

## A BRIEF INTRODUCTION TO THE PHILOSOPHY OF MIND

PROF. DAN FLORES  
HCC – NORTHWEST

### PART ONE: DESCARTES' *MEDITATIONS*: MEDITATION I

Let us first sketch out Descartes' *Argument from Skepticism* from Meditation I and then work our way through it.

Premise 1. I have believed many false things

Premise 2. My senses lie to me

Premise 3. It is possible that I am insane

Premise 4. It is possible that I am dreaming

Premise 5. It is possible that a maximally powerful "evil genius" is deceiving me

---

Conclusion 6. Everything that I claimed as true (knowledge) is illusory.

PREMISE 1. Descartes begins the meditations with the proposition that he has believed many false things since he was a child. Surely this is true. We have all believed many things to be true as children that we now regard as false, e.g., Santa Clause, ghosts, the monster under the bed, the Tooth Fairy, that cats have nine lives, etc... Further, isn't it at least possible that we, as adults, believe false things? You might believe your friend to be trustworthy, or that your grandmother loves you; but is this necessarily true? Hardly! The point, here, is that Descartes is trying to get at what counts as genuine knowledge. He is concerned with the state of knowledge and science – what can be claimed to be known beyond any shadow of doubt. Once he realizes that he has – and probably still does – believed in false propositions, then he immediately realizes that the very foundation of his knowledge (and ours) is completely in question. Doubt (skepticism) seems to undermine our knowledge.

PREMISE 2. If it wasn't bad enough that we believe false propositions, it is also certain that our senses deceive us. We have to learn as children that the water we see in front of us while on the road is simply an illusion and that it only appears that there is water. We learn that the wheels on a car are not really spinning in the opposite direction as they appear to do when on the road. We know that under certain conditions the lights can play tricks on us, and that alien abduction is symptomatic of sleep paralysis (and you should seek medical help!). We see things, hear things, and feel things all of the time that are not there. Descartes has us suppose that we

take this fact about our lives and consider the possibility that my hands, body, the clothes that we have on now, the place we're sitting are nothing but the product of our senses deceiving us. I might only think this is my hand in front of me when there is really nothing just as if I only think that there is water in front of me when it is nothing but a mirage. After all, is it not possible that my senses are, in fact, deceiving me about this being my own body? Of course it is. But Descartes tries to block this line of reasoning by suggesting that our senses are deceived, but that they are deceived in matters where objects are really far away or very difficult to perceive. Is there anything in my regular experience that suggests that I should doubt that I have hands? Or that this computer that I am typing on is illusory? Descartes' answer is, "No."

PREMISE 3. "No," that is, unless I were insane. If I were insane strapped up in a gurney somewhere, foaming at the mouth in some fit of delirium, then it is entirely possible that this is all illusory. And isn't that actually a possible state of affairs? What, after all, do the insane say is real? They think they are kings, when they are paupers, or that their bodies are made out of glass. They think that they mysterious people are listening to them and that their bodies don't really exist. How exactly would you know that you weren't insane? For all you know, you could be locked in a mental institution and your even reading this sentence is nothing but the mad ravings of a lunatic who is confined in a chair slobbering over herself! The key here is that, regardless of how likely or unlikely it is, it is a distinct *possibility*! Do you, or does anybody, have 100% proof that your report of the world or that of anybody else's is accurate? Very doubtful. Think about how much you would have to know about the world to say that it is certain that you are not actually insane – are insane people not absolutely convinced that what they claim is real? Wouldn't you do the same if you were insane?

PREMISE 4. However, there is one thing that the insane seem to share with the sane – we all sleep and dream. The lunatic and the sane person both have fantastic dreams. The sane will dream that they are made of glass, or that they can fly, or walk on water. They dream of monsters, and witches, and magic, and dragons, and that they are persecuted and being followed when they are not. These are the *same things* that an insane person imagines while awake. So, the common denominator between the sane and the insane is sleep and dreaming. This raises the question: if the sane and the insane give the same reports and both sleep and dream, then what are the odds that I am actually insane as opposed to merely dreaming? Descartes seems to think that it is more likely that you are dreaming. So, even if it turns out that you are not insane, you still might be dreaming such that the hands you use to grab the cup of coffee in front of you, the couch you're sitting on, the car you drive, and the philosophy course that you are sitting in are all fantastic images of your dreaming life. Further, is there a 100% way

of testing whether or not you are dreaming or are awake? Don't we all have memories of events that we are not sure whether or not it was a dream or an actual event? Did I say that, or was it just in my dream? Did I go to Disneyland as a child, or was it a pure fantasy that I did? Without a guaranteed way of telling the difference, then it is possible that you are, in fact, dreaming right now (and you'd have to be dreaming or insane to be taking part of this course, right?).

Nevertheless, says Descartes, even though this all might be a dream the fact remains that we dream of familiar objects. Consider our mythological stories and the fantastic creations of art; even the wildest and most outlandish of creatures in our stories and depictions of oddities in our paintings and sculptures are nothing more than the recombination of familiar objects. The unicorn is a horse with a horn; the Pegasus is a horse with wings; the chimera is a combination of a lion, goat, and bird; the Griffin is naught but a mix of a lion and an eagle; the Predator is bipedal, has two arms, and eyes, a nose, and mouth in the same places that we do; likewise for ghosts; Brahma, Shiva, and Vishnu assume human forms; even Jabba the Hutt has a familiar look. This all suggests that our dreams themselves are parasitic on our waking life. And even if we could somehow imagine something completely novel, the colors that one would dream in would still be dependent on what we experience in our waking lives. So, even if you are dreaming, there is small consolation that, at least, you once had to be awake and that, like all who sleep, you will wake yet again.

PREMISE 5. But Descartes is not done yet. In the midst of overcoming the most credulous skepticism, he presents the strongest of his premises. Imagine a maximally powerful being that is deceiving us about absolutely everything – not God, since God is benevolent and deception would run contrary to God's nature of benevolence (even though God could do absolutely anything on account of being *omnipotent*). So, not God, but an evil genius, Descartes has as being the source of all of our deception. We believe false things, our senses deceive us, we might be insane, or we might be dreaming. All are very real possibilities, unlikely as they may be. Nevertheless, all of these premises have objections that suggest that they are not likely. How, though, does one get past the possibility that some supernatural being is deceiving you – and not just any being, but a being powerful enough to manipulate the very appearance of the world; a being powerful enough to trick you into believing that that being doesn't exist or that  $1 + 1 = 9.874887$ . Isn't it, after all, a possibility? That's tough.

CONCLUSION. So, then, where are we left? We are left with the proposition that all of our knowledge is most likely false. This is an important point for Descartes' method. If you recall from the beginning of this meditation Descartes wants to find some piece of knowledge that is indubitable (not able to be doubted). The reason he wants this is to be able to do science and

have our knowledge rest on solid foundations. If it's possible that all we know is false and there is no 100% way of knowing which is true and which is false, then how can we inquire into the world and trust that what we claim to know is, in fact, true? We can't. Descartes' method is what is referred to as **hyperbolic skepticism**, or, exaggerated skepticism – doubting absolutely everything. And what makes matters worse is that one can (legitimately!) doubt everything. The end result is that our state of knowledge is, at best, on an unstable foundation. At worst, everything we know to be true is false! Now what?

## **PART II: MEDITATION II – PROVING THE EXISTENCE OF MATERIAL AND NONMATERIAL SUBSTANCE**

After having argued that all we know is most likely false, Descartes asks a simple but entirely fundamental question: *If I am fooled about absolutely everything, even the fact that my body is not real nor is the knowledge that I claim, then why is it that I am still not convinced that I don't exist?* His answer is that even if everything is an illusion, even if everything I hold to be true is false and I am deceived about absolutely everything by an evil genius, then it still remains that there has to be something there to be deceived. And since I cannot assume even that I have a body (remember, that can be legitimately doubted), then what must exist is a mind. Descartes asks what am I if not a body? Am I not a thinking thing? A thing that perceives, that thinks, questions, wonders, believes, hopes, loves, wishes, and so on? And are not all of these qualities and more qualities of minds? Therefore, he says, each time that I utter or even think the words "I think, therefore I am" (*cogito ergo sum*)<sup>1</sup> it must be a necessarily true statement, one that even God couldn't deceive you about simply because something has to exist in order to be deceived. And since minds don't need bodies, an obvious conclusion from Descartes' skeptical argument, this means that mind is nonmaterial stuff – a real thing that does not take up space, literally *res cogitans* (a thinking thing) whose primary characteristics are that it is nonmaterial substance that does not take up space and is not subject to the laws of nature and that perceives, wills, judges, affirms, denies, and so on.

But Descartes goes on to wonder about another curious fact. He has just proved that it is a necessary fact that *I think; I exist*. He then asks us to consider a ball of wax. As is, it has many characteristics that our senses can perceive such as having a certain odor, color, texture, weight, etc... But when the wax is brought close to the flame all of the characteristics change. The odor changes, its color changes, its texture changes, and so on. Nevertheless, we recognize the newly

---

<sup>1</sup> Actually, "cogito ergo sum" doesn't appear in the Meditations; it appears in Descartes' *Discourse on Methods*. The original phrase in the Meditations is "I am; I exist." The argument still hold. Everytime the phrase "I am; I exist" is recited, it must necessarily be true because only a thinking thing can recite it.

transformed wax as the *same piece* of wax as before. Descartes' point is this. Given that mind exists and is nonmaterial, or nonphysical, substance whose job it is to perceive things, then what is the mind tracking if all of the physical (possibly illusory) characteristics have completely changed? Presumably the mind isn't undergoing change and is, thus, not tracking itself. Nor is it tracking some other mind, since mind is immaterial; therefore, it must be tracking something else – the essence of the wax. But this means that it would be something that is not *res cogitans*. What the mind is tracking, then, is *res extensa* – something physical and, therefore, nonmental. And since change is real (otherwise what would the mind be tracking?), then the thing that the mind perceives outside of itself is also real. Therefore, mind and bodies exist: nonphysical and physical stuff exists. The position that claims that there exist both material and nonmaterial substance is called **substance dualism**.

### **PART III: THE PROBLEM OF MIND AND BODY INTERACTION**

But, then, we now seem to have a very basic and obvious problem. Consider what our normal story is about the interaction between mind and body. I have the belief that there is a beer in the fridge. Furthermore, I believe that I am thirsty and that drinking the beer would relieve my thirst. Subsequently, my beliefs about beer mixed with my desire to relieve my thirst – all mental states – cause me to get up and go get a beer. Similarly, my belief that my car is out of gas along with my belief that putting gas in the car is the best way to make it run mixed with my desire to make it run cause me to stop by the corner store and put gas in the car. This description would most likely not surprise anybody. It is a very natural way that we go about the world. But, let us now consider another scenario.

Consider our normal stories about ghosts. We say that ghosts are immaterial beings that pass through solid objects like walls and cannot be caught. They also do things like scare us by hurtling objects across rooms, slamming doors shut, and making various strange and bizarre sounds. Now, ghosts might actually turn out to be something quite different from what we have just mentioned. This is irrelevant, though. What is relevant is what we normally say about ghosts since this is what most people tend to believe about them: they are immaterial beings that interact with material stuff. Our question is simple: *how is that possible?* How is it, exactly, that a ghost can walk through a wall and, yet, move a solid object? By analogy, then, if our beliefs, desires, hopes, fears, dreams, and so on are immaterial stuff, then how do they interact with the physical body and get it to do things possible? How does the immaterial causally interact with the material?

#### PART IV: FROM THE MIND-BODY PROBLEM TO THE PROBLEM OF CONSCIOUSNESS

We begin from the issues raised by **Descartes' Meditations**. What do we say minds are? What do we say bodies are? Along with Descartes, we claim that mind is a thing which thinks, understands, wills, affirms, denies, loves, judges, plans, organizes, etc. and, as well, does not take up space. Body, on the other hand, does take up space (is subject to the laws of physics ...) and is stuff that is extended in space, is extended in a unique space, has form (shape), is perceivable, and so on. For Descartes, then, there are two fundamental substances that constitute the world, mind (**res cogitans** – thinking substance) and body (**res extensa** – extended substance).<sup>2</sup> This is a position known as **substance dualism**.

Now, the problem lay in the 'same old story' about mind and body, i.e., mind causally interacts with the body and directs it. The question is "How?" How is it that something that is essentially non-material can causally interact with something that is essentially material. How do ghosts (according to our 'same old story' about ghosts) walk through walls and, yet, move the vase? How does God heal the sick? And how does the mind direct the body? This is just what is known as the **Mind-Body Problem**; it is the problem of the causal interaction between mind and body.

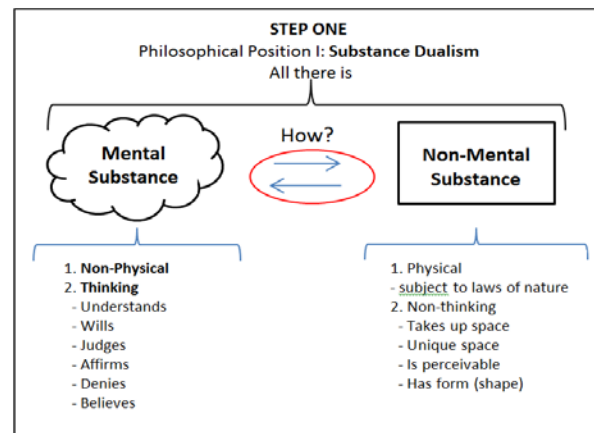


Fig. 1. Assume that the arrows show causation.

For the moment, set this problem aside, we will come back to it later. Let us ask another question: *what evidence do we actually have of a disembodied mind?* In other words, how is it that you know that the person next to you has a mind? There are three important points to

<sup>2</sup> To be fair, there were three substances for Descartes: mind, body, and God stuff. However, for our purposes, we need not introduce the latter.

notice, here. FIRST, we notice (if we are being honest with ourselves) that I begin with the assumption that I, from my first-person, personal, and subjective perspective, have a mind. I assume that I am a thinking thing that understands, wills, affirms, denies, loves, judges, plans, organizes, etc. But notice (and we'll come back to this later) that this is an assumption. SECOND, I notice that I do not have a first-person, personal, and subjective perspective of anybody else around me – that perspective applies only to myself. THIRD, I notice that my evidence for my neighbors' mentality comes by way of *what they do* or, to put it another way, *how they behave* and then make the analogy that when I 'think,' or 'feel,' or 'desire,' or 'judge,' or engage in any of the other so-called mental activities, that that is often accompanied by certain types of linguistic or, otherwise, bodily behavior. That behavior, linguistic or otherwise, is, nevertheless, **physical behavior** and is, thus, *observable* and, to varying degrees, *measurable*. So that when my neighbor put her hand up in the air and says, "Hello," I know that she *recognizes* me – where the term "recognition" is a **mental state term**, i.e., a term that refers to the mind and mentality – precisely because I who has a mind behaves in exactly the same way when I see my neighbor.

A second line of reasoning taken from neurological data further suggests that my evidence for others' mentality is physical. That is, there are large data sets that correlate my having of some **mental state** with some neural activation in the brain. This correlation is not simply a one-off instance. The report of mental states can regularly, systematically, and reliably be correlated with brain states. Furthermore, none of us in the Twenty-First Century are particularly shocked by this fact.

Our considerations from the two preceding paragraphs seem to suggest that substance dualism is either false or seriously flawed since all of our evidence for the non-physical mind comes from 1) physical evidence of behavior and 2) analogy by assumption of our own mentality. This does not support the thesis that there are two fundamentally different kinds of stuff – mental stuff and non-mental stuff. Rather, these considerations raise what is known as **The Problem of Other Minds**; namely, how do we know that others have them considering the actual evidence we have?

#### STEP TWO

The Problem of other Minds: How do we know there are any if all of our evidence for mentality is physical? We can see this in terms of an argument and ask, does the non-physical follow from the physical?

##### Standard Form.

[1. If there is physical Behavior, then there is mental behavior ( $P \rightarrow Q$ )]

2. Physical Behavior (P)

3. Mental Behavior (Q)

##### Logical Symbolic Form.

1.  $P \rightarrow Q$

2. P

3. Q

So, logically, we can construct a completely valid **modus ponens** argument with an implied first premise demonstrating that the mental does follow from the physical such that: *If there is physical behavior, then there is mental behavior. And in the case that there is physical behavior it would necessarily follow that there would be mental behavior.* Of course, logic has its limits. Outside of setting up such an exercise as a matter of logical interest, it simply seems weird to conclude that non-physical stuff follows from the fact of physical stuff. So, then, how is it that we can claim that others have a mind?

Fig. two.

This reasoning seems, then, to suggest that substance dualism is, again, either false or seriously wrong. What comes from this is the idea that, perhaps, all there is is physical stuff and that that physical stuff has properties or, perhaps, characteristics. This is a position known as **Property Physicalism**. In the broadest sense, all that is meant by the term “physical” is simply the assumption that whatever exists is matter that is subject to the laws of nature – however they actually turn out to be – and that there is nothing beyond the natural world. Property physicalism also assumes that the world is **causally closed**. What this means is that the universe does not admit of an infinite series of one cause to another and that all causation happens within the universe and does not come from without. The reason for this is that *ad infinitum* causal inferences and explanations are uniquely difficult, if not impossible, to understand. They are often as much or more of a mystery than the phenomenon being explained. Similarly, given that some phenomenon needs an explanation, if we allow that that explanation should come from outside the limits of the natural universe – the universe of space and time as we know it – then, if the explanation is to make any sense, that thing that is being used to do the explaining must also be understood. But since it is not clear what an extra-natural explanation would look like or how intelligible it really is, using it as an explanation is tantamount to using a mystery to explain a mystery. This is no real explanation at all.

So, then, property physicalism is the position that the fundamental stuff of the universe is physical and that this physical stuff has properties. Now, we notice that in our world of physical stuff and its properties, that some properties are observable or otherwise measurable and others are not. For example, my ceramic coffee mug is a certain color, weight, height, ability



to keep liquid hot or cold for certain periods of time, and so on. These properties can be seen or directly measured and we can tell some sort of physical story that explains the unique properties. The mug's property of color can be explained in terms of its microphysical structure and its interaction with other stuff in the world along with the physical story about how photons traveling from the sun refract off of the mug and onto our retinotopic nerves and so on, and so on.

Then there are the properties that we don't see and, yet, are still amenable to physical description and explanation. Such properties might be fragility, durability, inspirability, and so on. It is possible to give a complete physical and intelligible account of the mug's fragility in terms of its microphysical structure, yet, it is not a property that is observable in the way that color or height is. One sees the whole mug and then one sees the mug having shattered, but one never witnesses the property of being fragile. The same goes with the property of being inspirable. I can refer to the artwork on my mug as being inspiring simply by referring to the arrangement of atoms of the ink on the mug. A different arrangement of the same atom would not be inspirable. Thus, inspirability is a property that is explicable in physicalist terms, yet, is a property that is nevertheless not observable or measurable by any normal means.

Mentality seems to be a set of properties that is physical (no physicality, no mentality) and, yet, at least some mental states do not seem to be amenable to physicalist explanation. Consider a belief and let us be charitable and pretend that we have something like a completed neuroscience. It is entirely within the range of possibilities that we would be able to explain the having of a belief in physicalist terms by referring to the structure of the brain and the firing of neural networks, etc. ... Similarly, I would seemingly be able to give a physicalist account of someone's being in love by referring to their neural activity. However, what seems unlikely is that that physicalist account would provide any information as to *what it was like* to believe or what it was like to be in love. We could imagine an extra-terrestrial who saw someone get down on their knees and propose to someone. Imagine that ET asking why is he or she doing that. Our response would likely include the phrase "He loves her." Now imagine our ET asking what it was like to be in love. We good physicalists would proceed to tell the story of how the brain was working and what pheromones were being released and how ... But, of course, this wouldn't speak to the *feeling* of being in love, the phenomenal and qualitative aspect of *what it was like* to be in love and our ET would, undoubtedly, leave very confused. And it is this aspect of property physicalism that spells trouble for the theory. If one wishes to claim that all properties are physical and yet some properties are not amenable to physical explanation, then there is a

problem with using the physical as the manner of explanation. This suggests that property physicalism is either false or at least seriously flawed.

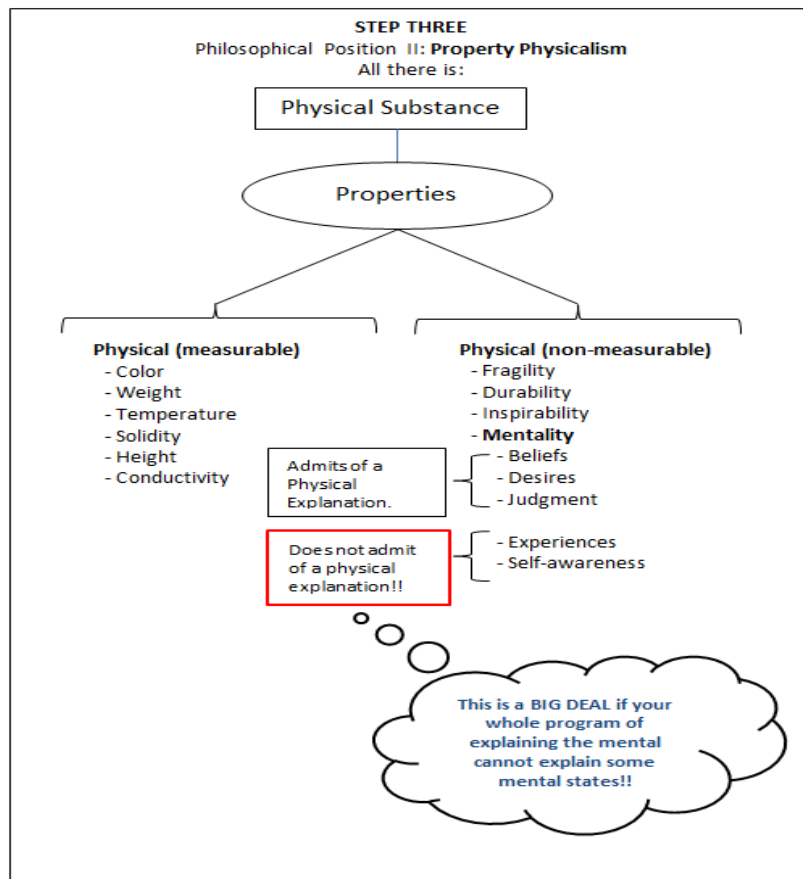


Fig. three.

**Step Four** in our little scheme, then, is simply to recognize the problem that some physical properties do not admit of a physicalist explanation. This leaves, yet, a third philosophical position open known as **Property Dualism**. Property dualism takes into account the fact that some physical properties do not admit of a physicalist explanation, or that some properties pertaining to the physical universe are not reducible to the physical. In other words, the property dualist will claim that, although, there is only one kind of stuff – physical stuff – in the universe and that the universe is causally closed, there are nevertheless two types of fundamental properties – **mental properties** and **non-mental properties**.

Non-mental properties are split up into those that are observable/measurable and those that are not. Observable, non-mental properties would be those such as color, weight, temperature, etc. ... Non-observable, non-mental properties would include properties such as

fragility, durability, inspirability, etc ... In each case, these properties are explicable by reference to their categorical, or micro-physical, structures – the structure of their basic particles and their interactions with the environment.

Mental properties, on the one hand, include things like desires, dreams, anger, joy, sadness, *et cetera*. These properties are explicable by natural and physical means. My desire for an orange, for example, will involve some physical story about irradiated photons, light waves, eyeballs, retinotopic lenses, neurons, brains, and so on. The same goes for anger, sadness, and the rest. On the other hand, there are some mental properties such as *being angry*, or *being in love*, or *having a religious experience*. The example of, say, 'what it is like to be angry' is somehow very different from the state of being angry. That is to say, that the state of anger can be perfectly explained in physical, natural terms; however, no amount of physical information will ever tell me what it is like to 'be in love' – it is seemingly an experience that one must go through one's self. Still, this is perfectly acceptable for a property dualist since there need be no final explanation for the properties; all one needs to know is that, explicable or not, in the physical universe, some properties are explicable by physical, natural means and other are not.

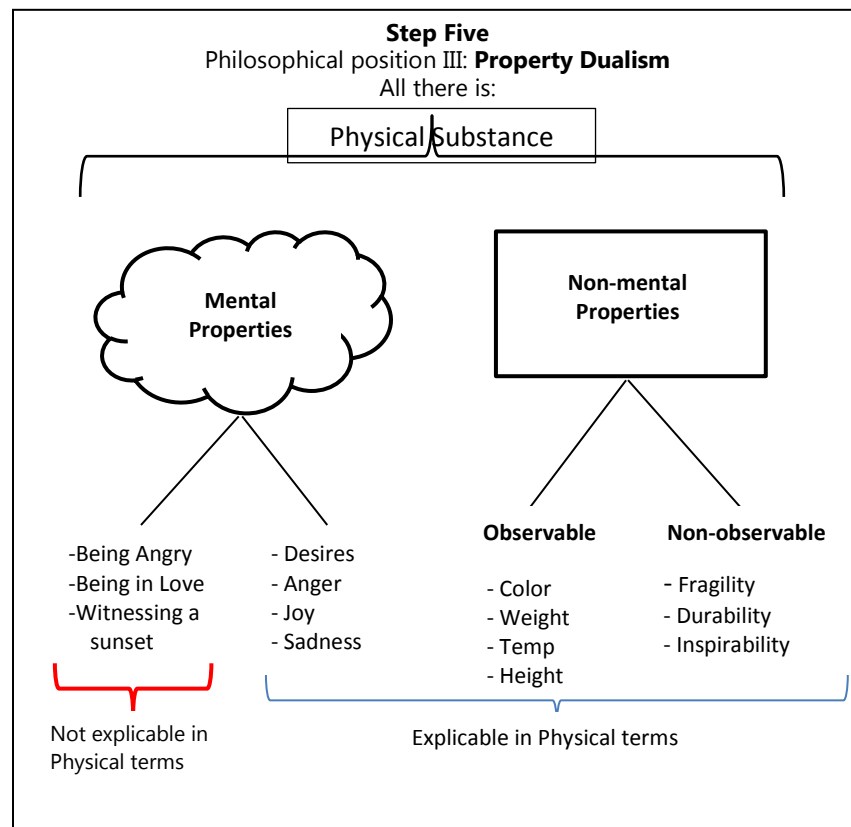


Fig. four.

But, of course, this still leaves it an open question as to what consciousness is. If what is wanted is some sort of theory of mind which takes into account the nature of consciousness, then none of the theories that we have encountered seems to be able to take up that challenge – property dualism included. This problem, the so-called **hard problem of consciousness**, is what seems to be so mysterious and elusive. We can simply phrase the problem as such: *how is it that a three and a half pound sack of neurons, cholesterol, some chemicals, and a bit of electricity can give rise to consciousness any more than a three and a half pound sack of potatoes?* It is all just physical stuff, after all.

To look at the problem differently, take a typical conscious, phenomenal experience, i.e., an experience that we are aware of and that is qualitative, such as the experience of striking a match. Children are often amazed at the experience of striking a match: it is a bright yellow color, surprising, mysterious, and so on. Now, the three and a half pound sack of cholesterol and chemicals that is sitting on top of your shoulders is grey. The question, then, is how is it that that three and a half pound of grey matter can have an experience that is *bright yellow*? Or, how is it that that three and a half pound sack of cholesterol and chemicals be surprised? Colin McGinn, in his 1989 article, *Can We Solve the Mind-Body Problem?*, phrased the problem like this:

How is it possible for conscious states to depend upon brain states? How can technicolour phenomenology arise from soggy grey matter? What makes the bodily organ we call the brain so radically different from other bodily organs, say the kidneys-the body parts without a trace of consciousness? How could the aggregation of millions of individually insentient neurons generate subjective awareness?

And what a problem to have!

## **PART V: THREE (MOSTLY) PHYSICALIST RESPONSES: BEHAVIORISM, IDENTITY THEORY, FUNCTIONALISM**

There are three recent responses to the mind-body problem and, to a lesser degree, the problem of other minds. We'll have a closer look at them later on. So, for now, we'll just briefly mention them and consider one striking objection to all of them. One response is **Behaviorism**. Building off of the idea that all we have access to is physical behavior for the mental, Behaviorism suggests that we abandon talk of the mental altogether or, at the very least, if we have to use mental language, then we should translate mental language into behavioral language. Behaviorists, then, refer to mental states as **dispositions**. A disposition, generally, is a

regular tendency to behave is certain sorts of ways. The regular tendency that glass has to shatter when struck is often referred to as the disposition of fragility. So, fragility is the disposition to shatter when struck. Likewise, mental state terms would receive a behaviorist translation. To know that, say,  $2 + 2 = 4$  where the word "know" is a mental state term simply *means* that one is disposed to answer "Yes" to the question 'Is two plus two equal to four?' Or, to take another example, be in love *means* to be disposed to giving flowers accompanied by a card that reads *I love you* to a particular individual. In other words, if I were to utter the following sentence (hence, my use of the double-quotation marks):

"I know that the TARDIS is the Doctor's time machine."

Then the behaviorist would translate this sentence into the following:

To *know* that the TARDIS is the Doctor's time machine is to respond "Yes" to the question, "Is the TARDIS the Doctor's time machine?"

since the normal way that we use the verb "to know" refers to a mental state. But since mental states are unobservable and behavior (even linguistic behavior) is observable, then the regular proposition which references mental states gets translated into behaviorist language which has the advantage of being reducible to and, hence, explicable by the study of behavior. It avoids the mind-body problem as well as the problem of other minds by simply referring to behavior. Other minds and causal efficacy need never be referenced!

Another response is the **Identity Theory** of mind. The Identity Theory, likewise, is a behaviorist model of mind. The difference being that the identity theorists make the rather more robust claim that the mind is *identical* to certain kinds of brain activity – the mind is the brain. This can be expressed with the following definition:

**DEF 1:** For all states  $(x)$ , if  $(x)$  is a mental state  $F$ , then  $(x)$  is a brain state  $G$ .

In symbolic notation, it reads

$$(\forall x)(Fx \rightarrow Gx)$$

Put differently, we might say:

**DEF 2:** For all mental states  $(x)$ , mental states are brain states,  $F$ .

Symbolically this reads

$$(\forall x)Fx$$

What is intended, here is the strict identification of mental states with brain states. What is not intended is to say that all brain states are mental states for, certainly, it is not the case that every instance of your brain doing something such as your thalamus keeping your heart pumping is a mental state. So, then, *it is not the case that for all states (x), if (x) is a brain state G, then (x) is a mental state, F* and represented by the formula  $\neg(\forall x)(Gx \rightarrow Fx)$ .

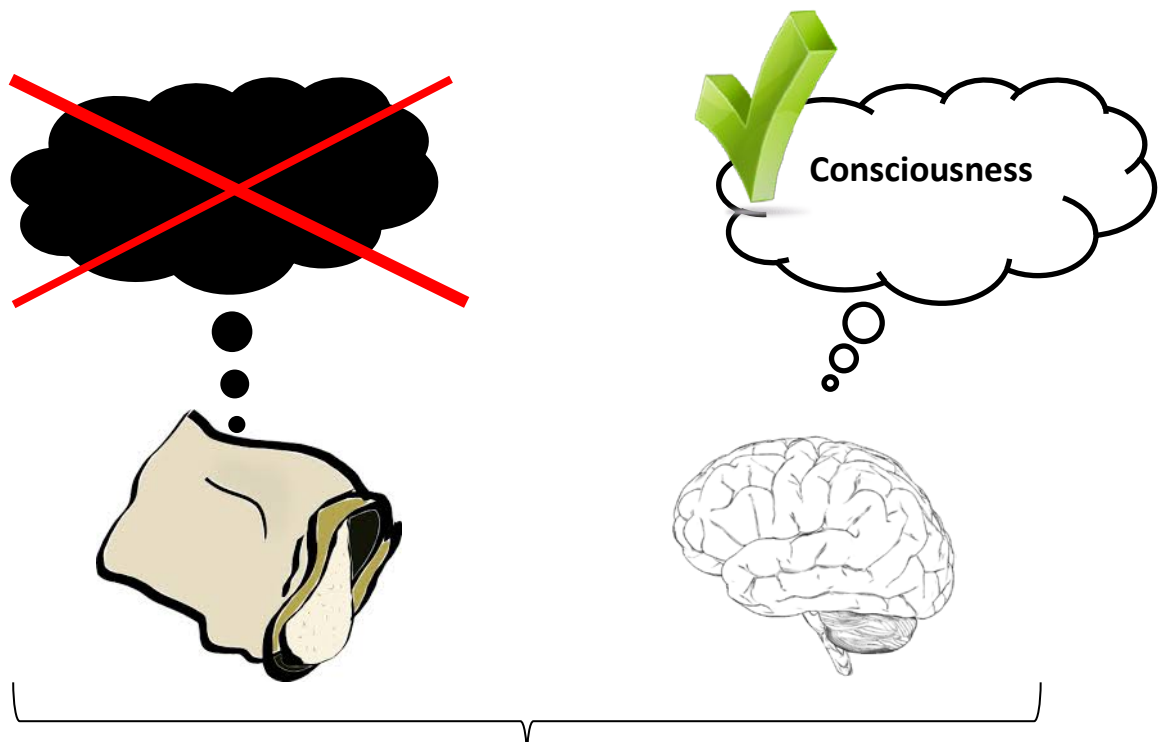
So, then, taken quite literally, your belief that grandmother loves you *is nothing more than* a certain neural pattern firing in your brain. Identity, here, is local to brain activity such that *if it were the case that there were no brain, then there would be no mind*. As with behaviorism the identity theory side-steps the problem of mind-body and the problem of other minds simply by identifying the mind as the brain.

Another response is known as **Functionalism**. Functionalism takes its name from the fact that what mental states are are place holders in some causal network of inputs and outputs. So, for example, let's say that I type the following sequence of keys on an electronic calculator: *1, +, 1. Enter*. What would the output be? *2*, of course. So, entering the information is the **input** and the conclusion *2* is the **output**. The output, *2*, was a result of, or was caused by, the input *1, +, 1. Enter*. In this case, it is said that the *function* of addition was realized by the electronic calculator. Now, let's say that I perform the same function on a wooden abacus. I arrive at this by a series of analogue steps where I move the wooden beads up and down the rows and columns whereby I end up with the two right most columns in Row II having three beads pushed to the top. Here, the *function* of addition was realized by the analogue abacus. In each of these cases, the inputs were some sort of information and the output was some sort of other information, but both the electronic calculator and the abacus – both made out of very different material – are properly said to be adding machines precisely because they performed the job of addition.

Thinking of this in mental terms, then, a mental state is simply a state that mediates between some input and some resulting output in a causal way. The joy I feel when I bite down into a York peppermint patty (an internal mental state) is nothing but the internal state triggered by putting a bit of the patty into my mouth and resulting in a smile on my face. In the case of human beings, as it turns out, that internal state seems to be occupied, or realized, by grey matter. But notice that there is nothing necessary about that role being occupied by grey matter. I might make a hammer out of frozen jelly and tap a nail into the wall or I might fashion a pair of shoes out of hard plastic. Now, they might not be a *good* hammer or shoe, but they are nevertheless a hammer and a shoe because they are performing the *hammer job* and the *shoe job* regardless of what materials those jobs are realized by. By analogy, then, mental states can be realized by any number of substrates; pain may be realized by grey matter in humans, and by

d-fibers in cuttlefish, and by siliconic-manipulation in ETs. The point is that mental states are occupiers of some causal role within a network of functions: *mental state X is as mental state X does*.

As you might have guessed by now, the problem (among others to be discussed in a later chapter) with all three theories just discussed is that they all fail to capture the phenomenological aspect of our experiences – the 'what it's like', qualitative aspect of *consciousness*. This is, of course, nothing more than the hard problem of consciousness.



Both are nothing but physical stuff made up out of atoms and molecules (DNA, RNA, iron, water, sugar, etc ...) that are subject to the same laws of nature. Why does one have conscious experiences and the other not?

**Ta' da!**

