1

Philosophy of Science

James Ladyman
Philosophy, University of Bristol, UK

1.1 What is science?

Dictionary definitions speak of a systematic body of knowledge, and the word ‘science’ comes from the Latin word for knowledge. However, not any old collection of facts – even one that is organized – constitutes a science. For example, an alphabetical list of all the words that are used in this book and all the others published on the same day would make no contribution to scientific knowledge. Something else is needed, and there are two obvious supplements to what has been said so far:

- First, the subject matter must be the workings of the physical world. There must be discovery of natural laws and the relations of cause and effect that give rise to the phenomena that we observe.
- Second, the relevant theories must be generated in the right way.

In fact, most philosophers of science and scientists define science in terms of its methods of production; science is knowledge produced by the scientific method. For many people then, asking the question with which we began really amounts to asking, ‘What is the scientific method?’

There are, of course, many methods, and this book is about some of them. The techniques and procedures of the laboratory and experimental trials and the measurement, recording and representation of data, as well as its statistical analysis, form at least as much a part of science as what it tells us about the world as a result. Clearly, the methods of geology and astrophysics differ from those of cell biology or pharmacology.

However, all the sciences we now take for granted have really only reached maturity and separation from each other within the last few hundred years. For example, biochemistry and neuroscience have only become separate disciplines in...
the last century, and whole areas of enquiry were impossible before the invention of electron microscopy and magnetic resonance imaging. Our gigantic science faculties, with their highly specialized disciplines, originated in the ancient and medieval systems of knowledge, and these made very few of the distinctions in subject matter that we now would. For example, many posited connections between the planets and human diseases and other conditions where we find none. Nonetheless, we can find some original truths from many subjects discussed a long time ago. For example, Aristotle recorded that bees pollinate flowers, and the 28-day cycle of the Moon’s phases has been known since prehistory.

Modern science is usually regarded as having originated at the turn of the 16th and 17th centuries. At this time, the established ways of predicting the motions of the planets, which placed the Earth at the centre of the solar system, were replaced by the Copernican theory placing the Sun at the centre, which was then modified by Kepler to incorporate elliptical orbits. The latter’s laws were precise mathematical statements that fitted very well with the detailed data that had recently been gathered using new optical technology. In the years that followed, telescopes, microscopes, the air pump and clockwork and other mechanical devices were invented and, over the next few generations, knowledge of chemistry, biology, medicine, physics and the rest of what was then called ‘natural philosophy’ grew enormously.

An amazing thing about all the scientific knowledge that we now take for granted is that the founders of modern science envisaged its production by the collaborative endeavour of people following the scientific method. They argued that there was a common core to all the methods mentioned above, and they advocated the collective use of a single set of principles or rules for investigation, whatever the subject matter. Different people had different ideas about exactly what the method should be, but everyone agreed that testing by experiment is fundamental to science. The task, therefore, is to say what exactly ‘testing by experiment’ means.

There two general kinds of answer:

- **Positive**, according to which the job of scientists is to gather data from which to infer theories, or at least to find out which theories are supported by it.
- **Negative**, according to which the real task is to try and prove theories false.

The latter may sound strange, but in fact many scientists put more emphasis on it than the former. The reason for that is that there is a very great tendency in human thought to find confirmation of preconceptions and received ideas by being selective in what is taken into account.

The phenomenon known as ‘confirmation bias’ has been studied extensively in psychology; it is manifested in many ways, including by people selectively remembering or prioritizing information that supports their beliefs. It is very difficult to overcome this tendency, so some people argue that science should always be sceptical and that attempts to prove theories false should be at its heart.
Modern science began with the upturning of many entrenched beliefs about the world, but since then the history of science has repeatedly involved the overturning of cherished doctrines and the acceptance of previously heretical ideas. Examples include the motion of the Earth, the common ancestry of the great apes and human beings, the expansion of the universe and its acceleration, the relativity of space and time, and the utter randomness of radioactive decay. Even the greatest scientific theories, such as Newton’s physics and Lavoisier’s chemistry, have been subject to substantial correction.

Hence, many scientists follow the philosopher of science Karl Popper in saying that the scientific method consists in the generation of hypotheses, from which are deduced predictions that can, in principle, be falsified by an experiment. When an experiment does not falsify the hypothesis, it may tentatively be employed to make predictions – but the aim should be to seek new kinds of test that may prove it false. A theory that makes specific and precise predictions is more liable to falsification than one that makes only general and vague claims; so, according to Popper, scientists should strive to formulate hypotheses from which very exact statements about experimental outcomes can be derived, and to say in advance what would count as falsification.

Popper emphasized that scientific knowledge is always revisable in the light of new empirical findings, and that science has succeeded in increasing its accuracy, depth and breadth, because even well-established theories are not regarded as immune from correction and revision. Science is not compatible with absolute certainty and the refusal to question.

However, it is also true that in practice, scientists do not immediately abandon core theories when experiments go against them. For example, Newton’s law of universal gravitation, the famous inverse-square law, gave beautifully accurate predictions for the paths of the planets in night sky and improved on those of Kepler, as well as generating successful new predictions such as the return of Halley’s comet and the flattening of the curvature of the Earth at the poles. However, in the 18th century it was found that the orbit of Uranus was not as predicted, but astronomers did not abandon Newtonian mechanics as a result. Instead, they looked at the other assumptions that they had made in order to calculate the orbit. They had assumed that only the gravitation attraction of the Sun and six other planets needed to be taken into account. If there was another planet, that might explain the anomaly; therefore, Neptune was looked for and found.

Modifying a theory to take account of data that contradicts the original is not, in itself, bad practice. In the case just mentioned, the modification led to a new prediction that could be tested. Science often proceeds like this and, indeed, Pluto was found in the same way. It is now common in astronomy to infer the existence of unobservable objects because of their hypothetical gravitational effect on observable ones.

These examples illustrate an extremely important feature of science, which is that predictions and, hence, tests are never of single hypotheses but always of a
collection thereof. To predict the orbit of a planet, one must know all the bodies to whose gravitational attraction it is appreciably subject, and also all of their masses and its mass. If the data do not fit, then logic dictates that there is a problem with at least one of the laws or the other assumptions – although not which one. This is called the Duhem problem (after Pierre Duhem). Scientists face this every day, but they rarely consider that a central theoretical component is false as Popper imagines. To do so would not be sensible, because those core beliefs have been at the centre of a vast number of successful predictions. On the other hand, there will often be many other plausible culprits among the other assumptions involved, and the art and practice of science involves teasing them apart and finding out which to amend.

It is not plausible to argue, as Popper did, that no matter how much a hypothesis has agreed with experiment and survived attempts to show it to be false, there are no positive grounds for belief in it. Since Francis Bacon proposed his new logic of ‘induction’, many others have sought to develop an account of how evidence can be said to support or confirm a theory. Thus we have two extreme positions:

- **Falsificationism** says science is about showing theories to be false.
- **Inductivism** says science is about showing theories to be true.

It is tempting to seek a happy medium able to incorporate the importance of both, but clearly we cannot do this without some notion of confirmation in science. It is often the case that we look to science to tell us positive facts, such as that a drug is efficacious and safe, or that a particular pathogen is the cause of some medical problem. Bayesian statistics provides measures for how much a given body of evidence supports a given hypothesis. On the other hand, statistical methods are also sometimes used in a falsificationist spirit, as when they are used to calculate the probability of the so-called ‘null hypothesis’, according to which some potential causal factor has no effect.

The fundamental problem with the scientific method is that it cannot tell us how confident we need to be in a theory before we accept it. Nor, if a research programme is in trouble, can it tell us exactly when to abandon it. For example, in the 19th century, more accurate measurements revealed that the orbit of Mercury did not fit with the predictions of Newtonian gravitation. The trick of positing another planet was tried but, because Mercury is so close to us, any such new planet ought to have been immediately obvious. Thus, it was thought that perhaps it was always the other side of the Sun from us. As it turned out, there is no such planet, and it took Einstein’s then new theory of General Relativity to solve the problem.

Similarly, when the evidence begins to come in about the efficacy of a new drug, there is no mathematical formula that can say when we should regard it as ‘known’ to be effective. Some scientists may feel sure very early on in the trials, and there may be patients who could benefit from its immediate prescription. However, others will insist that larger studies need to be done before the evidence is compelling.
the end, a committee will set the bar at some level, perhaps demanding that the probability of the null hypothesis for the drug acting on the condition be shown to be less than 0.05 per cent. That is reasonable, but it could also be set at 0.5 per cent or 0.005 per cent, or any other small value and which value is chosen is to some extent arbitrary. Clearly, if the chance of a drug being completely useless is 50 per cent. it should not be prescribed, and if it is .0000000005 per cent then it should be; but where exactly the line should be drawn between these extremes is a matter of choice and judgment.

It is therefore important to be very clear about the limitations of the scientific method, as well as its great power. How much evidence we demand before reaching a conclusion depends in part on whether we are more keen to have true beliefs or to avoid false ones. If all we care about is having true beliefs, then, for example, above all else we will wish to avoid failing to believe a drug works when it does; if all we care about is not having false beliefs, then, for example, we will wish above all else to avoid believing that a drug works when it does not. The former attitude emphasizes avoiding false negatives and the latter emphasizes avoiding false positives, and in general doing well in respect of one is at the cost of doing badly in respect of the other. Falsificationists emphasize avoiding false positives, so they always think of scientific theories as not yet falsified rather than as confirmed.

The problem is that, both in life and in science, we often need to stick our necks out and commit to the truth of a theory, because if we always wait for one more trial, patients will be denied treatments they need. Part of being a good scientist is developing good judgment about such matters, and it is also necessary to learn where reasonable disagreement is possible, how to identify the crux of such disputes, and how to use the scientific method to refine the evidential basis on which they can be resolved.

Further reading
