Metaphysics

Course notes by Richard Baron

This document is available at www.rbphilo.com/coursenotes

Contents

Version of January 2011

Introduction to metaphysics

The name comes from the Greek words *meta*, after, and *physica*, natural things. So the word means "the things that come after nature". The meaning is "after in the order of study".

The word has come to mean "the study of fundamental aspects of the world". A recurring theme is that of whether the world actually is like we think it is.

The subject has close connections with epistemology (the theory of knowledge), with the philosophy of mind and with the philosophy of science. In this course, we will branch out into the philosophy of science.

Reading

This handout can be used on its own. But if you would like to do some extra reading, there is plenty of material on all of the topics of metaphysics. A good place to start is the articles in:

The Stanford Encyclopedia of Philosophy, http://plato.stanford.edu

The Internet Encyclopedia of Philosophy, http://www.iep.utm.edu

What exists

The study of what exists is called ontology, from the Greek word for things that exist, *onta*.

Criteria we might use to decide whether things of a particular type exist

Are they detectable by our senses?

Do they play a useful role in our theories?

Can we identify their properties?

Can we imagine the world without those things?

Questions for discussion

What are the pros and cons of these criteria?

Is intuition more useful than criteria when deciding what exists?

Things that might exist

Physical objects of medium size
Sub-atomic particles
The force of gravity
Space and time, or space-time

Unicorns

Blueness

Emotions

Countries
Laws
Justice

The world economy
The gross domestic product of a country

The Industrial Revolution

Numbers
Pythagoras's theorem
The round square

Questions for discussion

How do these types of thing fare against the criteria we identified?

If bricks exist and numbers exist, do they exist in the same sense? If not, what do the two senses of existence have in common?

Things that must exist

There are different senses in which something might have to exist.

> It might be a logical impossibility for something not to exist. Are numbers like that?
>
> The Universe might have to include something, in order to be even remotely like our Universe. Space-time might be necessary in this sense.
>
> We might be unable to grasp the possibility of a Universe that did not include something. Space and time (or space-time if we think like physicists) might be necessary in this sense.
>
> Something might have to exist, in order for us to be here to worry about the question. Certain kinds of chemical structure might be necessary in this sense.
>
> Something might not have to exist: the Universe could have been quite similar to human eyes without it. But it may be such that if it does exist, it does not exist as the result of some physical laws or processes that might have turned out differently. That would give its existence a kind of necessity, if it exists at all. God might be like that.
>
> Something might have necessary existence built into its concept. Again, God might be like that.

We could object that the concept of necessary existence makes no sense. We can always think of something, and then think of its existing, or not existing. So existence cannot follow from the concept. But would this only apply to physical things? Would it, for example, apply to numbers?

Questions for discussion

Does the idea of necessary existence make sense?

If we cannot imagine something's not existing, how close does that come to saying that it must exist?

The properties of objects

Identifying and defining properties

An object may have lots of different properties.

> A piece of paper may be A4 sized, blue, light blue, useful for making notes, … .
>
> Susan may be tall, clever, ten metres south of Barbara, … .
>
> A mathematical problem may be difficult, interesting, significant, … .
>
> A government may be strong, unacceptable, … .

We may classify properties as intrinsic (in the object itself, such as size) or as extrinsic (dependent on a relationship to something else, such as tall for a human being, or interesting to human beings). Some extrinsic properties, like the property of being ten metres south of Barbara, look as though they have very little to do with the object itself. They could change, with no change in the object that anyone interested in the object would be likely to care about.

How can we define properties? We can give dictionary definitions, but they define properties in terms of other properties. How do we link our words for properties to reality? If we point to things and utter words, how do we know that we are latching on to the same properties when we agree on our practices of pointing and saying? When we all point and say "blue", how do we know that we are all latching on to the colour?

We might try to define properties as sets of objects that have them. But the set of animals with hearts and the set of animals with kidneys could be the same, without the properties "having a heart" and "having kidneys" being the same property.

Questions for discussion

Should we count as many properties as we can think of as real properties, or should we set some limits? For example, is a particular shade of blue a real property of a piece of paper, and "blue" an artificial composite that means "blue shade 1 or blue shade 2 or blue shade 3 or …"?

If we are going to set limits, as in the preceding question, how should we fix those limits?

Can extrinsic properties be translated into sets of intrinsic properties (for example, tall for a human being, translated into a height in metres and a statement of the height in metres that only one quarter of human beings exceed)?

Are there any genuinely intrinsic properties? (Being two metres tall, for example, seems to be related to the length of the metre.)

How should we classify properties that look intrinsic but that normally only show up in interactions with other objects, for example, magnetism?

Primary and secondary qualities

The distinction between primary and secondary qualities is particularly associated with John Locke (1632 – 1704), although it goes back further than him.

> Primary qualities are picked out without any reference to observation, not even implicit reference. Examples are the properties of having a mass of 15 kg, and of being spherical.
>
> Secondary qualities are picked out by the effects they have on observers. Examples are the properties of being blue, and of sounding discordant.

There does not seem to be much scope to disagree on whether an object has given primary qualities. If two people do disagree, the way to settle the question is to go back and take measurements more carefully.

Secondary qualities are different. Things can appear to have different colours in different lights, or to different people. Even people who do not have any identifiable lack of colour vision can disagree, for example as to whether something is yellow or green (assuming it is near the borderline). And for someone who lacks a sense, such as sight or hearing, a whole range of secondary qualities can seem not to be there at all.

Questions for discussion

Could we get rid of the problem of things appearing to have different secondary qualities to different people, or under different conditions, by assuming a standard person and standard conditions? How would we define the standard person and conditions?

Would we lose anything if we stopped talking in terms of secondary qualities, and started saying things like "reflects light of wavelength around 530 nm" instead of saying "is green"?

Is shape really primary, or is it relative to our way of perceiving the physical world, as made up of objects in space? Could there be creatures who did not perceive the world in spatial terms, and who were therefore unaware of shape?

Supervenience

Two sets of properties can be related in the following way: one is supervenient, and the other is subvenient.

This means that an object cannot have different properties drawn from the supervenient set, unless it has different properties drawn from the subvenient set.

Example: inherited properties of plants

> Supervenient set: resistant to disease X, resistant to disease Y, resistant to disease Z, … .
>
> Subvenient set: having genome P, having genome Q, having genome R, … .
>
> A plant cannot have different inherited resistances to diseases unless there is something different in its genome, its complete sequence of genes. It may be that only a very small difference in the genome makes a big difference to resistance to diseases, but there must be some difference or other. (We do need to overlook environmental influences that may switch some genes on or off in order to make this example work, or incorporate environmental conditions into the subvenient set.)
>
> Note that when there is supervenience, the requirement of difference does not have to go both ways. Plenty of differences in genomes do not imply any difference in resistance to diseases.

Example: mental and physical properties

> Supervenient set: thinking that it is sunny, thinking that philosophy is wonderful, … .
>
> Subvenient set: state of brain F, state of brain G, … .
>
> It looks as though a particular person, at a given time, in a given environment, and with a given history, cannot have a different mental state unless there is something different in the state of his or her brain.
>
> Even if that is so, it does not mean that our minds are identical to our brains, or that our thoughts are identical to patterns of activity in our brain cells. They might be, but it would not follow from supervenience alone.
>
> Again, the requirement of difference does not have to go both ways. Several different brain states could correspond to the same thought.

Question for discussion

Could we do without talk in terms of supervenient properties, and just stick to talking about the subvenient properties?

Emergence

Some properties arise in large, complex objects, but not in their parts separately. For example, an assembly of brain cells, or the mind to which they give rise, can be conscious, but no one brain cell is conscious. Consciousness seem to emerge as we move up to the more complex entity. Conscious states have higher-level properties that are set out in the terms of psychology. Brain cells only have lower-level properties that can be set out in the terms of physics, chemistry and cell biology.

Emergentists make claims like the following. (Different emergentists make different claims.)

> The higher-level properties could not be predicted from the lower-level properties. For example, we could not predict consciousness from knowledge of the chemistry and biology of brain cells, even if we considered the chemistry and biology of large assemblies of brain cells.
>
> There are laws governing relationships between higher-level properties that could not be replaced by laws governing the lower-level properties. For example, noticing that the sun is shining might be followed by happy thoughts, without there being a matching law in chemical or biological terms about one state of a brain leading to another state of the same brain.
>
> The occurrence of one higher-level property can cause the occurrence of another higher-level property. Noticing that the sun is shining actually causes happy thoughts, rather than the causal processes being at the level of brain cells.
>
> We, with our limited capacities, could not understand the world unless we thought in terms of the higher-level properties. A picture of the world in terms of lower-level properties would be too complex for us, and we could not make predictions about the complex entities.
>
> No-one, however great their intellectual powers, could understand the world properly without thinking in terms of the higher-level properties. For example, we can only understand how people behave if we think in terms of their beliefs and desires.

Questions for discussion

Which everyday properties of people, animals or objects might be emergent?

Is it true that we could not predict consciousness from a full understanding of brain cells and how they are put together in a brain?

Are there laws that link higher-level properties directly, or just rough generalizations that are effectively summaries of a lot of lower-level detail?

Should we rule out occurrences of higher-level properties causing occurrences of other higher-level properties, on the ground that there is enough causal action at lower levels to explain what happens?

Causation

Causal necessity

We say that one event causes another. A cricket ball's hitting a window causes the window to break. A person's pressing a switch causes a light to come on. And so on. We have the impression that the first event necessitates the second one – it forces the second one to occur. At least, if the second event does not occur, there will be some special explanation – a sudden gust of wind that lessened the impact of the ball, or a defect in the light switch.

But all we ever see are the two events. We do not see a force of necessity. According to David Hume (1711 – 1776), seeing the same sequence of events over and over again gives us the impression that the first event necessitates the second, but that is all. We could always conceive the first event happening, but not the second.

Questions for discussion

Is it true that we don't actually see causal necessity? We do say things like "I saw the cricket ball break the window", which seems to imply seeing more than just one event followed by another.

Can we say that causal necessity is established by detailed laws of physics? Pressing a switch causes a light to come on because the switch is solid, so the pressure of a finger causes a piece of metal inside the switch to move, which completes a circuit, and the voltage ensures that electrons move round the circuit, … . Or does that just push the problem down to the necessity in those laws?

Do we really get our idea of causal necessity from our experience of acting and making things happen in the world? (I decide to move a book from my desk to the shelf, I move my hand, and the world is forced to obey me – the book moves.) If we do get the idea from there, can it justify our assumption that causal necessity exists?

Can we imagine a world without causal necessity? Would mere regularities, which could cease to hold at any moment, be enough to make the world comprehensible?

Analysing causal claims

One way to analyse causal claims is to use counterfactuals: statements of what would have happened if things had been different. We might say that a ball's impact was the cause of a window's breaking, because if the ball had not struck the window, the window would not have broken.

One problem for this analysis is over-determination. For example, the ball might have struck the window just as an aircraft's sonic boom reached the ground. Either would have broken the window on its own. So without the ball, the window would still have broken. And without the aircraft, the window would still have broken.

We would also like an analysis that takes account of the importance of background conditions. A ball will only cause a window to break if the glass is not unusually tough and the wind does not lessen the impact of the ball too much. Perhaps we cannot simply say that the window broke because the ball hit it. Perhaps we should say that it broke because the ball hit it and it was weak enough and the wind did not intervene.

Questions for discussion

What should we do about over-determination? Is each of the events that would have led to the effect on its own a cause? Or should we say that they combine to form a single cause? Or is the concept of an identifiable cause in trouble?

If the counterfactual analysis only runs into trouble in special circumstances, can we rescue it by saying that it applies most of the time but needs amending in those special circumstances? What definition of special circumstances might we use?

Are background conditions causes? If not, why not?

The causal closure of the physical

A sufficient cause of an event is a cause which means that the event is bound to occur (perhaps against some background assumptions, for example that the usual laws of nature apply).

The causal closure of the physical is the principle that any physical event that has a sufficient cause, has a sufficient physical cause. Not everyone accepts this principle, but a lot of philosophers do.

If the principle is correct, then when we can find full causal explanations of physical events, they won't have to include non-physical causes. We won't, for example, have to identify thoughts as causes, although we might have to identify electro-chemical events in brains as causes.

Questions for discussion

Does the principle of the causal closure of the physical look plausible?

Is the principle well-defined? How can we say what counts as the physical?

Could the natural sciences ever show that the principle was mistaken, or would they miss any evidence that it was mistaken because they would treat anything they looked at as physical?

Should we accept a principle of the causal closure of the mental, for mental events like thoughts? That is, does every mental event with a sufficient cause have a sufficient mental cause? Are there any counter-examples that should lead us to reject this principle?

Determinism

Suppose that the principle of the causal closure of the physical is correct, and suppose that in addition, every physical event does have a sufficient cause. (Quantum physics, as it currently stands, would tell us that this was not so. Some physical events just happen, for no reason. But we will ignore that point for the moment, to see where the argument goes. And in any case, physics might change.)

It seems that in that case, the Universe would be physically deterministic. That is, if the Universe had a certain initial physical state, specified in complete physical detail (every particle's place and speed and direction of travel, every little bit of energy, and so on), then everything physical that happened from that point on would be inevitable. The argument would run as follows.

> The initial state would give all the physical detail of the Universe.
>
> It would therefore give all possible physical causes.
>
> Our assumptions mean that any given physical event at the next stage would have a sufficient physical cause.
>
> But that cause would have to lie in the initial state, so it would definitely be there.
>
> And because the cause was sufficient, it would ensure that the event occurred.
>
> So all physical events in the first round after the initial state would be inevitable.
>
> We could then repeat the argument for the next round of events, and so on throughout the history of the Universe. The sufficient causes of physical events in later rounds might lie in the immediately preceding rounds, or further back, but those causes would still ensure that the later events occurred.

If we could then add that all non-physical events supervened on physical events, for example, that all thoughts supervened on the occurrence of states of brains, then we would conclude that all events, physical and non-physical, were determined.

Questions for discussion

Does this argument for determinism work, given its assumptions?

If the Universe was deterministic, could we tell that it was deterministic?

If the Universe was deterministic and we had incredibly powerful computers and lots of detailed data, could we predict precisely what would happen, and be right every time?

Is determinism scary?

Necessity and possibility

What could have been different?

It looks as though some things could have been different from how they actually are, while others could not. Even if the Universe is deterministic, it seems that some things could have been different, because the initial conditions or the laws of nature could have been different. But it is hard to imagine how 2 + 2 could have equalled anything other than 4.

We can talk about contingent and necessary propositions.

> *Contingent propositions* could have been either true or false: for example, "it snowed in Norway on 17 May 2008" (true, but it could have been false); or "Mozart lived until 1792" (false, but it could have been true). The first of these examples is *contingently true*. The second is *contingently false*.
>
> *Necessary propositions* could not have been false: for example, "2 + 2 = 4".
>
> The negation of a necessary proposition, for example "not-(2 + 2 = 4)", is impossible. We can also say that it is *necessarily false*.

We can also talk about possibility. A proposition is possible if it is any of the following:

> Necessarily true
> Contingently true
> Contingently false

We can define each of necessity and possibility in terms of the other:

> A proposition is necessary if, and only if, its negation is not possible.
>
> A proposition is possible if, and only if, its negation is not necessary.

(We talk about propositions being necessary or possible, as a short way of describing them as necessarily true or possibly true.)

Questions for discussion

Is there a clear boundary between things that could have been different, and things that could not have been different?

If we draw a boundary based on what we can imagine, will the boundary just reflect how good or bad we are at imagining things? Will the position of the boundary vary from century to century, depending on the state of our scientific knowledge?

Possible worlds

Philosophers often handle questions of necessity and possibility by thinking in terms of possible worlds. A possible world is simply a way things could have been. The actual world is itself a possible world.

A possible world is a complete Universe, not just a reproduction of our own planet with some changes. So other possible worlds are not floating around in the space of our own Universe. We cannot reach them by travelling from here, and what goes on in one possible world has no causal influence on what goes on in another possible world.

Although we cannot see other possible worlds with telescopes, we can see them with our mind's eye. We can use them as examples of how things might have been. We can then use them to define necessity, possibility and contingency.

> A proposition is necessary if it is true in every possible world (including the actual world).
>
> A proposition is true if it is true in the actual world.
>
> A proposition is possible if it is true in at least one possible world (perhaps the actual world, perhaps another world).
>
> A proposition is contingent if it is true in at least one possible world and false in at least one possible world.
>
> A proposition is necessarily false if it is false in every possible world.

A whole system of logic, called *modal logic*, has been developed to handle questions of necessity and possibility. It is built around the idea of possible worlds.

Questions for discussion

Some possible worlds would be only a little different from the actual world, and some would be very different. What criteria might we use to get some kind of measure of how different a possible world was from the actual world? How precise could we be? Could we arrange possible worlds on a scale, starting with those that were most similar to the actual world and ending with those that were most different from the actual world?

Consider the claim that there is a possible world in which Mozart lived until 1792. What would be our justification for identifying a particular composer in that world as Mozart? What would be our justification for identifying a particular period of time in that world as 1792?

Types of necessity

A proposition may be necessary in several different ways.

> Logical necessity: the rules of logic require the proposition to be true. An example is, "If some elephants are friendly, then some elephants are friendly".
>
> Conceptual necessity: our concepts require the proposition to be true. An example is, "All triangles have three sides".
>
> Metaphysical necessity: the basic nature of things, properties and so on requires the proposition to be true. An example is, "Plato was a human being". Any creature who was not a human being would not have been Plato, even if that creature had written identical philosophical works.
>
> Physical necessity, also called nomological necessity (from the Greek word *nomos*, a law): the proposition must be true, given the laws of physics. An example is, "You cannot pick up an everyday object and accelerate it to a speed greater than the speed of light".
>
> > We can talk about physical necessity relative to a given set of laws of physics. We are usually most interested in the laws that apply in the actual world. But we may want to ask what would be possible under a different set of laws.

We can also talk about logical, conceptual, metaphysical and physical possibility. A proposition is logically possible if it is not ruled out by the laws of logic, conceptually possible if it is not ruled out by our concepts, metaphysically possible if it is not ruled out by the basic nature of things, properties and so on, and physically possible if it is not ruled out by the laws of physics.

Questions for discussion

Which types of necessity imply which other types? For example, if a proposition is conceptually necessary, is it also logically necessary?

Should any of the types be merged into one another, or be regarded as overlapping? Do any types include any other types? For example, are conceptual and metaphysical necessity the same thing? Or do they overlap? Or is one of them a special case of the other? Or are they really separate?

How could we develop the idea of a range of possible worlds in order to use it to picture different types of necessity?

Would objects be the same objects, even with different properties?

We can consider the properties of an object (including a person), and ask whether it would have been the same object, if its properties had been different. Would this red car have been the same car, if it had been painted blue instead of red? Would this person, brought up in England and speaking English, have been the same person if she had been taken to Japan at the age of one, had only ever spoken Japanese, and had never learnt English?

Most of us would say yes, in both cases. We see properties like the colour of a car, or the languages that someone speaks, as *accidental properties* of the car, or the person. They do not go to the very heart of the car's, or the person's, being.

On the other hand, some properties may be *essential properties*. They could not have been different, if the object was to be the same object. For example, it seems to be an essential property of any given person that he or she is a human being. A creature that was not a human being could not have been that person.

We are not talking about mid-life changes of species here. We are talking about a creature that was, from the beginning, some other species, for example, a bear. The suggestion is that it is metaphysically impossible that any given human being could have been a bear, even though a sufficiently talented bear could have lived at the same address as that person, had the same job, answered the telephone in the same way, and so on.

If this suggestion is correct, then it looks as though statements like "I could have been a bear", and "I would like to have been a bear", would make no sense. A bear would not have been the human being who said those things.

There is scope to argue about which properties are essential. But obvious candidates are as follows:

> for human beings, having the particular parents that one has (the parents who were the source of the sperm cell and the egg, disregarding surrogacy, adoption and so on); being the product of that particular sperm cell and egg; and being born within a reasonable time interval around the actual time of birth (someone born a century before you would not have been you);
>
> for inanimate physical objects, the particular lumps of material from which they are made, the kinds of material from which they are made (such as wood or steel), their overall design (a car of normal design could not have been a table of normal design), and their intended function (a statue of Descartes could not have been a statue of Leibniz);
>
> for natural kinds, their physical or chemical characteristics: gold must have 79 protons in each atom, and each water molecule must include two atoms of hydrogen and one atom of oxygen.

Saul Kripke (born 1940) has argued that we can go on to identify propositions that are both necessarily true and known a posteriori.

> A proposition is known a posteriori if it could be contradicted by experience. This implies that the proposition can only be known after we have gained experience of the world. We cannot come to know it by sitting in our armchairs and thinking.
>
> An example is, "Water has the chemical formula H_2O". We could not know this without experience, but given that it is true, anything that does not have the formula H_2O does not count as water.
>
> Another example is, "Saul Kripke was born after the year 1800". We need experience to check that there is such a person, and when he was born, but anyone born in or before 1800 would not have been him. There might have been someone of the same name born that long ago, and that person might also have been a philosopher, but it would not have been the same person as the Saul Kripke who was born in 1940.
>
> Kripke connects this with possible worlds by introducing the concept of a *rigid designator*. This is a term that designates the same object in all possible worlds in which that object exists. Thus "water" designates H_2O in all possible worlds where there is water, and "Saul Kripke" designates the same person in all possible worlds where he exists. A term gets to be a rigid designator when someone decides to use the term to refer to *that* substance, person or object, and the rest of us accept this decision and act accordingly.
>
> Then in every possible world, "Any water that exists has the chemical formula H_2O" and "If Saul Kripke exists, he was born after the year 1800" are true. They therefore qualify as necessary truths. In the actual world, both water and Saul Kripke exist, so the conditions "Any water that exists" and "If Saul Kripke exists" drop away. We are then left with "Water has the chemical formula H_2O", and "Saul Kripke was born after the year 1800".

Questions for discussion

What are the essential properties of amoebae (which reproduce by copying their DNA and then splitting), of pieces of music, and of computer programs?

Is there a clear boundary between essential properties and accidental properties?

Would an object have been the same object if most of its accidental properties had been different?

How should we think about someone in another possible world, who speaks a language just like English, and who uses the word "water" to refer to a colourless liquid that falls from the sky, fills the rivers, comes out of taps, and is drunk, but that has a different chemical formula from H_2O?

Can we just decree that a term will be a rigid designator, that it will pick out the same object in all possible worlds? Or is there a problem, because "I mean *that* thing" involves the speaker's pointing in the actual world, not in other worlds?

Scientific theories

The problem of scientific realism

We have scientific theories that are very successful. They refer to some unobservable entities, like electrons and the gravitational field. We can detect the effects of these things, but we cannot perceive them directly. So are these things really there, or are they just things that the theorists invent, in order to tell stories that correctly predict the events we can observe? (The events we observe include the occurrence of pictures on computer screens, and large and small stones falling from the top of a tower to the ground in the same length of time.)

Realism

Realism about scientific theories is the view that the unobservable entities are real. There really are electrons, there really is a gravitational field, and so on. There are several arguments for this position, but there are also challenges to the arguments.

> Realism is just common sense. We talk about chairs because we sit on them, bump into them and so on. It is obvious that they are real. The same should go for anything that we talk about, when our experience is what we would expect if the things we talked about were real.
>
> > Challenge: we have no right to apply common sense in the world of physical theory, where we discuss, in mathematical terms, a world that we experience in a form that is very different from what the mathematics would suggest. We experience a world of solid, stable objects. Physical theory does not show us a world like that. Electrons are not like very small billiard balls.
>
> The no miracles argument (Hilary Putnam, born 1926). Scientific theory is incredibly successful. It would be a miracle if it were that successful, without telling us what the world was really like. We should not believe in miracles. So the theory must tell us what the world is really like. This implies that the entities it talks about, are real.
>
> > Challenge: old theories, which have now been rejected, worked very well too. Newtonian physics presupposed absolute space and time. The caloric theory of heat explained heat by reference to a special fluid, called caloric. Now we know that absolute space and time, and caloric, do not exist. But if the no miracles argument is sound, they should exist.
>
> Realism gives an account of how our whole body of theories holds together. Physics is interwoven with chemistry, and chemistry with biology. If the entities that physics discusses are real, they can play their roles in making chemical reactions, and plants and animals, work in certain ways. But if they are just the imaginings of physicists, how can they do anything at those higher levels? In short, reality makes more sense if it is real all the way down.
>
> > Challenge: realism is about entities that appear in our theories. We cannot justify something that specific, on the basis of this argument about links between theories. Something that goes on at the level of physics could be real enough to explain chemistry and biology, without the specific entities that appear in our theories being real.

Experimental practice assumes realism. We act on things. How could we act on them, if they were not real? Ian Hacking (born 1936) gives the example of electrons being fired at a ball of niobium. He says, "If you can spray them then they are real" (*Representing and Intervening*, page 23).

> Challenge: we don't flick the electrons with our fingers. We set up complicated apparatus, which according to our theory will produce a stream of electrons and send them in a certain direction. So we can only say that we are really spraying electrons if we have already established that realism is correct.

Question for discussion

In what ways could the realist respond to the challenges given above?

Anti-realism

The anti-realist about scientific theories agrees that unobservable entities play crucial roles in theories that make accurate predictions about what we will observe and about the consequences of our actions. But the anti-realist denies that these entities are real. When we use words like "electron" or "gravitational field", we talk in a way that makes our theories easy to express and to use, but we do not refer to real things.

The arguments for anti-realism include the challenges to arguments for realism that were set out above. But there are some additional arguments for anti-realism, which we will consider now. These arguments are, in turn, open to challenges from the realist side.

Theories are under-determined. The only real constraint is that our theories must fit the experimental data. But lots of different theories could fit the same data. Different theories would involve different unobserved entities. Why should we assume that the theories we use identify entities that are real, and that alternative theories only identify imaginary entities?

> Challenge: not all possible theories are equal. We can prefer the theories that are simpler, more powerful, better integrated with existing theories, and so on. Such preferences narrow the field, perhaps to just one theory.

Anti-realism can easily handle changes in theories, including radical changes. For example, when relativity and quantum mechanics were introduced, they made very small differences to the observations we would expect in everyday life, but massive differences to our understanding of the nature of the Universe. If our theories are no more than adequate accounts of what to expect, rather than being statements of the true nature of things, that explains how we can make radical changes without feeling that we have lost touch with the world. Theory change is like learning a new language to describe the same world. It is not like being transported to a new planet, where everything is different and we feel completely lost.

> Challenge: that understates the significance of radical changes in theories. Such changes don't just show us that we can improve our language. They show us that the Universe really was different from how we thought it was.

Anti-realism does not present a false picture of what goes on in the natural sciences. Experimental practice, and the interpretation of observations, depend on current theory. There is a two-way process, in which theory and experiment influence each other, and our theories steer what we find. Realism, on the other hand, suggests that there is a fixed world out there, which actually contains the entities that our theories mention. Realism suggests that we simply look at those entities, and try to get a clearer and clearer view of them.

Challenge: this is unfair to realism. Yes, using our current theories, we do think that we are trying to find out more about certain particles and fields, and we choose our experiments accordingly. Indeed, we think about our experiments in those terms: "How can we use these new instruments to discover the precise value of … ?". But feedback from experiments could lead us to change our theories. We could decide that those particles and fields were not there after all. Then we would try to find out about other things.

Question for discussion

In what ways could the anti-realist respond to the challenges given above?

Structural realism

Structural realism about scientific theories is the view that while the unobservable entities may or may not be real, the structures that our theories capture are real. They are out there, and we can find out more and more about them.

For example, we can find out that when we compress a quantity of gas into a smaller volume, its temperature rises. And when we heat a quantity of gas that is trapped in a fixed volume, the pressure rises. We can discover all that, and build our theory of gases, without knowing anything about molecules and how they behave in gases. Then we can discover that molecules fly around in a gas, and that they collide with one another and with the walls of their container. At higher temperatures, they move faster. We get a whole new theory of how gases behave, and we start to talk about different things, like the energy of molecules, but the pattern of mathematical relationships that we had in the relationships between pressure, temperature and volume is preserved. So the structures are real, and they endure when we change our theories, even though the structures may get expressed in new ways.

There are two broad types of structural realism. Epistemic structural realism claims that we can only know the structures. Ontic structural realism claims that structures are either the only reality, or the primary reality. (An example of a claim of primary reality is a claim that while electrons and quarks may be real, their reality is only derived. The basic reality is the mathematical structure that is reflected in our table of all the particles mentioned in our theory. Individual particles get their reality from having the relationships to one another that are given by the mathematical structure.)

Here, we will look at arguments for ontic structural realism, and some challenges to those arguments.

Structural realism explains the success of science, without reference to miracles. If the structures are real, we should succeed when we have identified those structures.

> Challenge: when we use our science, for example in new technologies, we act on the things that are identified by our theories, not on their structures. So we can only explain our success if we accept that we have identified the things that are really there.

Structural realism can handle radical changes in theories. The structures endure from the old theories to the new ones. This explains why old theories and new ones tend to make very similar predictions about what we observe in the kinds of situation we have observed so far: observational consequences are governed by the same enduring structures. But it also means that after we have adopted a new theory, we don't have to say "We thought the old structures were real, but they obviously weren't, so we have no evidence that the structures that are implied by our new theory are real either".

> Challenge: we need to say which structures we are talking about here. If we are talking about the patterns in observations, and we just re-label the things that we measure, or make other simple changes, when we change our theories, those patterns are not exactly deep. They won't explain the success of our theories. If we are talking about deeper structures, then it is not clear that they remain unchanged when we change our theories. The space-time of relativity, for example, reflects a mathematical structure that is not just a matter of sticking space and time together. Philosophers have argued about examples, and have found some unchanged deep structures, but it does not seem acceptable to support structural realism with a "win some, lose some" scorecard. We might get continuity of deep structures nearly every time if we defined the structures in very general terms, but then we would not say enough about the structures involved to explain the success of science.

Mathematics is the key to scientific understanding in physics, in a large part of chemistry and in a fair amount of biology. And mathematics is all about structure. So we should expect to express the content of our sciences in terms of structure.

> Challenge: mathematics is indeed all about structure. Structures that have the same shape, get treated the same in mathematics. We cannot tell them apart mathematically. So if our science is only about structure, we cannot tell what kind of universe we really inhabit. It might be any one of a wide variety of universes, all having the same structure, but made up of different things. Can't we do better than that?

Question for discussion

In what ways could the structural realist respond to the challenges given above?

Physical laws

We have lots of physical laws, for example the law that when a metal bar gets hotter, it gets longer, and the law that hydrogen will burn in oxygen to form water. But what makes a law special? Here are some options, and some objections to them.

> A law is a true generalization about the whole Universe.
>
> > Objection: it may very well be true that all life came into being after the time that we would call 5 billion BC. That would not make it a law. Stars and planets could have formed early enough to have allowed life before that time.
>
> A law is a generalization about the whole Universe that has to be true.
>
> > Objection: it has to be true that all life came into being after the time that we would call 15 billion BC, because the Universe is not that old. That would not make it a law.
>
> A law is a generalization that has to be true always and everywhere, but not just by virtue of accidental facts. It is a law of nature that no life form can survive a temperature of 1 million degrees centigrade, because complicated molecules would disintegrate at that temperature.
>
> > Objection: how can we distinguish accidental facts from non-accidental facts? We cannot just say that a fact is non-accidental if it is a law of nature. That would leave us with a circular definition of “law of nature”.
>
> A law is a generalization that is true always and everywhere, and that we can test by deliberately setting up new conditions and seeing whether the outcome is what the law predicts.
>
> > Objection: we can do experiments, but the results are not often exactly what we expect. They are usually close, but this suggests that our laws are not strictly speaking true.
>
> A law is a generalization that forms part of a whole system of generalizations, with logical relationships between them, and the system as a whole does well in predicting the results of a wide range of experiments.
>
> > Objection: moving up to a whole system just means that we have a single big, complex generalization instead of lots of small, simple generalizations. We have to say what makes it a system of laws, rather than a system of statements that just happen to be true.

Questions for discussion

Which of the objections above can be answered, and how?

In what sense are our laws true, given that we do not often get the ideal conditions they presuppose? For example, we have the law that an object moving outside the influence of any forces will just carry on at the same speed. But we have never observed, and do not expect to observe, an object that is not influenced by any forces.

Human action

Agent causation

A human action, such as turning on a light, answering the telephone or climbing a mountain, is a peculiar thing. From the outside, it may look like a complicated physical event. A person has certain states of his or her brain, and receives sensory inputs. The combination triggers reactions in brain cells, in nerves and in muscles. As a result, limbs move in certain ways. From the inside, it does not feel like a complicated physical event, or at least not just like that. It feels as though the agent stands at the start of a causal chain, when he or she chooses the action and then acts. (“Agent” is the term that philosophers use for the person who performs an action.)

Some philosophers argue that the best way to make sense of this is to recognize a special form of causation, called agent causation. It is supposed to be different from event causation, in which one event causes another, for example the firing of one brain cell causing another cell to fire.

There are several possible attitudes to agent causation, including the following.

> Agent causation is just as real as event causation, however real that may be. Events cause events to occur, and agents cause events to occur too. Agents really do initiate new causal chains.
>
> Agent causation is not like event causation. Only events cause other events to occur. But we have to talk as if agent causation were real in order to be able to use concepts of decision and action. We talk as it it were real, whenever we talk about people deciding to do things and acting. And we have to do that, in order to make any kind of sense of our lives. This is not just a matter of needing to skip over a lot of detail about which brain cells do what.
>
> It is very useful to talk as if agent causation were real. We say “She decided to move the sofa, so she moved it”. But this is just a matter of skipping over the detail. It is a shorthand for “Her brain cells were in state X, and this caused region Y of her brain to be activated, which led to …”. When we talk about examples of decision and action, we are effectively talking about states of brains and bodies. We just don’t notice that fact.
>
> Talk as if agent causation were real is dangerously misleading. In everyday life, it is fine to talk about what people decide and do. But when we do philosophy, we try to work out what is really going on. In order to avoid making mistakes, we should stop talking about decision and action as soon as we start doing philosophy.

There are several arguments in favour of agent causation. They can be used to argue that agent causation is just as real as event causation. They can also be used to argue that we should at least talk as if agent causation were real, and that doing so is not just a way of skipping over detail. Here are some arguments to consider.

It seems to us as though we start off new causal chains. We do not deny what seems obvious to us in other circumstances: if we can see a tree in front of us, we accept that there is a tree there, except in very special circumstances where someone is playing tricks on us. So we should accept the obvious impression that we start off new causal chains, and our philosophy should work around that.

If events just occur, because they are caused by other events, we can be surprised by them. We may be surprised when snow suddenly falls off a roof, because we were unaware of the cause: poor insulation meant that when heaters were turned on in the house, the layer next to the roof melted. Our own actions are not like that. We are not usually surprised to find ourselves doing something. So our actions cannot simply be events that are caused by other events.

Talk of actions, which are initiated by agents and which are not seen as links in the middle of causal chains, is fundamental to how we live. We see people as choosing what to do, and we hold people responsible for their actions. We should not throw all that overboard, just to satisfy some philosopher who is worried about the details of agent causation.

There are also several arguments against agent causation. They can be used to argue that there is no real agent causation. They can also be used to argue that agent causation is a useless concept that explains nothing. Here are some arguments to consider.

Event causation is well-integrated into our scientific theories, and has useful roles to play all over the place. Agent causation has been invented to do one specific job, and is not well-integrated with the rest of our understanding of the world. So it is not scientifically respectable.

Agents would need to be special substances that could be causes in themselves, unlike the everyday natural substances that enter into causal processes only through events that involve them.

Only events, which have the property of occurring at particular times, could explain why other events, including actions, occurred at particular times.

If we invoke agent causation, we have no way to attribute adequate control to the agent when the agent could have performed any one of a range of different actions. Did the agent control his or her exercise of agent causation, and if so, how? The notion of responsibility for one's actions, whether moral responsibility or more general causal responsibility, can be undermined, because it can seem to be a matter of luck which action an agent chooses.

Questions for discussion

How strong a conclusion could we draw from each of the above arguments? For example, which ones could show that agent causation was real, and which ones could only show that it was useful to talk as if it were real?

What should our overall attitude to agent causation be?

Action, location and self-knowledge

We can, and need to, locate ourselves. We can align personal space ("two metres in front of me", "over there to the right") and objective space ("on the kitchen windowledge", "at the top of Primrose Hill"). This allows us to follow instructions like "Meet me at the top of Primrose Hill".

We also can, and need to, align personal time and objective time. If the instruction is "Meet me at the top of Primrose Hill at midday", and we know that it will take 15 minutes to walk there, we need to know when now (personal time) = 11.45 (objective time).

Each person can also identify a person as himself or herself. This is also important. Suppose that someone is walking through the woods, hoping to observe deer in the wild. He cracks a twig underfoot. Some deer are startled and run off. It is not enough for him to realise that someone must stop making sudden noises. He must realise that it is he himself who must tread more carefully.

More broadly, we all have a sense of a self in the world. We rely on it when we say things like "I live in London", or "I will go to Vienna again this year".

One can argue that we manage these things through action. Arguments include the following.

> Merely observing someone who fitted a given description, or whose limbs moved in certain ways, would not give anyone a clear sense that this person was himself or herself.
>
> Action, on the other hand, connects oneself to a particular physical person. Someone decides to move her arm, and observes it moving. She is directly aware of acting: that is not something she observes. And she is directly aware that she herself is acting. So the person whose arm moves must be herself. (See Lucy O'Brien, *Self-Knowing Agents*.)
>
> Action also gives us a sense of our own location in space and time, aligning our personal space and time with objective space and time. Most action is local, rather than at a distance. If someone wants to re-arrange the ornaments on the mantelpiece, she has to go over to the mantelpiece and pick them up. If someone wants to set off a firework display on New Year's Eve 1900, he has to be at that time (which is no longer possible). The place where we can act is the centre of personal space, and is also our personal time "now". But we can also be told where we are in objective space and time, for example, in the Stadtpark in Vienna at noon on 3 June 2011. Given that extra information, we can match up personal and objective space and time. We can add orientation. Personal orientation comes from action: if moving a hand forward leads to touching a statue, the statue is in front of oneself. Objective orientation is given independently of us: for example, the statue is 50 centimetres north of one's body. Putting these pieces of information together tells us that "in front" = "north". Mere observation, without action, would not give us our location in space and time, or our orientation. It would only tell us about the source of our sensory inputs. Those inputs might come straight into our brains from cameras and other sensors that were located far away from ourselves.

Question for discussion

Could someone who did not act in the world know who or where he or she was?