Physicalism

Welcome to Physics and Religion, a series of podcasts exploring the interaction between modern science and traditional religion, mostly Christianity. I am your host, Dr. A. S. We have been dealing with the legacy of Rene Descartes. Descartes held that there were two kinds of "substance," mind and body. For obvious reasons this point of view is called dualism. The mind can exist outside the body and has no physical properties such as mass and location; it can think the body cannot. The mind can control the body, but the body can also control the mind as for example when one does something out of passion. This seemed plausible at first but we were at a loss to explain how something completely non-physical could interact with something purely physical. We are forced back, at least for the time being, to monism, the idea that there is only one kind of substance. So what will it be, mind or body? I would like to start with the less promising of the two; mind -- nothing physical exists. Reality is purely mind, in fact, my mind! My mind is all that exists; the physical world is pure illusion. You are only an illusion and all the evidence I have of your existence is illusory. This is a philosophy called solipsism. (Bertrand Russel told of a woman who wrote to him announcing that she was a solipsist and wondered why there were not more people like her.) Can you refute solipsism? I think not. Any evidence you might bring to bear on the subject is pure illusion. So why are you not a solipsist? (I presume you aren't.) If the theory cannot be proved wrong, then it must be true – is that not so? I would like to assert an important philosophical principle: if a theory cannot be proved wrong under any circumstances, then for the same reason, it cannot be taken seriously. The most powerful theories are vulnerable. They constantly submit themselves to disproof. The more often they survive meaningful attempts to refute them, the more we consider them relevant and true. We will have recourse to this principle in coming lectures.

If we cannot take seriously the proposition that the one substance is mind, then we have to settle with the proposition that everything is physical. Among philosophers, the notion that "everything is physical" is called physicalism. Like many things in philosophy, this seems clear and intuitive enough until one starts to think about it. In fact, many alternative definitions of physicalism have been

proposed and disputed, and the disputes involve a host of subtle technical issues. One way to narrow the field is to define reductive physicalism, which holds that all phenomena can be "reduced" to physics. (This leaves open the possibility of non-reductive physicalism but this is widely considered a logically inconsistent position.) Philosophers will now quibble about the exact meaning of the word "reduced," but the following example should help. Quantum mechanics in principle gives a complete description of atoms and their interactions. Their interactions include binding with other atoms to make molecules. In this sense, chemistry is "reduced" to physics. The physicalist is committed to the notion that in the same way, biology can be reduced to chemistry and perhaps even that psychology and sociology can be reduced to biology. Certainly, consciousness can be reduced to neuro-biology.

An alternative point of view is called emergentism, the belief that there are emergent properties. A property of a system is said to be emergent if it is a new outcome of some other properties of the system and their interactions, while it is itself unexpected and different from them. Emergent properties are not identical with, reducible to, or deducible from the other properties. Take the example of snowflakes. It is said that no two are identical. It is obviously out of the question to predict the structure of every flake that has ever fallen. On the other hand, one could deduce from the molecular structure of water that it would crystalize in a hexagonal structure. In that sense the phenomenon of snowflakes is reducible to the properties of water. But is the same thing true of DNA? Given the properties of carbon, oxygen, and nitrogen atoms, could one anticipate the double helix? Beyond that, could one anticipate the emergence of life? The standard emergentist answer to those questions is probably not, and certainly human consciousness is an emergent property not reducible to the physiology of the brain.

Emergentists come in several varieties, however. One issue has to do with downward causation. Assume for the moment that DNA is an emergent phenomenon. The properties of the constituent atoms certainly have an effect on the structure of the DNA molecule, but is the reverse true? Do mitosis and meiosis affect the properties of carbon atoms? Obviously not. There is upward

causation, causation from lower-level systems to upper-level systems but not the reverse. With consciousness the question is trickier. Do our conscious lives affect the structure of our brains or is consciousness sort of a useless adjunct to our mental functions. The doctrine that there is no downward causation in this sense is called epiphenomenalism. The usual examples that philosophers give to illustrate epiphenomena are the rainbow that forms over a waterfall and the steam that comes out of a train whistle. The rainbow has no effect on the waterfall and the steam has no effect on the train. Is it true that our consciousness has no effect on the physiology of our brains? The position at first sight seems to be untenable. For example, I am now thinking about lunch. Later today I will remember thinking about lunch. This memory must be stored somewhere in the biochemistry of my synapses; how could it be otherwise? The committed epiphenomenalist would reply that I am confusing cause and effect. My brain cells were thinking about lunch and storing the memories. All this just happened to bubble up into my consciousness. A stronger argument can be made from the existence of esthetic pleasure. We can enjoy the sight of a beautiful landscape or the sound of a great symphony or the words of a poem heard in your head. All this involves conscious perception; without consciousness there would be nothing left. Even this argument is not conclusive. The validity of epiphenomenalism has been strenuously debated with a wide range of subtle arguments. There are no knock-down arguments for either camp but my sense is that it is like solipsism in the sense that it is logically constituted in such a way that it cannot be refuted and thus cannot be taken seriously.

A second issue that divides emergentists has to do with the matter of supervenience. This is a difficult and controversial term, but one definition accessible to the non-philosopher is this: A supervenes on B if it is the case that once God created a B, A would follow automatically. In fact, she couldn't do anything to *prevent* A. If consciousness does not supervene on neuro-biology, or to put it figuratively, if the mind doesn't supervene on the brain, there must be something about the nature of reality that is over and above the laws of physics and neuroscience. With this in mind, let's look at some of the arguments that

have been advanced against physicalism and see if there is some room for the "over and above."

The first argument has to do with zombies. We discussed zombies in connection with dualism. Let us bring them back from the dead and see how they reflect on physicalism. Your philosophical zombie would look like you in every respect. It would be identical to you in every atom of every cell in your body, except that it has no consciousness. To the extent that such a creature is logically conceivable, there must be something over and above physiology that is responsible for consciousness.

The next problem for physicalism is called the inverted spectrum argument, but I'm sure you can construct many different versions of the same argument. Suppose there is someone, not a zombie this time but an ordinary human being, whose color receptors are inverted in such a way that what we see as red he/she sees as green, etc. There is no way we could tell that this person was in any way different from the rest of us. We would look at a rose and say this is a red rose. Our hypothetical color-inverted person would agree, "Yes," she would say, "this is a red rose." Of course, it looks green to her but she has learned that the color of the rose is named "red." There is no way we could ever resolve the difference. So far as I know, my color spectrum is inverted relative to yours. This means there must be something to conscious perception that is beyond the range of physics-based knowledge.

Welcome to Physics and Religion, a series of podcasts exploring the interaction between modern science and traditional religion, mostly Christianity. I am your host, Dr. A. S. We have been discussing problems with physicalism, this notion that everything is physical. I told you about zombies and explained the inverted

spectrum argument. This brings us to what many people regard as the central problem of consciousness, the problem of qualia. Qualia are primary sense experiences, phenomenal experience the philosophers would say. Things like redness, pain, itching, the smell of an onion, the sound of a violin, or the touch of a caress. A famous thought experiment highlights the problem. It's called the problem of black and white Mary. Mary was a normal child in all respects except that there was something defective in her visual apparatus so that she had no color vision. Perhaps here cone photoreceptors were not hooked up correctly. During her deprived childhood she became an expert in neuroscience. In particular, she learned all about color vision. She knew all about the photoreceptors and ganglion cells. She was expert in matters pertaining to the lateral geniculate nucleus and the primary visual cortex, but she never actually saw color. Eventually a surgical procedure was developed that repaired her defective vision. When she awoke from the surgery, for the first time in her life, she saw RED! Nothing she could ever have learned could have prepared her for this experience. The experience of RED simply could not be explained in purely physical terms. The same argument could be made of course for any of the other qualia I mentioned above.

Bats pose a similar problem. Bats can see, but since they spend part of their lives in dark caves, they have become expert in echolocation. They "see" not with reflected light as we do but with reflected sound; but there is something about them that is still more remarkable. Mother bats raise their babies in dark caves with as many as a million other bats; but when mommy bat returns from a hunting trip with food for her offspring, she can find them immediately by the sound of their cries in a cave with a million other screaming bats. So the most famous and most influential question in the study of consciousness was raised by Thomas Nagel years ago, "What is it like to be a bat?" To which I would add, "And how could we possibly know?"

Then there is the problem of the self. I have a clear memory of standing up in my crib when I was an infant. I have learned a lot since then, and I am dragging around a big clumsy body. I suppose that every atom in my body has been replaced several times. And yet – It's still me! I am the same person I was then. It

is as if I had changed clothes over and over. I might have played different roles depending on the clothes I was wearing, but the person inside the clothes has always been the same. Where in your brain, in what neurological structure is your personal identity located?

Then there is the binding problem. At this moment I am looking at the computer, I hear the cat howling for some reason, I feel the touch of the keyboard, and I feel a vague itching in my scalp. I have no sense of smell, but if I did, I would probably smell the last of dinner in the air. All these elements of sense data are processed in different areas of the brain, and yet they are presented to my awareness as one unified experience. It is as if there were a central control room where all these sights and sensations were synchronized. If there is such a thing, we have no idea where it is or how it works.

The final problem is called the explanatory gap. Let me put it his way – I challenge you to come up with a wrong theory to explain consciousness! Come on now – it shouldn't be too hard – you don't have to come up with the right theory. There should be an infinite number of wrong theories for every correct one. But whenever I try to do this I find myself standing helplessly on the edge of this gap. I want to use grammar to explain the problem, but unfortunately, English grammar is defective in one important respect, as you will see. In English we say "I am going" or "You are going" or "It is going." *Am, are, and is* are the first-, second-, and third-person singular forms of the verb *to be.* The problem is that science works exclusively in the third-person perspective. It deals with things that can be studied, results that can be intersubjectively verified. The objects of scientific investigation can be properly called "its." Consciousness on the other hand, can only be experienced "from the inside." Only you have access to your own consciousness. It can only be sensed from the "I" perspective. This is the explanatory gap, and it is hard to see how it could ever be bridged. Except – the French have another verb form that is missing in English. It's called the *tu* form, as in "Tu vas." We have to translate this "You are going," but in French this is the language of love as in "Je t'aime," I love you. So the French have it right. The way to bridge the first- and third-person perspectives is with the language of love; but that's not very scientific either!

It seems to me that physicalism as I described it the opening paragraph is an untenable position. If so, what are the alternatives? We could simply deny that consciousness exists. This is the position taken by the philosopher Daniel Denett. This is not widely accepted, and some the arguments he has had have been quite acrimonious. I don't understand the position and I won't try to defend it. Alternatively, we could say that we can't understand consciousness and that's that. It's what philosophers call a brute fact. Physics is full of brute facts: the charge of an electron, the speed of light, Planck's constant, etc. These numbers are what they are. We have no idea why. Perhaps the nature of consciousness is just such a fact. This point of view is called mysterianism. There are several variations on this theme, however. It may be that we can't understand consciousness because our brains have simply not evolved to answer such questions. Our brains evolved to help us get food and avoid predators. Understanding consciousness had no survival value! I have a different idea, which might even be original. In the Copenhagen interpretation of quantum mechanics the concept of the meaningless question looms large. According to this interpretation, physical reality if so constituted that certain questions just don't make sense. This notion will be very prominent when we talk about free will in the next lecture.