here seems to be saying that, when I imagine a world in which there are no material things, I am still imagining that I exist. This suggests the following claim:

I can clearly imagine a scenario in which I exist, but no material things exist.

Suppose that this claim about imagination is true. Who cares? We are interested in what it could take for us to survive, not about what we can imagine.

The answer to this question comes in the following passage:

“I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. Hence the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct.”

“I know that everything which I clearly and distinctly understand is capable of being created by God so as to correspond exactly with my understanding of it. Hence the fact that I can clearly and distinctly understand one thing apart from another is enough to make me certain that the two things are distinct.”

Descartes seems to be saying that if I can clearly imagine something to be the case, then God could make it the case: God could bring it about. It seems to follow from this that Descartes would endorse the following principle:

If I can clearly imagine something being the case, then it is possible for it to be the case.

Is there any reason to think that this is true?

I can clearly imagine a scenario in which I exist, but no material things exist.

If I can clearly imagine something being the case, then it is possible for it to be the case.

It is possible that I exist and no material things exist.

If it is possible that I exist and no material things exist, then materialist theories of survival are false.

Materialist theories of survival are false.

THE CONCEIVABILITY ARGUMENT

1. I can clearly imagine a scenario in which I exist, but no material things exist.
2. If I can clearly imagine something being the case, then it is possible for it to be the case.
3. It is possible that I exist and no material things exist. (1,2)
4. If it is possible that I exist and no material things exist, then materialist theories of survival are false.

C. Materialist theories of survival are false. (3,4)

You might doubt that premise (1) of this argument is true. Can we really clearly imagine a scenario in which we exist but no material things do? What, exactly, are we imagining?

If one has doubts about this premise, it may help to consider a variant of the argument, which is due to a thought experiment from John Locke.

You might doubt that premise (1) of this argument is true. Can we really clearly imagine a scenario in which we exist but no material things do? What, exactly, are we imagining?

If one has doubts about this premise, it may help to consider a variant of the argument, which is due to a thought experiment from John Locke.

“Should the soul of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler ... everyone sees, he would be the same person with the prince, accountable only for the prince’s actions.”

Locke is imagining a scenario now familiar from various movies and TV shows: the idea that we can imagine one person waking up one morning in someone else’s body.

Suppose that we can clearly imagine this. It is easy enough to use this fact to come up with a variant on the conceivability argument, which we can call the ‘body-swapping argument.’

“Should the soul of a prince, carrying with it the consciousness of the prince’s past life, enter and inform the body of a cobbler ... everyone sees, he would be the same person with the prince, accountable only for the prince’s actions.”

Locke is imagining a scenario now familiar from various movies and TV shows: the idea that we can imagine one person waking up one morning in someone else’s body.

Suppose that we can clearly imagine this. It is easy enough to use this fact to come up with a variant on the conceivability argument, which we can call the ‘body-swapping argument.’

THE BODY-SWAPPING ARGUMENT

1. I can clearly imagine a scenario in which I exist in a different body.
2. If I can clearly imagine something being the case, then it is possible for it to be the case.
3. It is possible that I exist in a different body. (1,2)
4. If it is possible that I exist in a different body, then materialist theories of survival are false.

C. Materialist theories of survival are false. (3,4)

THE BODY-SWAPPING ARGUMENT

1. I can clearly imagine a scenario in which I exist in a different body.
2. If I can clearly imagine something being the case, then it is possible for it to be the case.
3. It is possible that I exist in a different body. (1,2)
4. If it is possible that I exist in a different body, then materialist theories of survival are false.

C. Materialist theories of survival are false. (3,4)

One might think that this is superior to the original conceivability argument, on the grounds that we can more clearly imagine the 'body-swapping' scenario than the scenario in which I exist but no material things exist.

How should the proponent of any one of our four materialist theories of survival respond?

Let's consider an objection to the second premise of both arguments:

2. If I can clearly imagine something being the case, then it is possible for it to be the case.

Here is a possible counterexample to this premise:

The barber

Deep in the Indiana countryside, there's a small town; and in this town there's a barber. Some of the men in this small town - the industrious ones - shave themselves every morning. But others (the lazy ones) don't; and the barber shaves all of them. (There's no one else around who will do it.) Moreover, he (the barber's a man) never shaves any of the industrious ones - he never shaves any of the men that shave themselves.

The barber

Deep in the Indiana countryside, there's a small town; and in this town there's a barber. Some of the men in this small town - the industrious ones - shave themselves every morning. But others (the lazy ones) don't; and the barber shaves all of them. (There's no one else around who will do it.) Moreover, he (the barber's a man) never shaves any of the industrious ones - he never shaves any of the men that shave themselves.

Can you clearly imagine this small town? Is it possible for there to be a town of this sort?

The barber

Deep in the Indiana countryside, there's a small town; and in this town there's a barber. Some of the men in this small town - the industrious ones - shave themselves every morning. But others (the lazy ones) don't; and the barber shaves all of them. (There's no one else around who will do it.) Moreover, he (the barber's a man) never shaves any of the industrious ones - he never shaves any of the men that shave themselves.

Does the town's barber shave himself, or not?

No.

But then he does shave himself, because he shaves every man that does not shave himself.

Yes.

But then he doesn't, because he doesn't shave any man that shaves himself.

So if he shaves himself, then he doesn't, and if he doesn't, he does.

This is thus an example of a scenario which, at first glance, seems possible, but then turns out, on closer inspection to be impossible, because it contains a hidden contradiction. Might the materialist plausibly say the same thing about the scenarios which figure in our two arguments?

This is an appealing thought, if you are a materialist. But you should ask yourself: what contradiction could this be? What could be impossible about these scenarios which we seem to be able to imagine?

Suppose, for the sake of argument, that you're convinced by the conceivability argument or the body-swapping argument. (We'll return to materialist views later.) That would leave three of our theories of survival on the table.



How can we decide which of these these theories is most likely to be true?

Two of these theories are committed to the existence of immaterial souls. The pure psychological theorist is likely to see this as a significant strength of their theory.

Most of us don't believe in fairies, there being no good evidence for their existence. Why should souls be any different?

Most of us don't believe in fairies, there being no good evidence for their existence. Why should souls be any different?

We can turn this into an argument against believing in immaterial souls. Consider the following points:

We have no sensory experience of immaterial souls.

It is not self-evident that there are immaterial souls.

We have no good argument for the existence of immaterial souls.

But the following principle looks plausible:

If P is not self-evident and your senses don't tell you that P and you don't have a good argument for P, you should not believe P.

We have no sensory experience of immaterial souls.

It is not self-evident that there are immaterial souls.

We have no good argument for the existence of immaterial souls.

If P is not self-evident and your senses don't tell you that P and you don't have a good argument for P, you should not believe P.

You should not believe in the existence of immaterial souls.

THE EVIDENTIALIST ARGUMENT AGAINST BELIEF IN IMMATERIAL SOULS

1. We have no sensory experience of immaterial souls.
2. It is not self-evident that there are immaterial souls.
3. We have no good argument for the existence of immaterial souls.
4. If P is not self-evident and your senses don't tell you that P and you don't have a good argument for P, you should not believe P.

You should not believe in the
existence of immaterial souls.
(1,2,3,4)

This might be called the 'evidentialist' argument against belief in immaterial souls, since it is based on the idea that we have no **evidence** for the existence of immaterial souls.

When we turn to the question of how we should determine what to believe, we'll ask whether 'rules of belief' like premise (4) are true. But the premise looks at least initially plausible; so it puts some pressure on the believer in immaterial souls to respond to the argument by rejecting premise (3).

But then we need an argument for the existence of immaterial souls.

The argument we are going to look at turns on the idea that immaterial souls are needed to explain the nature of **consciousness**.

Properties related to conscious experience include the property of feeling an itch or a pain, or the sensation of seeing red or hearing a loud noise. These properties are sometimes called **phenomenal properties**.

Human beings (obviously) have phenomenal properties — we experience all kinds of sensations.

Suppose that we were wholly physical things (as both the organism view and the brain view say). What would that imply about the nature of our phenomenal properties?

The following principle seems plausible:

Wholly physical things have only physical properties.

Suppose that we were wholly physical things (as both the organism view and the brain view say). What would that imply about the nature of our phenomenal properties?

The following principle seems plausible:

Wholly physical things have only physical properties.

But then it seems to follow from the fact that we have phenomenal properties that

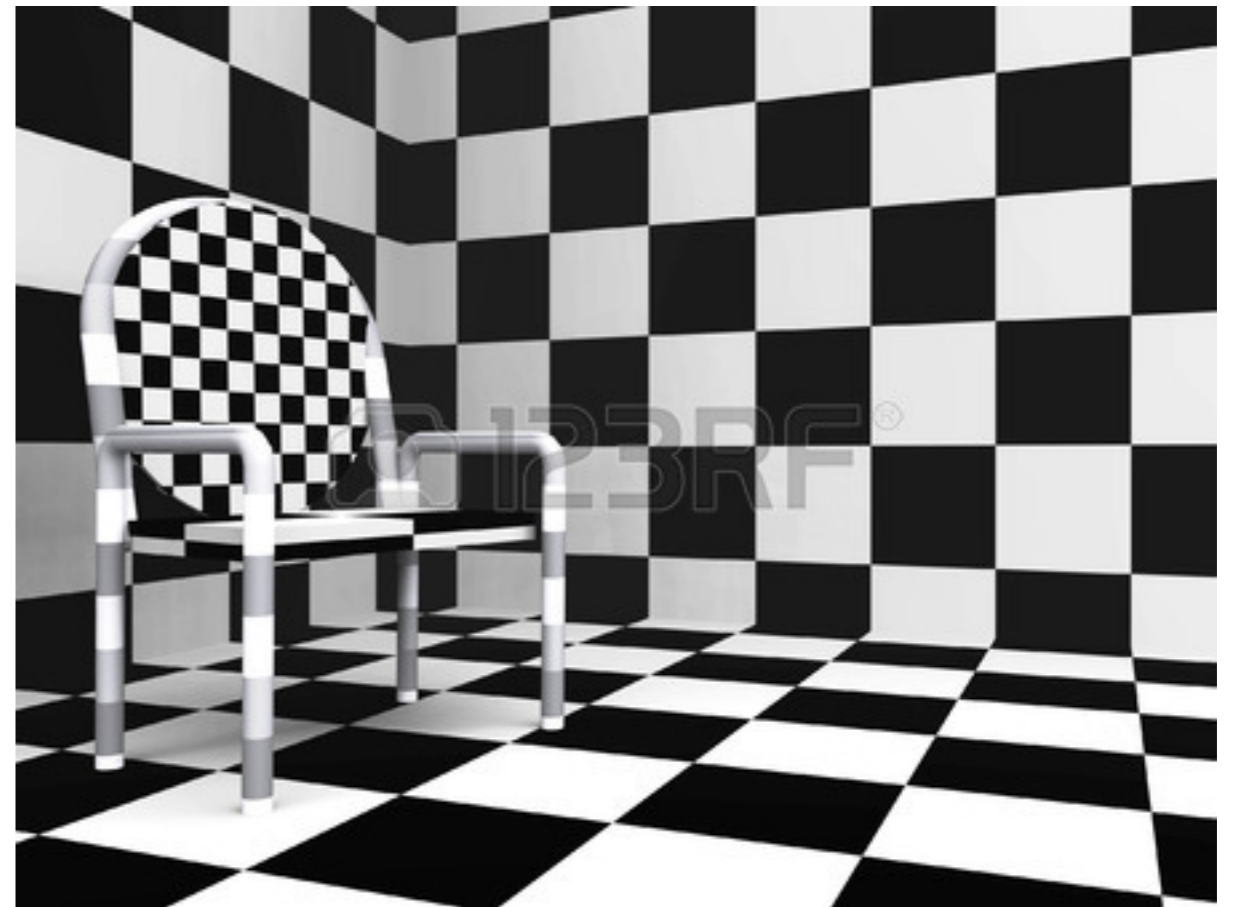
If we are wholly physical things, then phenomenal properties are physical properties.

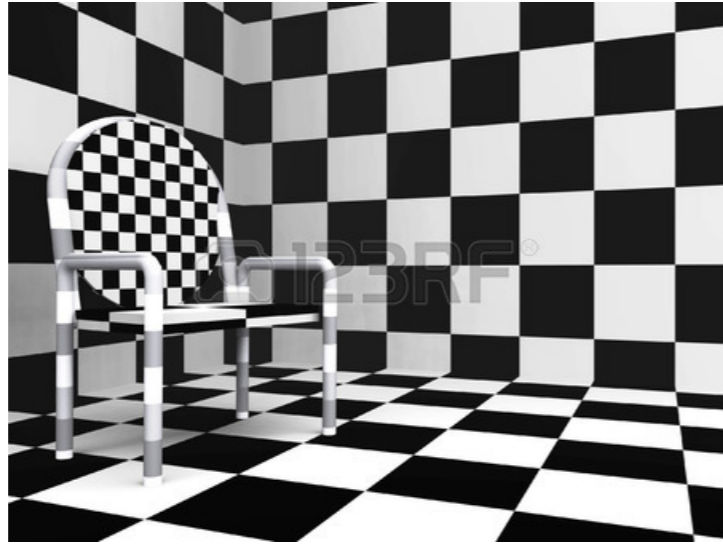
We'll now look at an argument that phenomenal properties are **not** physical properties — which, given the above, would show that we are not wholly physical things.

We'll now look at an argument that phenomenal properties are **not** physical properties — which, given the above, would show that we are not wholly physical things.

This argument is due to the philosopher Frank Jackson, and is based on his example of Mary and the black-and-white room.

Mary is a brilliant scientist who has been confined her entire life to an environment in which everything is colored white or black.



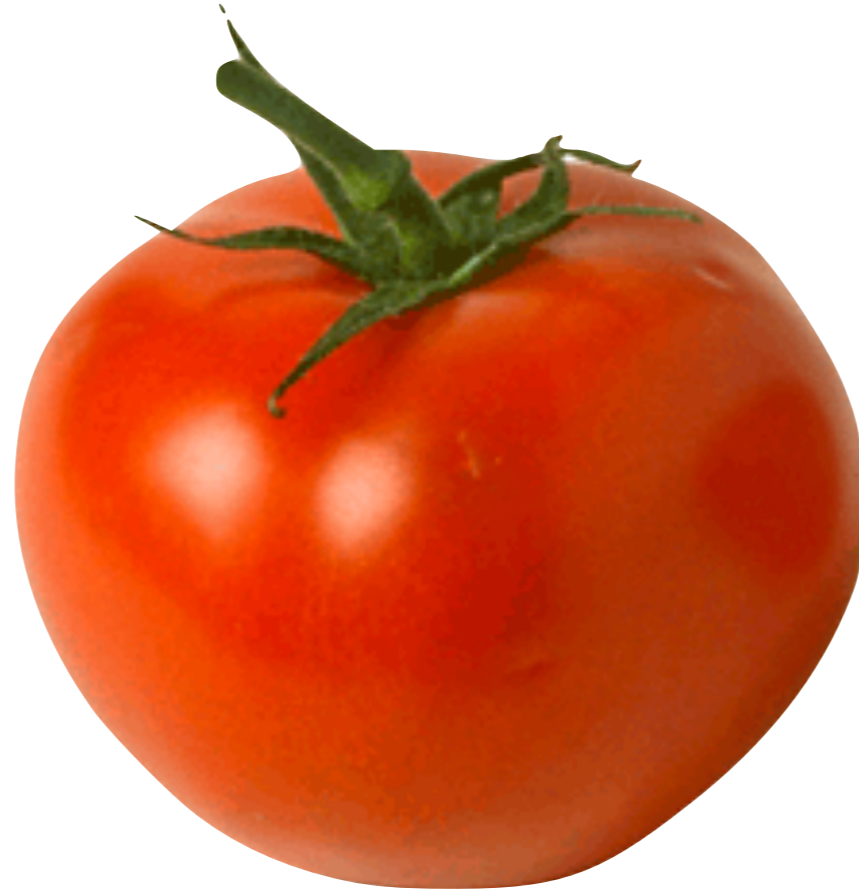


Mary is a brilliant scientist who has been confined her entire life to an environment in which everything is colored white or black.

She is so brilliant, in fact, that she has learned every fact that there is to learn about the physical world. In particular, she has learned all of the facts about the neurophysiology of color vision, and has studied extensively everything that happens to the brain when subjects are experiencing color.

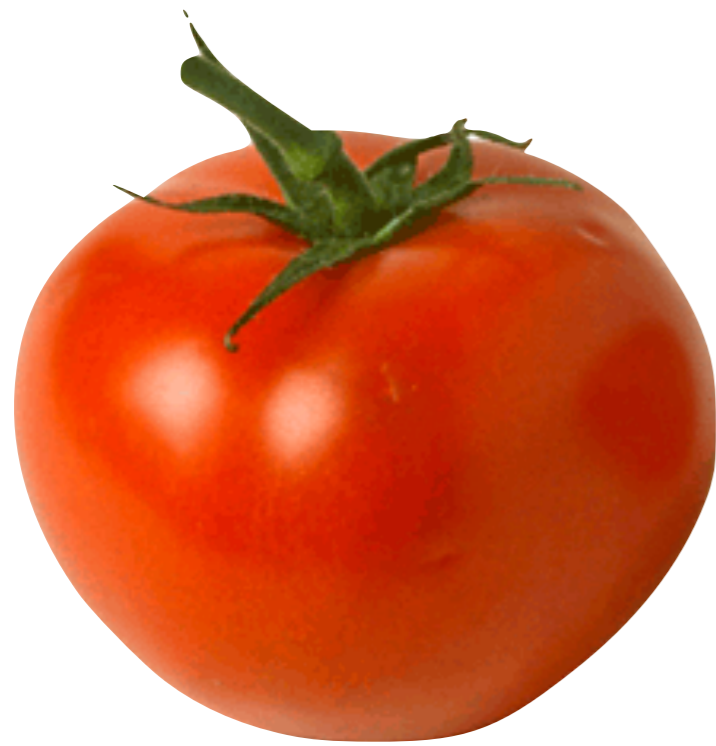
So she knows all of the physical properties that brains have when the person whose brain it is is experiencing color. It seems like this should be possible; people who are color blind can still learn physics.

One day, someone brings a new object into Mary's room:



Does Mary learn anything new when she sees the tomato?

Intuitively, yes: she learns **what it is like** to experience red things. It seems that she learns something about the phenomenal property of sensing redness.



But this seems to show that phenomenal properties are not physical properties.

Here is one way in which the argument, which is sometimes called the **knowledge argument**, can be laid out.

Upon seeing the tomato, Mary learns a new fact about conscious experiences.

Before seeing the tomato, Mary knows all the physical facts about conscious experiences.

There are non-physical facts about conscious experience.

Upon seeing the tomato, Mary learns a new fact about conscious experiences.

Before seeing the tomato, Mary knows all the physical facts about conscious experiences.

There are non-physical facts about conscious experience.

If we are wholly physical things, then phenomenal properties are physical properties.

Phenomenal properties are not physical properties.

We are not wholly physical things.

If we are not wholly physical things then immaterial souls exist.

Immaterial souls exist.

THE KNOWLEDGE ARGUMENT

1. Before seeing the tomato, Mary knows all the physical facts about conscious experiences.
2. Upon seeing the tomato, Mary learns a new fact about conscious experiences.
3. There are non-physical facts about conscious experience. (1,2)
4. Phenomenal properties are not physical properties. (3)
5. If we are wholly physical things, then phenomenal properties are physical properties.
6. We are not wholly physical things. (4,5)
7. If we are not wholly physical things then immaterial souls exist.

C. Immaterial souls exist. (6,7)

If we are not wholly physical things, then presumably we are (at least in part) immaterial souls. So, if the knowledge argument is sound, there are immaterial souls. And if there are immaterial souls, it seems plausible that they would play a role in answering the survival question.

THE KNOWLEDGE ARGUMENT

1. Before seeing the tomato, Mary knows all the physical facts about conscious experiences.
2. Upon seeing the tomato, Mary learns a new fact about conscious experiences.
3. There are non-physical facts about conscious experience. (1,2)
4. Phenomenal properties are not physical properties. (3)
5. If we are wholly physical things, then phenomenal properties are physical properties.
6. We are not wholly physical things. (4,5)
7. If we are not wholly physical things then immaterial souls exist.

C. Immaterial souls exist. (6,7)

One of the most popular responses to the knowledge argument from materialists involves denying that (2) is true.

Here's one way to make this seem plausible. Consider Lois Lane. She knows that Superman can fly; and she knows who Clark Kent is. But she does not know that Clark Kent is Superman.

Here's one way to make this seem plausible. Consider Lois Lane. She knows that Superman can fly; and she knows who Clark Kent is. But she does not know that Clark Kent is Superman.

Suppose that one day she sees Clark don the Superman costume and fly away. She is surprised — just like Mary is surprised.

If you asked her why she is surprised, she might say, "I did not know that Clark could fly!"

But of course in a way she did. She knew that Superman could fly. And Clark = Superman. So isn't the fact that Superman can fly just the same as the fact that Clark can fly?

It looks like Lois is surprised, not because there is some new fact that she learns, but because (in some sense) she learns a new way of thinking about a fact she already knew.

Could the materialist say that, similarly, Mary does not learn a new fact, but instead learns a new way of thinking about a physical fact she already knew?

Let's turn now to the case **against** immaterial souls. The historically most influential argument against dualism is one originally raised by Princess Elisabeth of Bohemia.

Elisabeth was one of the leading intellectuals of the 17th century. She worked in mathematics and physics as well as philosophy, and was active in German politics. She was known by her siblings as 'The Greek' because she mastered ancient Greek at such a young age.



Today Elisabeth is best known for her correspondence with Descartes, in which she raised the question of how an immaterial soul and a material body could interact.

Today Elisabeth is best known for her correspondence with Descartes, in which she raised the question of how an immaterial soul and a material body could interact.

Elisabeth pointed out that, if dualism is true, then it must happen quite often that the body causes effects in the soul, and that the soul causes effects in the body.

What might be some examples of your body causing effects in your soul?

How about examples of your soul causing effects in your body?

Consideration of examples show that, if dualism is true, then interactions between soul and body must happen all of the time. But Elisabeth argued that these kinds of causal interactions were entirely mysterious.



Consideration of examples show that, if dualism is true, then interactions between soul and body must happen all of the time. But Elisabeth argued that these kinds of causal interactions were entirely mysterious.

If you think about it, the idea that an immaterial soul is constantly interacting with your physical body is kind of weird. Could it really be true that every sensation you feel and every action you undertake involves an interaction between your body and some non-spatial immaterial thing?

If you think not, that would appear to provide a simple argument against the existence of immaterial souls.



If you think not, that would appear to provide a simple argument against the existence of immaterial souls.

THE INTERACTION ARGUMENT

1. If there are immaterial souls, they causally interact with material bodies.
 2. Immaterial things can't causally interact with material things.
-
- C. There are no immaterial souls.

Attention naturally focuses on premise (2). Sure, interaction between physical and non-physical things seems a little weird; but is there any way for the defender of the interaction argument to show that it never happens?



Elisabeth gave one defense of this premise:

“it seems every determination of movement happens from the impulsion of a thing moved, according to the manner in which it is pushed by that which moves it ... Contact is required for [this]. ... contact seems to me incompatible with the idea of an immaterial thing.”



What's her argument here?

Elisabeth here is considering a special case of causation: putting something in motion. For something to push something else, it seems that the two things must be in contact; and for two things to be in contact, both must occupy space (since being in contact is just a matter of occupying adjacent spaces). Since immaterial minds don't occupy space, it seems that they can't set things in motion — so, for example, my mind's desire for coffee can't be what sends my body down the hallway in search of some.

Elisabeth here is considering a special case of causation: putting something in motion. For something to push something else, it seems that the two things must be in contact; and for two things to be in contact, both must occupy space (since being in contact is just a matter of occupying adjacent spaces). Since immaterial minds don't occupy space, it seems that they can't set things in motion — so, for example, my mind's desire for coffee can't be what sends my body down the hallway in search of some.

Descartes replied that while some causation does work through contact, not all does. He gives the example of gravity. The earth interacts with the moon via its gravitational force. But this interaction does not require contact — so why, in general, should we think that contact is required for causal interaction?

Elisabeth was skeptical about the idea that the example of gravity could provide a model for the interaction between mind and body. And many philosophers since have been on Elisabeth's side. Let's consider one way of developing her argument further.

Elisabeth was skeptical about the idea that the example of gravity could provide a model for the interaction between mind and body. And many philosophers since have been on Elisabeth's side. Let's consider one way of developing her argument further.

If there are such things as souls, and those souls have phenomenal properties, then there must be some laws of nature which connect what happens in brains with the phenomenal properties of souls. Let's call these laws of nature **psychophysical laws**.

It seems that, if you believe in immaterial souls, you have to believe in the existence of psychophysical laws.

The problem is that it seems that these laws would be quite different from any other fundamental laws of nature with which we are acquainted.

The problem is that it seems that these laws would be quite different from any other fundamental laws of nature with which we are acquainted.

Laws of nature typically relate fundamental physical properties to other fundamental physical properties. Consider, for example, Newton's equation

$$F=ma$$

This law relates force, mass, and acceleration — all quite fundamental physical properties. But what would our psychophysical laws look like?

Consider the psychophysical law governing having a reddish sensation. It will presumably be of the form

RED. If x has physical property P , then x 's soul has a reddish sensation

Consider the psychophysical law governing having a reddish sensation. It will presumably be of the form

RED. If x has physical property P , then x 's soul has a reddish sensation

But what is physical property P ? It is not going to be any remotely natural physical property. It is going to be a massively complex physical property involving the more than 100 million neurons in the human visual cortex.

This already makes "laws" like RED way different than the kinds of laws we typically seek in physics.

The above statement does not look like a fundamental law. It looks more like a generalization which needs to be explained by more fundamental laws. In this sense it looks more like the equations which predict the movement of the tides. Precisely because those equations are about massively complex physical systems, we think that they must be explained by more basic laws.

The above statement does not look like a fundamental law. It looks more like a generalization which needs to be explained by more fundamental laws. In this sense it looks more like the equations which predict the movement of the tides. Precisely because those equations are about massively complex physical systems, we think that they must be explained by more basic laws.

The problem is that it is hard to see how the dualist can try to explain RED in more basic terms.

And note that the dualist won't have to just accept one or two weird laws of nature like RED. Humans can discriminate about 10 million different colors; so it looks like we will need 10 million different fundamental laws of nature, one corresponding to each type of sensation. And that's just the beginning. We will also need fundamental laws for all of tastes, smells, itches, pains, and other sensations we are capable of feeling.

And note that the dualist won't have to just accept one or two weird laws of nature like RED. Humans can discriminate about 10 million different colors; so it looks like we will need 10 million different fundamental laws of nature, one corresponding to each type of sensation. And that's just the beginning. We will also need fundamental laws for all of tastes, smells, itches, pains, and other sensations we are capable of feeling.

By contrast, the non-dualist can say that phenomenal properties just **are** complex physical properties. There's thus no need for fundamental psychophysical laws.

This seems to show that the dualist is committed to a much more complicated view of the laws of nature than the materialist.

But we regularly think that simpler theories are to be preferred over less-simple theories. So this seems to be a strike against the dualist.