

Chapter 6: Philosophy of Science

Overview: By focusing on ideas developed during the scientific revolution, this chapter takes us away (somewhat) from the metaphysics of God and leads us into the epistemology underlying science. We will discuss two competing worldviews of the time: rationalism, the idea that reason is primary source of knowledge, and empiricism, the idea that experience is the primary source of knowledge.

Major Ideas: After reading the material in this chapter and hearing the lecture, you should understand the following major ideas in depth, but other parts of the reading may appear on the assessment (besides names and dates).

Rene Descartes

Rationalism

Truth as Coherence

Methodic Doubt and the Cogito

The Ontological Argument

The Mind-Body Problem

David Hume

Empiricism

Impressions and Ideas

No Permanent Self

Personal Identity

Causality as Habit

Rene Descartes

We will get into Descartes soon, but first a few words on the historical period we are now turning to, and how it differs from the medieval period. In the medieval period, God was front and center. Reason mattered, but as we saw, God dominated. Faith and reason were discussed together, with faith typically coming out on top. As much as Aquinas respected Aristotle's rational worldview, he still had faith in a monotheistic Christian God.

In this chapter, we are looking at what's often referred to as *the scientific revolution* or *European enlightenment*, a period in which reason again dominated and the seeds that would later become the scientific method were sewn. Philosophers sometimes refer to the work done during this period as *modern philosophy*.

However, the belief in God did not disappear, though it did take a more individualist turn. Martin Luther (1483-1546 c.e.), who nailed a list of criticisms of the Catholic Church to its doors, is a famous example of the religious individualism of the time. Luther rejected the Catholic Church's claim that they were the true authorities on God, arguing that individuals could use their own powers of reason to understand the Bible. Luther started a whole new brand of Christianity, rejecting Catholic practices like confession, sometimes referred to as *The Protestant Reformation*.

Man's supposed place in the universe was also seriously challenged by Copernicus (1473-1543 c.e.), who put forth the idea that the sun, and not the earth, is the center of the solar system. Illustrating the complexity of the time, even Martin Luther, who had rebelled against the authority of the Catholic Church, did not believe that the sun was the center of the solar system. Not only did this discovery by Copernicus place less importance on earth—it's just another planet—but it placed less importance on those humans on that earth. What are humans but some random lifeform on some random planet in some random solar system in

some random universe? This insight about humans' place in the universe will be a central theme once we get to *existentialism* later.

Although the focus on reason clearly helped to develop the scientific method, as we'll see, it also had its detractors. There was an attitude of optimism during this period, as though a focus on reason would solve all of humanity's problems. Sometimes people were pejoratively said to have *faith* in reason, in other words an irrational, religious devotion to reason. Nevertheless, this focus on the use of reason to better our lives is a fundamental part of American thought today, since our founding fathers were children of the European enlightenment. In some cases, as we will see with political philosophy, elements of our founding documents include philosophical concepts created by enlightenment philosophers.

Philosophically speaking, there are a few influential and important ideas that arose during the scientific revolution. It is important to note that science as we know it did not arise in a vacuum, it had to be developed. In this chapter, we will look at some of the ideas that are today foundations of the scientific method. Philosophically speaking, the focus of this time was epistemology, theory of knowledge. Science today, after all, is supposed to give us knowledge/truth. We will focus on two of the central theories from the time, and two archetypes who represent those theories. The first is Rene Descartes (1596-1650 c.e.), the rationalist.

We could also call Descartes a philosopher, mathematician, or scientist. Back then, there were not really academic disciplines the way we have them today. Academic disciplines arose, and continue to arise, as knowledge gets more specialized. But back then, if you were lucky enough to have access to knowledge (which most people, especially women, did not) then you pretty much studied everything. For example, although we know Isaac Newton as a scientist today, he can be referred to as a *natural philosopher*, and his famous work is even called *The Principia: Mathematical Principles of Natural Philosophy*. Descartes'

famous work, *Meditations on First Philosophy*, was one of the most influential of the European enlightenment.¹

But unlike many other thinkers of his time, Descartes insisted on carrying out all his calculations on his own, rejecting methods that had been established by his predecessors. For example, whereas scholars typically wrote in Latin, he wrote in his home language of French to reach a wider audience.

Rationalism

Descartes started the school of philosophy known as *rationalism*: the belief that reason is the primary source of knowledge. Again, don't forget that reason comes down to making rational arguments using justified claims and evidence. Descartes believed that reason could serve as a foundation for science itself. If we had a rational foundation, he thought, then every other piece of knowledge could be built upon it.

But how will we reach that foundation? By using reason, and reason alone. What's the alternative? One famous alternative view of the time was *empiricism*, or the idea that experience, and not reason, is how we primarily gain knowledge. We will look at empiricism in more depth below with Hume, but it's important to understand the implications of rationalism. For Descartes, one could simply sit and think in an easy chair and arrive at truth, without needing to venture into the outside world.

Sometimes the two Latin phrases *a priori* and *a posteriori* are used here. *A priori* literally means before experience, whereas *a posteriori* typically refers to during/after experience. When Descartes argues that rationalism is the path to truth, he is arguing that knowledge is fundamentally *a priori*, or happens before, or without reference to, experience. Hume, the empiricist, focused on the *a posteriori*. Notice that the difference

¹ Descartes, R. (1996). *Meditations on First Philosophy* (2nd ed.). J. Cottingham (Ed.). New York, NY: Cambridge University Press. Perhaps his other most famous work is: Descartes, R. (1998). *Discourse on Method* (3rd ed.). D. A. Cress (Trans.). United States: Hackett Publishing.

between the two is also the difference between deductive (a priori) and inductive (a posteriori) reasoning, as discussed in the first chapter.

Truth as Coherence

What is truth? is a very philosophical question, and one that probably seems pointless in a way. Indeed, some things seem obviously true, like math, or scientific facts. But how does something become a fact in the first place? Presumably a fact is somehow related to truth. So we have to have an understanding of a truth before we can determine what a fact is. Thankfully, today we have the scientific method to determine facts for us through logic and experimentation—but keep in mind that the very thing we are studying right now is the origin of the scientific method. Before the scientific method developed, there were different theories of truth, different hypotheses about how we can convincingly call something *true*.

One of these theories said that something is true if it is coherent with established truths, this is known as *truth as coherence* or *the coherence theory of truth*. In other words, we can call something true if it does not directly contradict what we already know to be true. A major component of this theory is *rational consistency*. For an idea to be a candidate for truth according to rationalists of the time like Descartes, it must be clear, distinct, and consistent.

Again, the rationalist method focused on what we would today call deductive reasoning. For example, I can deduce an important truth from the fact that a bachelor is an unmarried man. I can deduce that if I meet a man in the future who says he's a bachelor, then he must be unmarried. Crucially, I can do this a priori, without needing to experience anything. I can even write out the deduction logically:

1. All bachelors are unmarried men.
 2. Scott is a bachelor.
- So Scott is an unmarried man.

Deductive arguments like this, developed from rationalists like Descartes, are now staples of required college critical thinking classes. According to the coherence theory of truth, all truth is like this example in that it can be logically derived from other truths.

A serious objection to this theory is that it does not specify how something becomes true *to begin with*. Sure, there may be some obvious truths that can basically be assumed (like $2+2=4$), but this simplicity is not a characteristic of *all* truth. For example, the truth that the earth is the third planet from the sun was not arrived at by deducing it from other truths.

Related to the last objection is the implication that, if the coherence theory is accurate, then it could potentially lead us to believe a bunch of consistent *untruths*. What if we begin with a false idea, then derive a bunch of ideas that are consistent with it? Sure, the ideas would be consistent with the existing idea, but since the existing idea was false, then all the resulting ideas would be false too. Such is the problem with the coherence theory of truth, if we can't figure out what is true to begin with, then consistency doesn't seem to help all that much.

While this theory certainly has issues as a complete theory of truth, there is no doubt that rational consistency is necessary to science and to general reasoning. For example, scientists must understand terms consistently when they run their experiments. It might also behoove us to be rationally consistent in our personal lives. If I believe that everyone should be treated equal, but I find that I am treating one friend differently than another, then I am not being rationally consistent in the way I apply my beliefs, which may cause a problem in my personal relationships.

Methodic Doubt and the Cogito

Considering what was just discussed, it's not too surprising that Descartes tried to use the rationalist, a priori, deductive method to find

an ultimate foundation for all knowledge. What ultimate truth can we find that all other knowledge can be built upon? Descartes will give us an answer, but let's look first at his method of methodic doubt.

Methodic doubt is exactly what it sounds like, giving a method to the process of doubting. But unlike informal doubting, methodic doubt has an aim, a goal: the ultimate truth. Methodic doubt is a method of doubt that meticulously goes through different ideas to test their certainty, rejecting them when they are uncertain. Whatever is left after this process of doubt will, by definition, be certain, something incapable of being doubted. It is important to remember that Descartes' method does not just entail doubting for the sake of doubting; rather his method entails doubting to find certainty.

Some accounts of Descartes' life suggest that he closed himself off in his room to begin the process of doubt so he would be undisturbed. Descartes began by considering ideas that he could state clearly and distinctly, rejecting them if they did not stand up to scrutiny. It's important to note that Descartes' process here must be understood from a subjective, personal level. You must try to go through what Descartes went through to fully understand his reasoning. In this sense, this material is more personal than anything we've discussed thus far.

Descartes first rejects our senses as being certain. By *senses* Descartes is referring to seeing, hearing, tasting, touching, and smelling. Strange isn't it? If anything, it seems our senses are certain. Aren't your senses telling you right now that you're seeing this reader in front of you? How could they be wrong about that? We should remember that senses are a *medium*, not a direct window into the world. And we should remember that our senses can deceive us, as is clear with sights like optical illusions and mirages. And if senses can deceive us, then they are not certain. Descartes is looking for absolute certainty, not just reliability. While senses might be said to be reliable, they are not certain. Additionally, our senses can be effected by other factors, like by whether

we're asleep or dreaming. In fact, sometimes we conflate the dream world with the real world.

What other candidates do we have for certainty? Another go-to candidate is math, or mathematical truths. How could a truth like $2+2=4$ ever be uncertain? Isn't math the ultimate certainty? It would seem so, and yet Descartes does not find math to be the ultimate certainty due to the possible existence of something he calls *the evil demon*. The evil demon, Descartes said, is an all powerful being like God, but evil instead of good. Imagine a God-like being with the power to create and destroy, but who does not have our best interests in mind—kind of a like a little kid with a magnifying glass ruling over humans like ants. To be clear, Descartes is not saying that the evil demon actually exists; he's saying that it's *possible* such a being exists. And if such a being did exist, he could deceive me about math. He could make me believe whatever he wants since he is all powerful. He could make me believe that $2+2=5$. Although the demon's existence is a remote possibility, the possibility alone suggests that math cannot be an ultimate, absolute certainty. Is math reliable like our senses? Yes. But not certain.

So what are we left with, then? Unfortunately, some students who study this material get the impression that Descartes' reasoning stops here with total uncertainty about everything. If this is where Descartes' reasoning actually stopped, then it would lend credibility to the claim that philosophy is just pointless theorizing. Thankfully, Descartes is presenting a logical argument with a conclusion, so his reasoning does have a structure and a purpose. Let's look at his full argument before getting to his conclusion, the ultimate certainty.

1. Anything that can be doubted is not certain.
 2. I can doubt my senses.
 3. I can doubt math due to the possibility of an evil genius.
 4. I cannot doubt my internal thoughts,
 5. My internal thoughts confirm that I exist as something.
- Thus, existence itself is certain.

We just covered the first three premises so let's look at the last two. Premise 4 says that you cannot doubt your internal thoughts—again, try to consider this from your own subjective perspective. Even if your senses are deceiving you, and an evil demon is messing with your mind, you would still know that you have internal thoughts. You don't know that *others* have internal thoughts, but you know that *you* have them. This suggests that you exist as something, which is premise 5, leading to the conclusion that existence alone is the only certainty.

But wait a minute, if the evil demon can deceive you about math, can't he deceive you about existing as well? Couldn't existence itself be an illusion? Well, yes, the nature of your existence could be an illusion, but not existence itself. No matter how much the evil demon attempts to deceive you—makes you believe $2+2=4$, that the earth is flat, that your name is Aristotle—you would still know that you are thinking when you look at your internal thoughts. For Descartes, the certainty comes from *thinking in the present moment*. In other words, Descartes is saying that thinking itself is certain, not the content of thought. Descartes is *not* saying that whatever you think is certain. For example, if you think that you are an honest human being, it's not your honesty that is certain, it's *the fact that you are thinking it* that is certain. Your thoughts may or may not refer to a reality beyond you, but they are present while you think them, and that's the point.

Descartes is most famous for the phrase, *I think, therefore I am*. Now we can see why, because for Descartes, thinking itself is the evidence/justification for existence. Sometimes the argument is called *the cogito*, since I think, therefore I am in Latin is *cogito ergo sum*.

The selection that follows comes from Descartes' *Meditations on First Philosophy* (cited above) and shows his cogito argument in its original form.

It is now some years since I detected how many were the false beliefs that I had from my earliest youth admitted as true, and how doubtful was everything I had since constructed on this basis; and from that time I was convinced that I must once for all seriously undertake to rid myself of all the opinions which I had formerly accepted, and commence to build anew from the foundation, if I wanted to establish any firm and permanent structure in the sciences... And for that end it will not be requisite that I should examine each in particular, which would be an endless undertaking... I shall only in the first place attack those principles upon which all my former opinions rested.

All that up to the present time I have accepted as most true and certain I have learned either from the senses or through the senses; but it is sometimes proved to me that these senses are deceptive, and it is wiser not to trust entirely to anything by which we have once been deceived.

But it may be that although the senses sometimes deceive us concerning things which are hardly perceptible, or very far away, there are yet many others to be met with as to which we cannot reasonably have any doubt... there is the fact that I am here, seated by the fire, attired in a dressing gown, having this paper in my hands... And how could I deny that these hands and this body are mine, were it not perhaps that... certain persons... have an earthenware head or are nothing but pumpkins or are made of glass.

At the same time I must remember that I am a man, and that consequently I am in the habit of sleeping, and in my dreams representing to myself the same things or sometimes even less probable things, than do those who are insane in their waking moments. How often has it happened to me that in the night I dreamt that I found myself in this particular place, that I was dressed and seated near the fire, whilst in reality I was lying

undressed in bed! At this moment it does indeed seem to me that it is with eyes awake that I am looking at this paper; that this head which I move is not asleep, that it is deliberately and of set purpose that I extend my hand and perceive it; what happens in sleep does not appear so clear nor so distinct as does all this. But in thinking over this I remind myself that on many occasions I have in sleep been deceived by similar illusions, and in dwelling carefully on this reflection I see so manifestly that there are no certain indications by which we may clearly distinguish wakefulness from sleep that I am lost in astonishment. And my astonishment is such that it is almost capable of persuading me that I now dream

... whether I am awake or asleep, two and three together always form five, and the square can never have more than four sides, and it does not seem possible that truths so clear and apparent can be suspected of any falsity [or uncertainty]

Nevertheless I have long had fixed in my mind the belief that an all powerful God existed by whom I have been created such as I am. But how do I know that He has not brought it to pass that there is no earth, no heaven, no extended body, no magnitude, no place, and that nevertheless [I possess the perceptions of all these things and that] they seem to me to exist just exactly as I now see them? And, besides, as I sometimes imagine that others deceive themselves in the things which they think they know best, how do I know that I am not deceived every time that I add two and three, or count the sides of a square, or judge of things yet simpler, if anything simpler can be imagined? But possibly God has not desired that I should be thus deceived, for He is said to be supremely good. If, however, it is contrary to His goodness to have made me such that I constantly deceive myself, it would also appear to be contrary to His goodness to permit me to be sometime deceived, and nevertheless I cannot doubt that He does permit this.

I shall then suppose, not that God who is supremely good and the fountain of truth, but some evil genius not less powerful than deceitful, has employed his whole energies in deceiving me; I shall consider that the heavens, the earth, colours, figures, sound, and all other external things are nought but the illusions and dreams of which this genius has availed himself in order to lay traps for my credulity; I shall consider myself as having no hands, no eyes, no flesh, no blood, nor any senses, yet falsely believing myself to possess all these things...

I suppose, then, that all the things that I see are false; I persuade myself that nothing has ever existed of all that my fallacious memory represents to me. I consider that I possess no senses; imagine that body, figure, extension, movement and place are but the fictions of my mind. What, then, can be esteemed as true? Perhaps nothing at all, unless that there is nothing in the world that is certain. But how can I know there is not something different from those things that I have just considered, of which one cannot have the slightest doubt? Is there not some God, or some other being by whatever name we call it, who puts these reflections into my mind? That is not necessary, for is it not possible that I am capable of producing them myself? I myself, am I not at least something? But I have already denied that I had senses and body. Yet I hesitate, for what follows from that? Am I so dependent on body and senses that I cannot exist without these? But I was persuaded that there was nothing in all the world, that there was no heaven, no earth, that there were no minds, nor any bodies: was I not then likewise persuaded that I did not exist? Not at all; of a surety I myself did exist since I persuaded myself of something [or merely because I thought of something]. But there is some deceiver or other, very powerful and very cunning, whoever employs his ingenuity in deceiving me. Then without doubt I exist also if he deceives me, and let him deceive me as much as he will, he can never cause me to be nothing so long as I think that I am something. So that after having reflected well and carefully

examined all things, we must come to the definite conclusion that this proposition: I am, I exist, is necessarily true each time that I pronounce it, or that I mentally conceive it.

I am, I exist, that is certain. But how often? Just when I think; for it might possibly be the case if I ceased entirely to think, that I should likewise cease altogether to exist.

I am, however, a real thing and really exist; but what thing? I have answered: a thing which thinks.

But what then am I? A thing which thinks. What is a thing which thinks? It is a thing which doubts, understands, [conceives], affirms, denies, wills, refuses, which also imagines and feels.²

The Ontological Argument

Let's recall that Descartes was trying to build a new foundation for the sciences through his use of methodic doubt. But if the only thing we can be certain of is existence from a subjective perspective, where do we go from there? How can that serve as any sort of foundation, especially if math itself cannot be said to be certain? Abstract as it seems, Descartes' cogito influenced many scientists and continues to. If the external world can be doubted because our senses can be doubted, then that casts doubt on science itself, since scientific tools are merely extensions of our senses (telescopes, microscopes, etc.). Most scientists today acknowledge Descartes' argument, but claim that reliability is enough. Sure, our senses can deceive us and should not always be trusted, but that doesn't mean that our senses aren't reliable. In other words, we can trust our senses enough to strongly support the case for an external world around us. A professor of cognitive science at UC Irvine, Donald Hoffman believes that Descartes is partly right, arguing for what is

² From *Meditations on First Philosophy* (cited above).

known as *the interface theory of perception*.³ Hoffman claims that our perceptual apparatus is more like a computer interface in that it tells us what we need to know to function, but not the actual truth about the external world.

However, Descartes' solution to the deep skepticism about an external world that his cogito argument leads to is: God. That's right, Descartes says that an external world exists because God exists too, and a good God, by his very nature, would not deceive us. Additionally, the existence of a good God suggests that an evil demon could not exist. If you're thinking that Descartes' solution here is a bit of a cop out, you're not alone. In fact, some philosophy professors do not even teach about this part of his philosophy. However, I believe that it is important to follow his entire train of thought. Descartes is perhaps worthy of more criticism here too because his argument for God is not original and was first presented by an earlier theologian, St. Anselm.⁴ The argument is called *the ontological argument*. *Ontology* is a branch of metaphysics, and it focuses on the nature of existence (for example, being a ghost or spirit might be ontologically different from being a human or ape). Here is the argument:

1. A being is God if and only if it is the greatest conceivable being.
2. God exists in thought but not in reality.
3. Existence both in thought and in reality is greater than existence in thought alone.
4. We can conceive of God's existence in reality and in thought.
5. So we can conceive of a being greater than God.
6. So we can conceive of a being greater than the greatest conceivable being.

³ Hoffman has a TED talk on the topic, but here is the original article: Hoffman, D. (2009). The interface theory of perception: Natural selection drives true perception to swift extinction. In S. Dickinson, M. Tarr, A. Leonardis, B. Schiele (Eds.), *Object Categorization: Computer and Human Vision Perspectives* (pp. 148-165). Cambridge, MA: Cambridge University Press.

⁴ Find the argument here: Anselm, St. (2008). *Anselm of Canterbury: The Major Works*. B. Davies & G. R. Evans (Eds.). New York, NY: Oxford University Press.

Thus, God exists in reality as well as in the understanding.

To understand this argument fully, it is important to note that it fits an argument form known as *reductio ad absurdum*, or reduction to absurdity. A *reductio ad absurdum* argument tries to show why an idea or assumption leads us to a ridiculous or absurd conclusion, suggesting that we should *not* accept that idea or assumption. The ontological argument tries to show why it is ridiculous to believe that God exists in thought only, but not in reality, ultimately trying to get us to accept that God exists in *both* thought and reality. This argument suggests that if something is perfect, it means it must exist: perfection requires existence. The concept of God as a perfect being itself requires the existence of that being.

The main counter to the ontological argument revolves around the term *existence*. The argument seems to assume that existence is an *object*, something that we can point to, like redness or roundness. But, this counterargument suggests, existence is not an object, but a *property*. If existence is a property, then we cannot make a distinction between existing in reality versus existing in the understanding, invalidating the argument.

However, even if we did assume that perfection requires existence, then we would have to say that other perfect things exist just because we can conceive of them. For example, if I can conceive of a perfect island, then it must exist according to this logic. If *that* conclusion doesn't seem justified, this objection says, then neither does the conclusion that God exists.

The Mind-Body Problem

Whatever we think of Descartes' views, they seem to lead to a distinction between physical stuff and mental stuff. Sometimes Latin phrases are used to denote this difference: *res extensa* refers to extended physical things like cells, atoms, bodies, planets, and so forth. Whereas

res cogitans refers to mental things, like thoughts and internal experience. Recall that at the end of his cogito argument, Descartes establishes thinking as indicative of the one certainty, suggesting a difference between thought and the physical world around him. Sometimes this distinction is understood as a metaphysical view called *Cartesian dualism*, since Descartes is suggesting that the nature of reality is two-fold: *res cogitans* and *res extensa*.

However, a deeper problem arises here that is still being discussed today: how is it that mental stuff influences physical stuff, and vice versa? This problem of the interaction between mental and physical is called *the mind-body problem*. Throughout the years, many different solutions have been proposed.

One solution is *materialism*—a view related to Aristotle’s naturalism—suggesting that all mental processes can be reduced to physical ones. Thus, this view suggests that as much as there *appears* to be a difference between *res extensa* and *res cogitans*, *in reality* there is no difference. In reality, according to this view, the mind is just a very complicated physical process. The current problem with this view is that every effort to reduce the mind to physical properties has been either unsuccessful or dubious.⁵

Other solutions try to maintain some sort of dualism. *Interactionism* suggests that the mind and body are causally related, like a cue ball hitting the 8 ball. *Parallelism* suggests that God sets up mind and body to work simultaneously when he creates the universe.

Like materialism, dualism has issues as well. As one who corresponded with Descartes, Princess Elizabeth of Bohemia (1618-1680 c.e.) argued that interactionism is untenable since it is logically impossible for something non-physical (*res cogitans*) to have a causal effect on something physical (*res extensa*). Additionally, the law of conservation

⁵ Indeed, much of this research can be seen regularly in the pages of *The Journal of Consciousness Studies*.

of energy tells us that the level of energy in the universe remains consistent (where *energy* is defined as a property that must be transferred to an object for it to have an influence on that object). Energy cannot be created or destroyed, it can only change forms. However, if mental stuff really is ontologically different from physical stuff as Cartesian dualism suggests, then we would expect a spike or increase in the overall level of energy in the physical universe when the two distinct substances interact (for example, when a thought leads to a physical action). Unfortunately, this spike in energy is not observed, shedding doubt on dualism. Additionally, dualism suffers from the difficulty of even knowing how to test for it, potentially being vulnerable to the untestable explanation fallacy. Finally, similar to the criticisms of the first cause argument, dualism adds an extra level of explanation to the scientific method: we have to explain not just the physical universe, but the mental universe as well. Ockham's razor might be used to cut out the *res cogitans*.

The mind-body problem is a difficult one, and we are kind of like the early Presocratics talking about atoms when we discuss it today: we have a lot of hypotheses, but at this point we have little idea how to determine which one is true. In fact, it's not even clear at this point if we're asking the right questions. For these reasons and more, an entire new academic discipline has arisen called *cognitive science*. Donald Hoffman, mentioned above, is a cognitive scientist. As noted earlier, academic disciplines are created as knowledge gets more specialized. Cognitive science is an example of that: it is an interdisciplinary field in which people have come together to study similar issues about the mind. Cognitive science brings together the disciplines of philosophy, psychology, computer science, linguistics, and more. Largely, cognitive scientists are tackling the mind-body problem, which today is cashed out in terms of *consciousness*. Consciousness is the state of being an *I*, the feeling of being a thinking thing, likely what Descartes was referring to in his cogito argument. The question that fascinates many cognitive scientists is: how does consciousness arise in the first place? Even if we assume materialism, the answer to this question is far from clear. And

pursuing this question gives rise to related questions like: can machines think? Could we create consciousness in a machine artificially, or does consciousness have to arise through a natural process? Are other beings conscious, like animals? To what degree?

David Hume

Let's back up. We have been discussing Descartes, the rationalist, who believed that truths could generally be arrived at through reason alone. We see this belief reflected in the arguments presented above from Descartes in premise/conclusion format—all are deductive, a priori arguments. Now, we turn to another archetype of the scientific revolution, the empiricist David Hume, who focused on the a posteriori.

Hume is often credited as one of the first philosophers to really, truly follow his arguments to their ultimate conclusions, without, for example, invoking God to save his argument as Descartes did. In Hume's writings, one gets a sense of his unquestionable search for knowledge and a desire to know why, really, things are the way they are. Two of his most famous works are *Treatise of Human Nature* and *Dialogues Concerning Natural Religion*.⁶ As a person, Hume was known to be incredibly friendly and sociable. Although Hume is usually identified as an empiricist (one who focuses on experience over reason), he is sometimes identified as a skeptic as well, since he set a high bar for belief. A *skeptic* is someone who demands clear, undoubtable evidence before accepting something as knowledge.

Indeed, Hume was skeptical of religion, reason, and even some aspects of the scientific method that were developing while he was alive. For example, we already saw Hume's critique of the argument by design. For Hume, our lack of experience of universe-creation is enough to be skeptical that some type of God created the universe. However, it is important to note that he did not identify as an atheist, but an agnostic. An *agnostic* is typically defined as someone who believes that we do not know enough to say that God exists—maybe he does, maybe he doesn't. For Hume, all the arguments for God are not enough to establish God's

⁶ *Dialogues* was cited in the last chapter, but here's the other one: Hume, D. (1978). *Treatise of Human Nature* (2nd Ed.). L. A. Selby-Bigge & P. H. Nidditch (Eds.). Oxford, England: Oxford University Press.

existence, but the arguments against are also not enough to establish God's lack of existence.

Additionally, although Hume valued reason as a tool, he believed its limitations should be appreciated, famously saying that "Reason is the slave of the passions." Finally, Hume had something interesting to say about the scientific idea of cause and effect, which we will look at below.

It is important to note how far skepticism can go. If I accept Descartes' logic that my existence is the only thing I can be certain of, and I am skeptical that an external world exists at all, I might be led to solipsism. *Solipsism* is the extreme skeptical view that you are the only thing that exists in the universe. I said extreme, didn't I?

Hume's skepticism didn't go as far as solipsism, and Hume thought we could know that there is an external world through *inference*. Our minds create patterns, or infer patterns, that make the world coherent to us, suggesting that it actually exists. For example, I have not been to every town on the planet, but I can still infer that they all exist by inference.

Notice that Hume's logic here is connected to our *experience* of reality, so let's take a look at empiricism in more depth.

Empiricism

First of all, note that empiricism is a theory of knowledge, like rationalism. Again, the scientific revolution focused on epistemology. However, sometimes empiricism gets confused with naturalism, which is a metaphysical view about what exists in the universe.

Empiricism is the view that experience is the primary source of knowledge. A key component of any scientist's toolbox today, empiricism back then was one theory of how we acquire knowledge. Recall that Descartes' believed that our brains contain *innate ideas* from

which we can deduce conclusions. For example, a bachelor being an unmarried man is an innate, universal idea—or $2+3=5$.

Many empiricists leading up to Hume disagreed completely with Descartes. One empiricist named John Locke (1632-1704 c.e.) famously said that the mind is a *tabula rasa*, or blank slate, meaning that there are no innate ideas. He argued that rationalists like Descartes were merely assuming that ideas are innate when in fact, Locke said, these ideas come from experience. Ultimately, Locke argued that the universality in ideas come from the objects themselves; he even said objects have primary qualities that exist independently of the perceiver, and secondary qualities that the perceiver imparts to the object. The kicker is that we can never directly experience primary qualities, but we can experience secondary qualities (things like color, shape, etc.). So we see that the problem of truly knowing the external world continues. Even if we can infer that the world around us exists, can we ever really know it *as it actually is*? Or are we doomed to know only what our brains tell us, which may not represent the way things really are? This view of Locke's, that there is a distinction between the knower and the known, is called *epistemological dualism*.

Another empiricist leading up to Hume was George Berkeley (1685-1753 c.e.), the inspiration for UC Berkeley. Berkeley famously supported the Latin phrase *esse est percipi*, to be is to be perceived. He argued that nothing exists unless we perceive it, and that an external world does not actually exist. If a tree falls in the forest does it make a sound? Nope, not unless someone was there to hear it. This view, that only mental ideas exist, is sometimes called *immaterialism* or *idealism*.

Hume was definitely not an idealist in this sense, though he was influenced by empiricists like Locke and Berkeley. Hume's empiricism was similar to and different from the empiricism of contemporary scientists. It was similar in that both focus on experiencing and experimenting with the physical world around us to find truth—when scientists today conduct an experiment, for example, on health they are

using empiricism. However, Hume's empiricism was different in that he believed our internal thoughts are a legitimate object of objective experience. In other words, he believed that while you and I might have different beliefs, the literal process of our thinking is similar and can be studied. Hume believed that *introspection* can be empirical. While this belief may seem strange to us now in the west, people in the east have been using introspective empiricism for many years. When Buddhists meditate, for example, they often report similar experiences, such as having a wandering or *monkey* mind. If we did not have similar mental processes, how could we agree on a similar experience like this? Hume argued that our thoughts occur in three ways: contiguity (thinking of San Diego makes me think of the padres), resemblance (seeing a picture of an apple makes me think of an apple), and cause and effect (seeing a seed makes me think of a tree).

Impressions and Ideas

A prime component of Hume's empiricism is the *theory of impressions and ideas*. Basically, this is Hume's theory about how the mind works. Again, note that this material might actually be considered psychology today, and is in fact sometimes cited by current psychologists.⁷ But back then, it was merely one area of focus for philosophers.

For Hume an impression is an immediate perception, what you are seeing, hearing, tasting, touching, or smelling in the present moment. You are having an impression of the reader in front of you right now. But an idea is a later, duller impression, like an afterthought. (Hume also said that impressions and ideas could both be divided into simple and complex, but we won't get into that in this class).

There are a couple of consequences to Hume's theory. For one, for Hume, every idea can be traced to an impression. Even ideas of things that do not exist can be traced to multiple impressions; my idea of a

⁷ The current psychologists Paul Bloom and Jonathan Haidt, for example, frequently cite philosophers from the time of Hume.

castle on mars can be traced to my impression of mars and my impression of castles. Another consequence to the theory is that an impression is *always* more vivid than an idea since an impression is happening *now*. An idea always comes later, and is removed from direct experience.

Hume developed from his theory what is called *the empirical criterion of meaning*, which basically says that if an idea does not lead to a clear impression, then that idea is meaningless and might be rejected. In fact, an idea that doesn't have an impression backing it up isn't really an idea at all. A group of philosophers in the 20th century called *logical positivists* had a similar view, arguing that nearly all metaphysics is meaningless since it does not refer to anything concrete. However, it is questionable whether the theory itself can pass its own test.⁸

Nevertheless, Hume believed that the idea of a concrete or permanent self could not stand the test of the empirical criterion of meaning, leading him to conclude that there is no permanent self.

No Permanent Self

Recall the doctrine of anatta, the illusion of self, from the Buddha. Although Buddhists suggest that the self is an illusion, they are not really saying that the self doesn't exist. Rather, they are saying that the self as we know it, as a permanent, separate thing, doesn't exist. Hume made a similar point about the self, but from the western perspective using rational argument.

Using his theory of impressions and ideas, Hume concluded that there is no permanent self. It is important to remember that Hume's empiricism can include introspection. Upon looking at his internal experience, Hume found no impression of a permanent self. He did find impressions of thoughts and emotions, but he did not believe that these together

⁸ Is the claim that all claims that aren't verifiable are meaningless, *itself* meaningless? Does *that* claim have any empirical correlate? Not according to many critics of positivism.

constituted a permanent thing. At least, not the same way, for example, my impression of a pen constitutes a permanent thing.

Remember that the empirical criterion of meaning says that if an idea has no definite impression, then we must reject that idea. For Hume, the permanent self is worthy of rejection on these grounds. To be clear, here is Hume's argument for no permanent self:

1. All ideas can be traced to an impression.
 2. The idea of self has no definite impression.
 3. Only fleeting perceptions (thought after emotion, etc.).
- Thus, there is no permanent self.

So, the Buddha and Hume both tell us that there is no permanent self. Are they right? If we believed them simply because they are smart, famous people from history, we'd be committing the appeal to authority fallacy. And in fact, many thinkers have disagreed with the Buddha and Hume over the years, as we'll see now.

Personal Identity

When many people learn about the idea that there isn't a permanent self, they reject it, because there is a strong sense that we *are* something permanent in our day to day lives. The question is, though, what? What is that something that makes us permanent? The Buddha and Hume couldn't find anything permanent, so if there is something permanent, where can it be found? How can we understand it? These questions are addressed in an area within philosophy called *personal identity*, or the question of what makes you the same person from one day to the next. Do not confuse personal identity with the sociological question of how you *developed* your current likes and dislikes, your current personality. Personal identity goes deeper than personality. To illustrate, consider a famous thought experiment called *the Ship of Theseus*.⁹ In this story, the

⁹ This thought experiment has been discussed by many philosophers throughout the years, from Plato to Locke.

owner of a ship, Theseus, gradually replaces all the parts of his ship over a long period of time. However, the mechanic who replaces the parts keeps them all, constructing his own “ship of Theseus” out of the original parts. The question is, whose ship is the real ship? What is the ship’s *identity*?

Naturally, some of you will see these questions as pointless, and dependent upon the answerer. While this is the case to some extent, it is important to see that the different answers to the ship of Theseus thought experiment are metaphors for different understandings of identity, or the self. The point of this thought experiment is *not* to illustrate that diversity of opinion on philosophical topics exists. Indeed, one of the difficulties in this debate is that each of three possible responses to the thought experiment can be backed up by legitimate reasons, giving rise to three different understandings of identity.

The first position one can take with regard to the ship is that it doesn’t have an identity at all. This view coincides with Hume’s and the Buddha’s view that there is no permanent self. In other words, the ship’s physical parts do not constitute its identity, but the idea of it as Theseus’ ship does not constitute identity either. We have already explored the depths of this view. The second view one can take is that the ship on land, the mechanic’s ship, is the real ship. Relating this to people, this view understands the self as being constituted by the physical body. In other words, what makes you the same person from one moment to the next is that your physical body remains the same. That is why the ship in the garage would be seen as the true ship: the physical parts are the same. This second view is called *the substance criterion of identity*. The third view one can take on the thought experiment, and probably most common, is that the ship at sea is the true ship. On this view, the ship at sea retains its identity as an idea, even though its parts are altered. Similarly, according to this view, a person is the same person from one moment to the next even as your body changes (when you age, for example). This third view is called *the continuity criterion of identity*.

The substance criterion is problematic because our bodies do, in fact, change over time: your body is very different now from what it was like when you were ten years old. But if our bodies change so often, then how can our bodies be used to establish our selves as *permanent*? A stronger version of this theory suggests that the physical brain is where identity lies, so even if I put my brain in another's body, it would still be me. But then, what if my brain state could be duplicated in another person's body? Would that mean there are two mes?

The continuity criterion tries to address these problems. According to continuity, our bodies can and do change, because there is something continuous about us that goes beyond our bodies that is the source of a permanent identity. That something might be *memory*. Why are you the same person you were when you were five, even though your body has changed? According to continuity, you are the same person because you have memory links leading back to that time. However, memory runs into problems when we consider memory loss. Do people with memory loss not have an identity, a self? The movie *Memento* touches on this issue: the main character can only remember the last few seconds of his life and has to constantly relearn everything, and yet he still seems to have an identity. Another version of continuity focuses not on memory but on our *stream of consciousness*. Our conscious stream seems to continue through memory changes *and* bodily changes.

But then, consider the following thought experiment. Imagine there is a machine that can duplicate physical things perfectly in a new location, including humans. But after it duplicates something, the original thing is destroyed. Let's say that you step into the machine and you are duplicated, the body you stepped in with being destroyed. Because the new body was created exactly at the point when the old one was destroyed, your stream of consciousness in the new body would be continuous with the old. So are you the same person? What about if the machine made *two* duplications of your original body?

In the end, there's no doubt that this is an abstract philosophical debate. Nevertheless, it sometimes comes up in court cases, such as when a defendant seeks to excuse his behavior for "not being himself." How does our identity factor into our behavior? Additionally, the question of what we are, our identity, is a core question in Eastern religions like Buddhism.

Causality as Habit

Hume's criticism of science as it was developing during the enlightenment period focused on the idea of cause and effect. We saw previously that Aristotle had a whole theory of causality that centered on telos, or purpose. However our current, more basic, idea of cause and effect was being developed during Hume's era. The basic idea was that there are particular natural processes that are built into our universe that are so strongly correlated that we can convincingly say that one caused the other. Every time I apply the right amount of heat to a piece of wood, it catches fire. I can therefore say that the heat caused the fire to burn. Seems simple, right?

Not for Hume. For Hume, when we say that one thing causes another, we are overestimating the strength of the relationship between those two things. There is not a necessary, or deductive, connection between cause and effect, but a contingent, or inductive, connection. In other words, for Hume, causality is a prediction we make, and not built into the structure of the universe. We infer causality based on *habit or custom*. In other words, the only reason we think that sufficient heat causes wood to burn is because we have observed it burning in the past. There is no necessary connection between the two events. Today, Hume's argument about causality has helped us understand the limitations of inductive reasoning—it can only ever give us 99% truth, never 100%.