Richmond Journal of Philosophy 3 (Spring 2003)

On Scientific Realism

Pierre Cruse

Should We Believe In Unobservable Entities?

One of the most notable features of modern science is that it explains phenomena we observe by postulating entities that we do not. Many of the explanations appeal to entities that are too small to be seen – the TV works because electrons are being fired at the screen and causing it to illuminate in certain patterns; you have the eye colour you do because you inherited DNA that codes for eye colour when you were conceived; you have a fever because a virus is attacking your immune system. Some appeal to entities that are too big, or too far away – radiation from distant stars has a slightly lower frequency than we expect because the universe (too big to see) is expanding; it isn't expanding as fast as we think it should be because the universe contains 'dark matter' – objects (e.g. planets) in space that do not emit light, so cannot be detected by normal astronomical observation. And some entities are unobservable not because of their size, shape, or distance, but simply because of their nature – metal objects get attracted to a magnet because of the magnetic field it causes, the mass of the object causes it to resist acceleration, but neither the field nor the mass itself can directly be observed.

Entities falling into these categories pervade modern science. But the fact we don't actually directly observe them raises the question: do we know they are really there, and if so how? These questions divide philosophers of science into *scientific realists* – who believe in the reality of theoretical entities – and *anti-realists* or *instrumentalists* (after the view that theoretical postulates are just instruments for generating predictions) who do not. In this article I will look at some of the reasons which have led philosophers to take up these positions, and put forward my own view on how some of these disputes might be resolved.

Scientific Realism and Explanation

Let us begin by looking at how one might justify realism about theoretical entities. Realism is probably the most intuitive position, since most would probably assume that what our best scientific theories say about the world is true. After all, do scientists not have *evidence* for the claims that they make about the world?

Well, let us ask what sort of evidence there is for holding that, say, electrons exist. Although the name 'electron' was used for various hypothetical posits beforehand, electrons are said to have been 'discovered' by physicist J. J. Thomson in 1897. A potted history leading up to Thomson's discovery is this. If one places an 'anode' and a 'cathode' – two metal terminals connected to a wire carrying an electric current – and places them inside a vacuum tube, one can create a visible 'ray' that travels between the anode and the cathode.

During the latter part of the nineteenth century there was some controversy about what these rays are – do they consist of waves or particles? J. J. Thomson was the first to succeed in showing that the rays were deflected by an electric field, showing that they must consist of negatively charged particles; and he was able to measure the mass/charge ratio of these particles. The evidence for the existence of the electron, then, seems in Thomson's time to have been that the existence of tiny charged particles could explain the various results available about the behaviour of cathode rays – their deflection, their power of penetration, the fact that they can do 'work' (they can, for example, turn a 'windmill' placed in their path). In short, electrons were thought to exist because they were the *best explanation* of results relating to cathode rays.

The fact that the existence of electrons would best explain experimental results would seem to give us strong grounds for thinking that electrons are real. But let us ask why we think this is true. To draw this conclusion we need to use the following principle, known as the principle of 'inference to the best explanation' or IBE:

> We observe evidence E
>
> Theory T is the best explanation of evidence E,
>
> Therefore,
>
> Theory T is true

But we might well ask what justification there is for thinking that this principle will give us true conclusions when the premises are true. Note that it doesn't seem to do this as a matter of *logic*. Compare the principle of IBE with the principle of *modus ponens* (the principle that if it is true that *P*, and it is true that *if P then Q*, then it is true that *Q*). In the case of modus ponens, it looks as though it is *impossible* for *P* to be true and *if P then Q* to be true, but *Q* false. But it doesn't seem impossible that the best explanation of some evidence is false. For example, in the case of electrons, the evidence certainly seems to suggest that electrons are particles – since waves don't get deflected by electric fields – but it is surely *possible* that they are not particles at all, but a special sort of wave that violate the usual laws about the way that waves behave, and *do* get deflected by electric fields. The reason we think they are not seems not to be that we think this impossible, but simply that it seems highly *unlikely* to be the case, since it involves all sorts of needless complications to our theory. But this just raises the question again – is there any reason to think that the best explanation of our evidence is *likely* to be true?

The No Miracles Argument

There is another argument that scientific realists often appeal to in order to justify the view that the best explanations of our evidence are likely to be true. We can see how it works by looking at an analogy. Suppose that a friend of yours comes up with an implausible conspiracy theory, for example, that the government has rigged the result of the national lottery, so that a prominent member will win it at the next draw. Then suppose that a prominent government member does in fact win the next lottery draw. In that case it is *possible* that your friend's theory is false, and that the government member won just by chance. But for this to be the case your friend's prediction would have to have been incredibly lucky. Given her successful prediction, a more reasonable conclusion seems to be that her theory was actually true.

Something similar might be said in the case of electrons. Suppose I make a theoretical claim about unobservable entities (e.g. that cathode rays consist of electrons), as an explanation for certain phenomena (e.g. the fact that they get deflected in electric fields). Suppose I then predict on this basis that certain further things will happen (e.g. the cathode rays will turn a 'windmill') and that then these things *do* in fact happen. Now if my theory was false, it seems that the truth of my prediction would have to be put down to mere good luck – if cathode rays *don't* consist of electrons, then there is no reason to expect that they should behave as my theory says. But if my theory is true – and cathode rays really do consist of electrons – then we should expect them to behave just as I said they would. As with the conspiracy theory, if the events that the theory predicts genuinely occur – especially if the events are surprising or novel events – it seems much more reasonable to conclude that the theory was true than that it was false.

This observation is the basis of what has become known as the 'no-miracles' argument for scientific realism. Philosopher Hilary Putnam puts it succinctly as follows: 'the positive argument for realism is that it is the only philosophy that does not make the success of science a miracle'.[1] His idea is basically the one we have looked at. Many scientific theories throughout history – which were put forward because they were thought to be the best explanations for known evidence – led to further surprising and novel predictions, which in fact came true. Now according to the scientific realist this is just what we would expect, since those theories were really true. But if anti-realism or instrumentalism is true – and theories do *not* make true claims about the unobservable – then the ability of theories to make true predictions would be inexplicable; nothing short of an incredibly lucky fluke or a 'miracle'. Thus, the scientific realist says that it is much more reasonable to suppose that successful theories really are true.

Scientific Realism and the History of Science

Many scientific realists hold that the no-miracles argument provides some of the strongest support available for realism. However, by no means all philosophers of science have been convinced by the argument. For the remainder of this article I will look at one of the major reasons why not. The no-miracles argument asks us to conclude from the fact that certain theories in science have been very successful in predicting phenomena that they are true, and that the entities they postulate exist. But when we reflect on what has happened in the *history* of science, it becomes obvious that this is a very problematic claim.

We can begin to see why it is problematic by looking at our current attitude to theories of the past. It is undeniable that the history of science contains a large number of theories which we now regard as false. Many we now regard as *completely* false. For example, Aristotle believed that the planets moved around the earth embedded in crystalline spheres; Galen held that blood was produced in the liver, pumped by the heart around the body and then simply consumed; Descartes thought that gravity worked by the contact of swirling 'vortices' of solid particles that permeated the whole of space. These theories, though they were put forward with what at the time were good reasons, seem from the perspective of modern science to be fundamentally mistaken. Other theories we regard as false, but for more subtle reasons. For example, Newtonian mechanics is almost exactly true when we consider medium-sized objects moving at low speeds, but needs to be corrected to deal with phenomena at the very small or very large levels, or objects moving at very high speeds. Although the corrections are crucial, the discrepancies are so subtle they would probably have been undetectable in Newton's time.

A second feature of the history of science is that many of these theories despite being false were to a greater or lesser degree successful in explaining and predicting observed facts. For example, Aristotle's theory of crystalline spheres, though false, was able to predict the motions of the planets to quite a high degree of accuracy. And of course Newton's theory was able throughout the period of over 200 years when it was the dominant theory in physics to generate a huge variety of successful predictions.

For example, the observation that the orbit of Uranus diverged from the predictions of Newton's theory led to the prediction that a planet existed outside it, pulling it slightly out of its orbit – this led to the discovery of the planet Neptune in 1846. Note, then, what consequence these features of the history of science have for scientific realism. According to the realist we should believe that successful theories are true because otherwise their success in making predictions would be completely inexplicable. But in the history of science we find a number of theories which are undeniably both successful, *and false*. If this is right then the claim that successful theories are always true looks downright untenable. Moreover, since it looks as though the only justification we have for thinking that *any* theories are true is that they are successful in explaining and predicting experimental results, the fact that many theories have been successful without being true even undermines our justification for thinking that the theories we have *now* are true. This problem has become known as the 'pessimistic induction', since it suggests that generalising from the history of failures in science, we should be pessimistic about the chances of our own theories being true. The realist clearly needs some way of responding to this argument.

A More Careful Realism

In fact, given the way we have described the problem, the realist has quite a convincing response. Consider again the list of theories that we have said were false. Some of these theories are false, a realist might say, but not disastrously so. For example, although Newton's theory is false, it is also thought to be a 'limiting case' of subsequent theories such as special relativity, that is to say, Newton's theory is true if we assume that the speeds at which objects travel are vanishingly small in comparison with the speed of light (an assumption that is almost exactly true of the 'middle-sized' bodies we find on earth). One could therefore claim that although Newton's theory is strictly false, it is still *approximately* true, or *close to the truth*. The realist's claim can therefore be stated slightly more carefully to incorporate cases like this. Instead of saying that successful theories are *completely true*, the realist can claim that they are generally *approximately* true, and that this, rather than their actual truth, explains their success.

The realist is also in a position to deal with the *completely* false theories that we mentioned. Contrast the sorts of predictions that Newton's theory was able to make with those of Aristotle's theory of crystalline spheres. Newton's theory, as we saw, was able to generate predictions of phenomena that were unknown before, such as the existence of the planet Neptune. Aristotle's theory of crystalline spheres, on the other hand, while being able to explain the known motions of the planets to a certain degree of accuracy, was unable to come up with anything genuinely *novel*, that wasn't known to occur before the theory was proposed. The realist can thus point out that it isn't really a miracle if a false theory explains results that are already known to occur: these could just be 'built in' to the theory, and give us no reason to think it is true. The real 'miracle' would be if a false theory made a number of correct, *novel* predictions, about things that weren't known to occur before the theory was formulated. Thus, the realist can claim that we should only believe that theories that are successful in making novel predictions are approximately true, pointing out that many of the theories we now think are completely false did not make genuinely novel predictions.

Deeper Problems

Unfortunately, there are some remaining historical examples which seem to pose a problem for even this more cautious version of realism. One of the most notorious examples is the case of the optical aether. Before the nineteenth century, most scientists believed, after Newton, that light consisted of small, solid particles. However, around the beginning of the 19th century, it began to be realised that many optical phenomena – for example the pattern of light and dark patches created caused by interference when light is shone through two slits – could better be explained by supposing that light was constituted of waves. At the time, though, all known cases of waves occurred in a *medium*, for example in water, or on a string; and it was assumed that light waves must too. Thus scientists postulated an *aether*, which was a jelly-like, though very rigid, substance that permeated the whole of space: the immensely fast transverse (i.e. side to side) vibrations of this substance were thought to be light waves.

Aether theories were certainly successful in predicting results. In a famous case, French scientist Augustin Fresnel had submitted an entry for the Grand Prix of the Académie des Sciences in Paris in 1819, and a member of the judging panel, Siméon-Denis Poisson, worked out that Fresnel's theory – based on the waves-in-aether theory of light – would imply that if light were shone at a small opaque disc the shadow would have a small bright spot in the centre. This rather surprising prediction was later demonstrated to be true, and Fresnel was awarded the Grand Prix, despite the opposition of most of the panel members to the wave theory he defended.

The problem for scientific realism is that we now believe that there is no aether. Since Einstein's work around the beginning of the 20th century, light waves have been explained in terms of oscillations in electromagnetic fields: but fields can exist in empty space, and do not require the solid aether that nineteenth-century physicists postulated. This makes the aether a rather more problematic case for the realist than theories such as Newtonian mechanics or Aristotle's theory of crystalline spheres. In the case of Newtonian mechanics the realist could admit the theory was successful, but call it approximately (though not completely) true, and in the case of Aristotle's theory the realist could deny that the theory was really successful. But in the case of aether theories it looks as though neither move is going to work. Aether theories were certainly successful. But given that the central entity they talked about just doesn't exist, it is very difficult to argue that they are in any meaningful sense 'approximately true'.

Although we have focussed on the aether, critics of realism point out there are a number of further theories which were genuinely successful, but which also seem crucially to involve non-existent entities. These include the phlogiston theory of chemistry, the caloric theory of heat, and others.[2] If each of these theories is indeed successful, but completely mistaken, then the scientific realist's claim that successful theories are generally close to being true looks in dire trouble.

A Realist Reply

The problem of apparently successful theories that are now known to be based on fundamentally wrong assumptions has been known in recent years to be one of the major difficulties with scientific realism, and a number of philosophers of science have attempted to address it.[3] In this final section I will describe part of what I consider to be the best response to this problem.

To begin to see how we might reply, let us compare the case of the optical aether with another entity which was being investigated at around the same time, the chemical atom. Although the idea that all matter consists of small indivisible particles goes back to the ancient Greeks, an important stage in the development of the modern concept of the atom was the proposal of English chemist John Dalton's atomic theory in 1808. Dalton proposed that each chemical element was comprised of identical atoms of a certain mass (different for each element) which combined in fixed small whole-number ratios to form compounds. This theory allowed him to use available data about the relative weights of different elements that go into forming compounds to determine the relative weights of atoms of different elements, something that had not been possible before him. However, Dalton's conception of an atom was that atoms of each element are merely solid, indivisible spheres. This model is clearly quite different to the modern view of atoms as comprised of protons and neutrons in a nucleus surrounded by orbiting electrons.

Suppose we compare, then, the relative merits of Fresnel's aether theory and Dalton's atomic theory from the modern point of view. The aether theory was correct in saying that light was essentially a wave-like phenomenon, and led Fresnel to what are from our point of view mathematically very accurate descriptions of interference and polarisation and other optical effects. However, Fresnel was wrong in thinking that the carrier of light waves was a solid aether. Dalton, on the other hand was right about the fact that atoms of different elements could combine in whole-number ratios, and right in thinking that this fact could be exploited to discover the relative atomic weights of different elements. However, he was wrong in some of his particular claims about relative weights (e.g. those of hydrogen and oxygen, since he thought water was formed from hydrogen and oxygen in a 1:1 ratio, rather than 2:1 as we think nowadays), and wrong in his view about the nature of atoms.

There is therefore surely a case to be made that Fresnel's theory is at least as accurate or 'close to the truth' as Dalton's atomic theory. Maybe it is even *closer* to the truth, since it was more accurate in describing optical phenomena like polarisation and interference than Dalton's was in describing relative atomic weights. But despite this fact most people would still think that when Dalton talked about atoms he was talking about the same things as we now call 'atoms' – Dalton's 'atom' existed, in other words – whereas when Fresnel talked about the ‘aether’, he was talking about something that didn’t exist.

With this in mind let us go back to the main claim of the ‘pessimistic induction’. According to this argument, Fresnel’s theory cannot be said to be approximately true because the central entity that it postulates – the aether – does not really exist. But comparing the aether theory with Dalton’s theory it seems the aether theory is arguably *better* – i.e. closer to being true – than Dalton’s theory even though Dalton is talking about an entity that *does* exist. This suggests that whether the entities a theory postulates ‘really exist’ is not, after all, so crucial a question in deciding whether the theory is approximately true.

I claim we can explain this as follows. Part of the reason we think that the aether doesn’t exist is that some of the crucial features that Fresnel attributed to the aether – e.g. the fact that it is a solid medium – don’t in reality apply to anything. But part of the reason is just the fact that at a certain point in science the community simply decided as a matter of convention not to use the term ‘aether’ any more. In fact it was Einstein who was first responsible for dropping the concept of the ‘aether’, when he claimed it was ‘superfluous’ according to his special theory of relativity. However, we could easily imagine that Einstein had instead decided to continue to use the word ‘aether’ to refer to electromagnetic fields; if he had we would probably now say that the aether really existed. The fact that this didn’t happen, and that we continue to say there is no aether, does not seem to be relevant to whether theories that went before him – e.g. Fresnel’s – were approximately true. My claim, then, is that it is partly just a matter of convention that we say the aether doesn’t exist, and it doesn’t mean that aether theories could not have been approximately true.

Can the realist also give some positive grounds for thinking that aether theories *were* approximately true? I think this is easily done. Whether or not we want to say that the aether exists, we know what Fresnel said *about* the aether. For example, he claimed that:

> a) It vibrates transversely (i.e. from side to side): these vibrations constitute light waves, and explain interference, polarisation, etc.
>
> b) It is a universal, solid, very rigid, jelly-like substance.

Now we know that b) is not true of anything. However, there *is* something that has property a) – the electromagnetic field. So we can certainly say that there *exists something* with many of the important features that Fresnel thought the aether had. Moreover, that there really does exist something with feature a) seems like a good explanation of why Fresnel was able to correctly predict the results of interference and polarisation experiments. It seems to me that this gives us sufficient grounds to say that Fresnel’s theory was approximately true whether or not we say that the aether ‘really existed’.

Conclusion

We have seen that a strong argument for scientific realism is that if our most successful theories were not at least approximately true then their success in predicting novel, previously unknown, phenomena would be very difficult to explain – it would look like a lucky coincidence. However, a challenge for the realist comes from the fact that many theories in history were successful, whereas modern science tells us that they were fundamentally mistaken in the way they describe the world.

My claim is that this problem is not as bad for the realist as it seems at first. One of the major reasons we think certain scientific theories of the past were mistaken is that we think the central entities they postulated – such as the aether – don't exist. But the fact we say this is just a matter of convention, and doesn't necessarily mean that the theory in question isn't approximately true. By looking at exactly what the theories said *about* the entities they postulated, we find that even theories that postulated entities that we think didn't exist say a lot of things that we would regard as true. The realist still has a lot of work to do to persuade us that successful theories are almost *always* approximately true, but I don't think things are quite as bad as the 'pessimistic induction' makes out.

Pierre Cruise
King's College London

[1] Hilary Putnam, *Mathematics, Matter and Method: Philosophical Papers vol. 1* (Cambridge: Cambridge University Press, 1975), 73

[2] The original problem was raised in an article by Larry Laudan, 'A Confutation of Convergent Realism', *Philosophy of Science* 48 (1981), 19-49, reprinted in D. Papineau ed. *The Philosophy of Science* (Oxford: OUP, 1999)

[3] Two recent examples are Philip Kitcher, *The Advancement of Science* (New York: Oxford University Press, 1993) ch 5, and Stathis Psillos, *Scientific Realism: How Science Tracks Truth*, chs 4-7, (London: Routledge, 1999)