about the truth of observation statements depend on what is already known or assumed, thus rendering the observable facts as fallible as the presuppositions underlying them. Both kinds of difficulty suggest that maybe the observable basis for science is not as straightforward and secure as is widely and traditionally supposed. In the next chapter I try to mitigate these fears to some extent by considering the nature of observation, especially as it is employed in science, in a more discerning way than has been involved in our discussion up until now.

Further reading

For a classic discussion of how knowledge is seen by an empiricist as derived from what is delivered to the mind via the senses, see Locke (1967), and by a logical positivist, see Ayer (1940). Hanfling (1981) is an introduction to logical positivism generally, including its account of the observational basis of science. A challenge to these views at the level of perception is Hanson (1958, chapter 1). Useful discussions of the whole issue are to be found in Brown (1977) and Barnes, Bloor and Henry (1996, chapters 1–3).

CHAPTER 2

Observation as practical intervention

Observation: passive and private or public and active?

A common way in which observation is understood by a range of philosophers is to see it as a passive, private affair. It is passive insofar as it is presumed that when seeing, for example, we simply open and direct our eyes, let the information flow in, and record what is there to be seen. It is the perception itself in the mind or brain of the observer that is taken to directly validate the fact, which may be "there is a red tomato in front of me" for example. If it is understood in this way, then the establishment of observable facts is a very private affair. It is accomplished by the individual closely attending to what is presented to him or her in the act of perception. Since two observers do not have access to each other's perceptions, there is no way they can enter into a dialogue about the validity of the facts they are presumed to establish.

This view of perception or observation, as passive and private, is totally inadequate, and does not give an accurate account of perception in everyday life, let alone science. Everyday observation is far from passive. There are a range of things that are *done*, many of them automatically and perhaps unconsciously, to establish the validity of a perception. In the act of seeing we scan objects, move our heads to test for expected changes in the observed scene and so on. If we are not sure whether a scene viewed through a window is something out of the window or a reflection in the window, we can move our heads to check for the effect this has on the direction in which the scene is visible. It is a general point that if for any reason we doubt the validity of what seems to be the case on the basis of our perceptions, there are various actions we can take to remove the problem. If, in the example above, we have reason to suspect that the image of the tomato is some cleverly contrived optical image rather than a real tomato, we can touch it as well as look at it, and if necessary we can taste it or dissect it.

With these few, somewhat elementary, observations I have only touched the surface of the detailed story psychologists can tell about the range of things that are done by individuals in the act of perception. More important for our task is to consider the significance of the point for the role of observation in science. An example that illustrates my point well is drawn from early uses of the microscope in science. When scientists such as Robert Hooke and Henry Powers used the microscope to look at small insects such as flies and ants, they often disagreed about the observable facts, at least initially. Hooke traced the cause of some of the disagreements to different kinds of illumination. He pointed out that the eye of a fly appears like a lattice covered with holes in one kind of light (which, incidentally, seems to have led Powers to believe that this was indeed the case), like a surface covered with cones in another and in yet another light like a surface covered with pyramids. Hooke proceeded to make practical interventions designed to clear up the problem. He endeavoured to eliminate spurious information arising from dazzle and complicated reflections by illuminating specimens uniformly. He did this by using for illumination the light of a candle diffused through a solution of brine. He also illuminated his specimens from various directions to determine which features remained invariant under such changes. Some of the insects needed to be thoroughly intoxicated with brandy to render them both motionless and undamaged.

Hooke's book, *Micrographia* (1665), contains many detailed descriptions and drawings that resulted from Hooke's actions and observations. These productions were and are public, not private. They can be checked, criticised and added to by others. If a fly's eye, in some kinds of illumination, appears to be covered with holes, then that state of affairs cannot be usefully evaluated by the observer closely attending to his or her perceptions. Hooke showed what could be *done* to check the authenticity of the appearances in such cases, and the procedures he recommended could be carried out by anyone suitably inclined and skilled. The observable facts about the structure of a fly's eye that eventuate result from a process that is both active and public.

The point that action can be taken to explore the adequacy of claims put forward as observable facts has the consequence that subjective aspects of perception need not be an intractable problem for science. Ways in which perceptions of the same scene can vary from observer to observer depending on their background, culture and expectations were discussed in the previous chapter. Problems that eventuate from this undoubted fact can be countered to a large extent by taking appropriate action. It should be no news to anyone that the perceptual judgments of individuals can be unreliable for a range of reasons. The challenge, in science, is to arrange the observable situation in such a way that the reliance on such judgments is minimised if not eliminated. An example or two will illustrate the point.

The moon illusion is a common phenomenon. When it is high in the sky, the moon appears much smaller than when it is low on the horizon. This is an illusion. The moon does not change size nor does its distance from earth alter during the few hours that it takes for its relative position to undergo the required change. However, we do not have to put our trust in subjective judgments about the moon's size. We can, for example, mount a sighting tube fitted with cross-wires in such a way that its orientation can be read on a scale. The angle subtended by the moon at the place of sighting can be determined by aligning the cross-wires with each side of the moon in turn and noting the difference in the corresponding scale readings. This can be done when the moon is high in the sky and repeated when it is near the horizon. The fact that the apparent size of the moon has remained unchanged is reflected in the fact that there is no significant variation in the differences between the scale readings in the two cases.

Galileo and the moons of Jupiter

In this section the relevance of the discussion in the previous chapter is illustrated with an historical example. Late in 1609 Galileo constructed a powerful telescope and used it to look at the heavens. Many of the novel observations he made in the ensuing three months were controversial, and very relevant to the astronomical debate concerning the validity of the Copernican theory, of which Galileo became an avid champion. Galileo claimed, for instance, to have sighted four moons orbiting the planet Jupiter, but he had trouble convincing others of the validity of his observations. The matter was of some moment. The Copernican theory involved the controversial claim that the earth moves, spinning on its axis once a day and orbiting the sun once a year. The received view that Copernicus had challenged in the first half of the previous century was that the earth is stationary, with the sun and planets orbiting it. One of the many, far from trivial, arguments against the motion of the earth was that, if it orbited the sun as Copernicus claimed, the moon would be left behind. This argument is undermined once it is acknowledged that Jupiter has moons. For even the opponents of Copernicus agreed that Jupiter moves. Consequently, any moons it has are carried with it, exhibiting the very phenomenon that the opponents of Copernicus claimed to be impossible in the case of the earth.

Whether Galileo's telescopic observations of moons around Jupiter were valid was a question of some moment then. In spite of the initial skepticism, and the apparent inability of a range of his contemporaries to discern the moons through the telescope, Galileo had convinced his rivals within a period of two years. Let us see how he was able to achieve that — how he was able to "objectify" his observations of Jupiter's moons.

Galileo attached a scale, marked with equally spaced horizontal and vertical lines, to his telescope by a ring in such a way that the scale was face-on to the observer and could be slid up and down the length of the telescope. A viewer looking through the telescope with one eye could view the scale with the other. Sighting of the scale was facilitated by illuminating it with a small lamp. With the telescope trained on Jupiter, the scale was slid along the telescope until the image of Jupiter viewed through the telescope with one eye lay in the central square of the scale viewed with the other eye. With this accomplished, the position of a moon viewed through the telescope could be read on the scale, the reading corresponding to its distance from Jupiter in multiples of the diameter of Jupiter. The diameter of Jupiter was a convenient unit, since employing it as a standard automatically allowed for the fact that its apparent diameter as viewed from earth varies as that planet approaches and recedes from the Earth.

Using these, Galileo was able to record the daily histories of the four "starlets" accompanying Jupiter. He was able to show that the data were consistent with the assumption that the starlets were indeed moons orbiting Jupiter with a constant period. The assumption was borne out, not only by the quantitative measurements but also by the more qualitative observation that the satellites occasionally disappeared from view as they passed behind or in front of the parent planet or moved into its shadow.

Galileo was in a strong position to argue for the veracity of his observations of Jupiter's moons, in spite of the fact that they were invisible to the naked eye. He could, and did, argue against the suggestion that they were an illusion produced by the telescope by pointing out that that suggestion made it difficult to explain why the moons appeared near Jupiter and nowhere else. Galileo could also appeal to the consistency and repeatability of his measurements and their compatibility with the assumption that the moons orbit Jupiter with a constant period. Galileo's quantitative data were verified by independent observers, including observers at the Collegio Romano and the Court of the Pope in Rome who were opponents of the Copernican theory. What is more, Galileo was able to predict further positions of the moons and the occurrence of transits and eclipses, and these too were confirmed by himself and independent observers, as documented by Stillman Drake, (1978, pp. 175–6, 236–7).

The veracity of the telescopic sightings was soon accepted by those of Galileo's contemporaries who were competent observers, even by those who had initially opposed him. It is true that some observers could never manage to discern the moons, but I suggest that this is of no more significance than the inability of James Thurber (1933, pp. 101–103) to discern the structure of plant cells through a microscope. The strength of Galileo's case for the veracity of his telescopic observations of the moons of Jupiter derives from the range of practical, objective tests that his claims could survive. Although his case might have stopped short of being absolutely conclusive, it was incomparably stronger than any that could be made for the alternative, namely, that his sightings were illusions or artifacts brought about by the telescope.

Observable facts objective but fallible

An attempt to rescue a reasonably strong version of what constitutes an observable fact from the criticisms that we have levelled at that notion might go along the following lines. An observation statement constitutes a fact worthy of forming part of the basis for science if it is such that it can be straightforwardly tested by the senses and withstands those tests. Here the "straightforward" is intended to capture the idea that candidate observation statements should be such that their validity can be tested in ways that involve routine, objective procedures that do not necessitate fine, subjective judgments on the part of the observer. The emphasis on tests brings out the active, public character of the vindication of observation statements. In this way, perhaps we can capture a notion of fact unproblematically established by observation. After all, only a suitably addicted philosopher will wish to spend time doubting that such things as meter readings can be securely established, within some small margin of error, by careful use of the sense of sight.

A small price has to be paid for the notion of an observable fact put forward in the previous paragraph. That price is that observable facts are to some degree fallible and subject to revision. If a statement qualifies as an observable fact because it has passed all the tests that can be levelled at it hitherto, this does not mean that it will necessarily survive new kinds of tests that become possible in the light of advances in knowledge and technology. We have already met two significant examples of observation statements that were accepted as facts on good grounds but were eventually rejected in the light of such advances, namely, "the earth is stationary" and "the apparent size of Mars and Venus do not change appreciably during the course of the year".

According to the view put forward here, observations suitable for constituting a basis for scientific knowledge are both objective and fallible. They are objective insofar as they can be publicly tested by straightforward procedures, and they are fallible insofar as they may be undermined by new kinds of tests made possible by advances in science and technology. This point can be illustrated by another example from the work of Galileo. In his *Dialogue Concerning the Two Chief World Systems* (1967, pp. 361–3) Galileo described an objective method for measuring the diameter of a star. He hung a cord between himself and the star at a distance such that the cord just blocked out the star. Galileo argued that the angle subtended at the eye by the cord was then equal to the angle subtended at the eye by the star. We now know that Galileo's results were spurious. The apparent size of a star as perceived by us is due entirely to atmospheric and other noise effects and has no determinate relation to the star's physical size. Galileo's measurements of star-size rested on implicit assumptions that are now rejected. But this rejection has nothing to do with subjective aspects of perception. Galileo's observations were objective in the sense that they involved routine procedures which, if repeated today, would give much the same results as obtained by Galileo. In the next chapter we will have cause to develop further the point that the lack of an infallible observational base for science does not derive solely from subjective aspects of perception.

Further reading

For a classic discussion of the empirical basis of science as those statements that withstand tests, see Popper (1972, chapter 5). The active aspects of observation are stressed in the second half of Hacking (1983), in Popper (1979, pp. 341–61) and in Chalmers (1990, chapter 4). Also of relevance is Shapere (1982).

CHAPTER 3

Experiment

Not just facts but relevant facts

In this chapter I assume for the sake of argument that secure facts can be established by careful use of the senses. After all, as I have already suggested, there are a range of situations relevant to science where this assumption is surely justified. Counting clicks on a Geiger counter and noting the position of a needle on a scale are unproblematic examples. Does the availability of such facts solve our problem about the factual basis for science? Do the statements that we assume can be established by observation constitute the facts from which scientific knowledge can be derived? In this chapter we will see that the answer to these questions is a decisive "no".

One point that should be noted is that what is needed in science is not just facts but relevant facts. The vast majority of facts that can be established by observation, such as the number of books in my office or the colour of my neighbour's car, are totally irrelevant for science, and scientists would be wasting their time collecting them. Which facts are relevant and which are not relevant to a science will be relative to the current state of development of that science. Science poses the questions, and ideally observation can provide an answer. This is part of the answer to the question of what constitutes a relevant fact for science.

However, there is a more substantial point to be made, which I will introduce with a story. When I was young, my brother and I disagreed about how to explain the fact that the grass grows longer among the cow pats in a field than elsewhere in the same field, a fact that I am sure we were not the first to notice. My brother was of the opinion that it was the fertilising effect of the dung that was responsible, whereas I suspected that it was a mulching effect, the dung trapping moisture beneath it and inhibiting evaporation. I now have a strong suspicion that neither of us was entirely right and that the main explanation is simply that cows are disinclined to eat the grass around their own dung. Presumably all three of these effects play some role, but it is not possible to sort out the relative magnitudes of the effects by observations of the kind made by my brother and me. Some intervention would be necessary, such as, for example, locking the cows out of a field for a season to see if this reduced or eliminated the longer growth among the cow pats, by grinding the dung in such a way that the mulching effect is eliminated but the fertilising effect retained, and so on.

The situation exemplified here is typical. Many kinds of processes are at work in the world around us, and they are all superimposed on, and interact with, each other in complicated ways. A falling leaf is subject to gravity, air resistance and the force of winds and will also rot to some small degree as it falls. It is not possible to arrive at an understanding of these various processes by careful observation of events as they typically and naturally occur. Observation of falling leaves will not yield Galileo's law of fall. The lesson to be learnt here is rather straightforward. To acquire facts relevant for the identification and specification of the various processes at work in nature it is, in general, necessary to practically intervene to try to isolate the process under investigation and eliminate the effects of others. In short, it is necessary to do experiments.

It has taken us a while to get to this point, but it should perhaps be somewhat obvious that if there are facts that constitute the basis for science, then those facts come in the form of experimental results rather than any old observable facts. As obvious as this might be, it is not until the last couple of decades that philosophers of science have taken a close look at the nature of experiment and the role it plays in science. Indeed, it is an issue that was given little attention in the previous editions of this book. Once we focus on experiment rather than mere observation as supplying the basis for

science, the issues we have been discussing take on a somewhat different light, as we shall see in the remainder of this chapter.

The production and updating of experimental results

Experimental results are by no means straightforwardly given. As any experimentalist, and indeed any science student, knows, getting an experiment to work is no easy matter. A significant new experiment can take months or even years to successfully execute. A brief account of my own experiences as an experimental physicist in the 1960s will illustrate the point nicely. It is of no great importance whether the reader follows the detail of the story. I simply aim to give some idea of the complexity and practical struggle involved in the production of an experimental result.

The aim of my experiment was to scatter low-energy electrons from molecules to find out how much energy they lost in the process, thereby gaining information related to the energy levels in the molecules themselves. To reach this objective, it was necessary to produce a beam of electrons that all moved at the same velocity and hence had the same energy. It was necessary to arrange for them to collide with one target molecule only before entering the detector, otherwise the sought-for information would be lost, and it was necessary to measure the velocity, or energy, of the scattered electrons with a suitably designed detector. Each of these steps posed a practical challenge. The velocity selector involved two conducting plates bent into concentric circles with a potential difference between them. Electrons entering between the plates would only emerge from the other end of the circular channel if they had a velocity that matched the potential difference between the plates. Otherwise they would be deflected onto the conducting plates. To ensure that the electrons were likely to collide with only one target molecule it was necessary to do the experiment in a region that was highly evacuated, containing a sample of the target gas at very low pressure. This required pushing the available vacuum technology to its limits. The velocity of scattered electrons was to be measured by an arrangement of circular electrodes similar to that used in producing the mono-energetic beam. The intensity of electrons scattered with a particular velocity could be measured by setting the potential difference between the plates to a value that allowed only the electrons with that velocity to traverse the circle and emerge at the other end of the analyser. Detecting the emerging electrons involved measuring a minutely small current which again pushed the available technology to its limits.

That was the general idea, but each step presented a range of practical problems of a sort that will be familiar to anyone who has worked in this kind of field. It was very difficult to rid the apparatus of unwanted gases that were emitted from the various metals from which the apparatus was made. Molecules of these gases that were ionised by the electron beam could coagulate on the electrodes and cause spurious electric potentials. Our American rivals found that gold-plating the electrodes helped greatly to minimise these problems. We found that coating them with a carbon-based solvent called "aquadag" was a big help, not quite as effective as gold-plating but more in keeping with our research budget. My patience (and my research scholarship) ran out well before this experiment was made to yield significant results. I understand that a few more research students came to grief before significant results were eventually obtained. Now, thirty years later, low-energy electron spectroscopy is a pretty standard technique.

The details of my efforts, and those of my successors who were more successful, are not important. What I have said should be sufficient to illustrate what should be an uncontentious point. If experimental results constitute the facts on which science is based, then they are certainly not straightforwardly given via the senses. They have to be worked for, and their establishment involves considerable know-how and practical trial and error as well as exploitation of the available technology.

Nor are judgments about the adequacy of experimental results straightforward. Experiments are adequate, and interpretable as displaying or measuring what they are intended to display or measure, only if the experimental set-up is appropriate and disturbing factors have been eliminated. This in turn will require that it is known what those disturbing factors are and how they can be eliminated. Any inadequacies in the relevant knowledge about these factors could lead to inappropriate experimental measures and faulty conclusions. So there is a significant sense in which experimental facts and theory are interrelated. Experimental results can be faulty if the knowledge informing them is deficient or faulty.

A consequence of these general, and in a sense quite mundane, features of experiment is that experimental results are fallible, and can be updated or replaced for reasonably straightforward reasons. Experimental results can become outmoded because of advances in technology, they can be rejected because of some advance in understanding (in the light of which an experimental set-up comes to be seen as inadequate) and they can be ignored as irrelevant in the light of some shift in theoretical understanding. These points and their significance are illustrated by historical examples in the next section.

Transforming the experimental base of science: historical examples

Discharge tube phenomena commanded great scientific interest in the final quarter of the nineteenth century. If a high voltage is connected across metal plates inserted at each end of an enclosed glass tube, an electric discharge occurs, causing various kinds of glowing within the tube. If the gas pressure within the tube is not too great, streamers are produced, joining the negative plate (the cathode) and the positive plate (the anode). These became known as cathode rays, and their nature was a matter of considerable interest to scientists of the time. The German physicist, Heinrich Hertz, conducted a series of experiments in the early 1880s intended to shed light on their nature. As a result of these experiments Hertz concluded that cathode rays are not beams of charged particles. He reached this conclusion in part because the rays did not seem to be deflected when they were subjected to an electric field perpendicular to their direction of motion as would be expected of a beam of charged particles. We now regard Hertz's conclusion as false and his experiments inadequate. Before the century had ended, J. J. Thomson had conducted experiments that showed convincingly that cathode rays are deflected by electric and magnetic fields in a way that is consistent with their being beams of charged particles and was able to measure the ratio of the electric charge to the mass of the particles.

It was improved technology and improved understanding of the situation that made it possible for Thomson to improve on and reject Hertz's experimental results. The electrons that constitute the cathode rays can ionise the molecules of the gas in the tube, that is, displace an electron or two from them so that they become positively charged. These ions can collect on metal plates in the apparatus and lead to what, from the point of view of the experiments under consideration, are spurious electric fields. It was presumably such fields that prevented Hertz producing the deflections that Thomson was eventually to be able to produce and measure. The main way that Thomson was able to improve on Hertz's efforts was to take advantage of improved vacuum technology to remove more gas molecules from the tube. He subjected his apparatus to prolonged baking to drive residual gas from the various surfaces within the tube. He ran the vacuum pump for several days to remove as much of the residual gas as possible. With an improved vacuum, and with a more appropriate arrangement of electrodes, Thomson was able to establish the deflections that Hertz had declared to be non-existent. When

Thomson allowed the pressure in his apparatus to rise to what it had been in Hertz's, Thomson could not detect a deflection either. It is important to realise here that Hertz is not to be blamed for drawing the conclusion he did. Given his understanding of the situation, and drawing on the knowledge available to him, he had good reasons to believe that the pressure in his apparatus was sufficiently low and that his apparatus was appropriately arranged. It was only in the light of subsequent theoretical and technological advances that his experiment came to be seen as deficient. The moral, of course, is this: who knows which contemporary experimental results will be shown to be deficient by advances that lie ahead?

Far from being a shoddy experimentalist, the fact that Hertz was one of the very best is borne out by his success in being the first to produce radio waves in 1888, as the culmination of two years of brilliant experimental research. Apart from revealing a new phenomenon to be explored and developed experimentally, Hertz's waves had considerable theoretical significance, since they confirmed Maxwell's electromagnetic theory, which he had formulated in the mid-1860s and which had the consequence that there be such waves (although Maxwell himself had not realised this). Most aspects of Hertz's results remain acceptable and retain their significance today. However, some of his results needed to be replaced and one of his main interpretations of them rejected. Both of these points illustrate the way in which experimental results are subject to revision and improvement.

Hertz was able to use his apparatus to generate standing waves, which enabled him to measure their wavelength, from which he could deduce their velocity. His results indicated that the waves of longer wavelength travelled at a greater speed in air than along wires, and faster than light, whereas Maxwell's theory predicted that they would travel at the speed of light both in air and along the wires of Hertz's apparatus. The results were inadequate for reasons that Hertz already suspected. Waves reflected back onto the apparatus from the walls of the laboratory were causing unwanted interference. Hertz (1962, p. 14) himself reflected on the results as follows:

> The reader may perhaps ask why I have not endeavored to settle the doubtful point myself by repeating the experiments. I have indeed repeated the experiments, but have only found, as might be expected, that a simple repetition under the same conditions cannot remove the doubt, but rather increases it. A definite decision can only be arrived at by experiments carried out under more favorable conditions. More favorable conditions here mean larger rooms, and such were not at my disposal. I again emphasize the statement that care in making the observations cannot make up for want of space. If the long waves cannot develop, they clearly cannot be observed.

Hertz's experimental results were inadequate because his experimental set-up was inappropriate for the task in hand. The wavelengths of the waves investigated needed to be small compared with the dimensions of the laboratory if unwanted interference from reflected waves was to be removed. As it transpired, within a few years experiments were carried out "under more favorable conditions" and yielded velocities in line with the theoretical predictions.

A point to be stressed here is that experimental results are required not only to be adequate, in the sense of being accurate recordings of what happened, but also to be appropriate or significant. They will typically be designed to cast light on some significant question. Judgments about what is a significant question and about whether some specific set of experiments is an adequate way of answering it will depend heavily on how the practical and theoretical situation is understood. It was the existence of competing theories of electromagnetism and the fact that one of the major contenders predicted radio waves travelling with the speed of light that made Hertz's attempt to measure the velocity of his waves particularly significant, while it was an understanding of the reflection behavior of the waves that led to the appreciation that Hertz's experimental set-up was inappropriate. These

particular results of Hertz's were rejected and soon replaced for reasons that are straightforward and non-mysterious from the point of view of physics.

As well as illustrating the point that experiments need to be appropriate or significant, and that experimental results are replaced or rejected when they cease to be so, this episode in Hertz's researches and his own reflections on it clearly bring out the respect in which the rejection of his velocity measurements has nothing whatsoever to do with problems of human perception. There is no reason whatsoever to doubt that Hertz carefully observed his apparatus, measuring distances, noting the presence or absence of sparks across the gaps in his detectors, and recording instrument readings. His results can be assumed to be objective in the sense that anyone who repeats them will get similar results. Hertz himself stressed this point. The problem with Hertz's experimental results stems neither from inadequacies in his observations nor from any lack of repeatability, but rather from the inadequacy of the experimental set-up. As Hertz pointed out, "care in making the observations cannot make up for want of space". Even if we concede that Hertz was able to establish secure facts by way of careful observation, we can see that this in itself was insufficient to yield experimental results adequate for the scientific task in question.

The above discussion can be construed as illustrating how the acceptability of experimental results is theory-dependent, and how judgments in this respect are subject to change as our scientific understanding develops. This is illustrated at a more general level by the way in which the significance of Hertz's production of radio waves has changed since Hertz first produced them. At that time, one of the several competing theories of electromagnetism was that of James Clerk Maxwell, who had developed the key ideas of Michael Faraday and had understood electric and magnetic states as the mechanical states of an all-pervasive ether. This theory, unlike its competitors, which assumed that electric currents, charges and magnets acted on each other at a distance and did not involve an ether, predicted the possibility of radio waves moving at the speed of light. This is the aspect of the state of development of physics that gave Hertz's results their theoretical significance. Consequently, Hertz and his contemporaries were able to construe the production of radio waves as, among other things, *confirmation of the existence of an ether*. Two decades later the ether was dispensed with in the light of Einstein's special theory of relativity. Hertz's results are still regarded as confirming Maxwell's theory, but only a rewritten version of it that dispenses with the ether, and treats electric and magnetic fields as real entities in their own right.

Another example, concerning nineteenth-century measurements of molecular weights, further illustrates the way in which the relevance and interpretation of experimental results depend on the theoretical context. Measurements of the molecular weights of naturally occurring elements and compounds were considered to be of fundamental importance by chemists in the second half of the nineteenth century in the light of the atomic theory of chemical combination. This was especially so for those who favoured Prout's hypothesis that the hydrogen atom is the basic building block from which other atoms are constructed, for this led one to expect that molecular weights measured relative to hydrogen would be whole numbers. The painstaking measurements of molecular weights by the leading experimental chemists last century became largely irrelevant from the point of view of theoretical chemistry once it was realised that naturally occurring elements contain a mixture of isotopes in proportions that had no particular theoretical significance. This situation inspired the chemist F. Soddy to comment on its outcome as follows (Lakatos and Musgrave, 1970, p. 140):

> There is something surely akin to if not transcending tragedy in the fate that has overtaken the life work of this distinguished galaxy of nineteenth-century chemists, rightly revered by their contemporaries as representing the crown and perfection of accurate scientific measurements. Their hard won results, for the moment at least, appear as of little significance as the determination of the average weight of a collection of bottles, some of them full and some of them more or less empty.

Here we witness old experimental results being set aside as irrelevant, and for reasons that do not stem from problematic features of human perception. The nineteenth-century chemists involved were "revered by their contemporaries as representing the crown and perfection of accurate scientific measurement" and we have no reason to doubt their observations. Nor need we doubt the objectivity of the latter. I have no doubt that similar results would be obtained by contemporary chemists if they were to repeat the same experiments. That they be adequately performed is a necessary but not sufficient condition for the acceptability of experimental results. They need also to be relevant and significant.

The points I have been making with the aid of examples can be summed up in a way that I believe is quite uncontentious from the point of view of physics and chemistry and their practice. The stock of experimental results regarded as an appropriate basis for science is constantly updated. Old experimental results are rejected as inadequate and replaced by more adequate ones, for a range of fairly straightforward reasons. They can be rejected because the experiment involved inadequate precautions against possible sources of interference, because the measurements employed insensitive and outmoded methods of detection, because the experiments came to be understood as incapable of solving the problem in hand, or because the questions they were designed to answer became discredited. Although these observations can be seen as fairly obvious comments on everyday scientific activity, they nevertheless have serious implications for much orthodox philosophy of science, for they undermine the widely held notion that science rests on secure foundations. What is more, the reasons why it does not has nothing much to do with problematic features of human perception.

Experiment as an adequate basis for science

In the previous sections of this chapter I have subjected to critical scrutiny the idea that experimental results are straightforwardly given and totally secure. I have made a case to the effect that they are theory-dependent in certain respects and fallible and revisable. This can be interpreted as a serious threat to the idea that scientific knowledge is special because it is supported by experience in some especially demanding and convincing way. If, it might be argued, the experimental basis of science is as fallible and revisable as I have argued it to be, then the knowledge based on it must be equally fallible and revisable. The worry can be strengthened by pointing to a threat of circularity in the way scientific theories are alleged to be borne out by experiment. If theories are appealed to in order to judge the adequacy of experimental results, and those same experimental results are taken as the evidence for the theories, then it would seem that we are caught in a circle. It would seem that there is a strong possibility that science will not provide the resources to settle a dispute between the proponents of opposing theories by appeal to experimental results. One group would appeal to its theory to vindicate certain experimental results, and the opposing camp would appeal to its rival theory to vindicate different experimental results. In this section I give reasons for resisting these extreme conclusions.

It must be acknowledged that there is the possibility that the relationship between theory and experiment might involve a circular argument. This can be illustrated by the following story from my schoolteaching days. My pupils were required to conduct an experiment along the following lines. The aim was to measure the deflection of a current-carrying coil suspended between the poles of a horseshoe magnet and free to rotate about an axis perpendicular to the line joining the poles of the magnet. The coil formed part of a circuit containing a battery to supply a current, an ammeter to measure the current and a variable resistance to make it possible to adjust the strength of the current. The aim was to

note the deflection of the magnet corresponding to various values of the current in the circuit as registered by the ammeter. The experiment was to be deemed a success for those pupils who got a nice straight-line graph when they plotted deflection against current, revealing the proportionality of the two. I remember being disconcerted by this experiment, although, perhaps wisely, I did not transmit my worry to my pupils. My worry stemmed from the fact that I knew what was inside the ammeter. What was inside was a coil suspended between the poles of a magnet in such a way that it was deflected by a current through the coil causing a needle to move on the visible and evenly calibrated scale of the ammeter. In this experiment, then, the proportionality of deflection to current was already presupposed when the reading of the ammeter was taken as a measure of the current. What was taken to be supported by the experiment was already presupposed in it, and there was indeed a circularity.

My example illustrates how circularity can arise in arguments that appeal to experiment. But the very same example serves to show that this need not be the case. The above experiment could have, and indeed should have, used a method of measuring the current in the circuit that did not employ the deflection of a coil in a magnetic field. All experiments will presume the truth of some theories to help judge that the set-up is adequate and the instruments are reading what they are meant to read. But these presupposed theories need not be identical to the theory under test, and it would seem reasonable to assume that a prerequisite of good experimental design is to ensure that they are not.

Another point that serves to get the "theory-dependence of experiment" in perspective is that, however informed by theory an experiment is, there is a strong sense in which the results of an experiment are determined by the world and not by the theories. Once the apparatus is set up, the circuits completed, the switches thrown and so on, there will or will not be a flash on the screen, the beam may or may not be deflected, the reading on the ammeter may or may not increase. We cannot make the outcomes conform to our theories. It was because the physical world is the way it is that the experiment conducted by Hertz yielded no deflection of cathode rays and the modified experiment conducted by Thomson did. It was the material differences in the experimental arrangements of the two physicists that led to the differing outcomes, not the differences in the theories held by them. It is the sense in which experimental outcomes are determined by the workings of the world rather than by theoretical views about the world that provides the possibility of testing theories against the world. This is not to say that significant results are easily achievable and infallible, nor that their significance is always straightforward. But it does help to establish the point that the attempt to test the adequacy of scientific theories against experimental results is a meaningful quest. What is more, the history of science gives us examples of cases where the challenge was successfully met.

Further reading

The second half of Hacking (1983) was an important early move in the new interest philosophers of science have taken in experiment. Other explorations of the topic are Franklin (1986), Franklin (1990), Galison (1987) and Mayo (1996), although these detailed treatments will take on their full significance only in the light of chapter 13, on the "new experimentalism". The issues raised in this chapter are discussed in a little more detail in Chalmers (1984).

CHAPTER 4

Deriving theories from the facts: induction

Introduction

In these early chapters of the book we have been considering the idea that what is characteristic of scientific knowledge is that it is derived from the facts. We have reached a stage where we have given some detailed attention to the nature of the observational and experimental facts that can be considered as the basis from which scientific knowledge might be derived, although we have seen that those facts cannot be established as straightforwardly and securely as is commonly supposed. Let us assume, then, that appropriate facts can be established in science. We must now face the question of how scientific knowledge can be derived from those facts.

"Science is derived from the facts" could be interpreted to mean that scientific knowledge is constructed by first establishing the facts and then subsequently building the theory to fit them. We discussed this view in chapter 1 and rejected it as unreasonable. The issue that I wish to explore involves interpreting "derive" in some kind of logical rather than temporal sense. No matter which comes first, the facts or the theory, the question to be addressed is the extent to which the theory is borne out by the facts. The strongest possible claim would be that the theory can be logically derived from the facts. That is, given the facts, the theory can be proven as a consequence of them. This strong claim cannot be substantiated. To see why this is so we must look at some of the basic features of logical reasoning.

Baby logic

Logic is concerned with the deduction of statements from