Science: Why, and How?

Some Basic Ideas in Scientific Method

1

PREVIEW In the first two chapters we will present some central concepts in the philosophy of science. We will discuss what knowledge is, and how knowledge claims might be justified. Almost every concept in the field has been the subject of intense debate, and while we cannot introduce these without some philosophical discussion, we will try here to present the consensus with respect to fundamental concepts in the theory of science. Chapter 3 deals with different and sometimes conflicting philosophical views and ideals of science.

After a first primer on the nature of science (1.1), the second section (1.2) introduces two extreme views on the nature of scientific knowledge and the scope of truth: realism and idealism. These epistemological positions have been brought to bear upon the question of how much objectivity science can claim: the present authors' answer is an intermediate position, pragmatism. A further question is how scientific knowledge may be different from common sense.

In Chapter 1.3 a number of epistemological concepts, that is concepts we use in our knowledge claims – such as deductive and inductive arguments; laws, theories and facts; justification and discovery of theories – will be discussed. The nature of causal laws is the subject of Chapter 1.4 where we explore some aspects of typically scientific knowledge, such as explanations, laws, observations, and causes.

1.1 INTRODUCTION: WHY SCIENCE?

Demarcating science

In modern societies, science is held in high regard, and the results of scientific research seem almost unconditionally trusted. Laboratory tests count as a guarantee of the quality of drugs, food and cosmetics: logic and mathematics are the hallmarks of certainty and objectivity. No one seems to question the almost magical ability of scientists to estimate the safety of a new nuclear plant in terms of the probability of an accident per million years and health scientists will specify the effects of smoking, overeating, and pollution in percentages. Science is apparently seen as the embodiment of rationality, objectivity and truth (or at least our best approximation of the truth). Most people believe that science and technology have steered us on our way to more welfare, health, freedom and prosperity (Toulmin, 1990) and in our society common sense yields to scientific knowledge: psychological testing takes the place of empathy, evidence-based medicine replaces lore and intuition. Briefly, over the past four hundred years (or so), scientific thinking and research have proved a huge success. 'Unscientific' and 'pseudo-science' are (almost) terms of abuse, although there can of course be doubts about the validity of certain pieces of research, such as the suspicion that in climate science data have been manipulated, or in drugs research business interests have distorted published results, and individual investigators have indeed been caught faking or embellishing their data. But these cases are more like aberrations from, or abuse of, a basically correct model, than real worries about the status of science as such.

Strangely enough, philosophers of science have as yet failed to find out exactly what defines science and its methods, what accounts for its success, and how to make an airtight separation between science and pseudo-science. Even more surprising, in light of the omnipresence of science nowadays, some philosophers reject the idea of a difference in principle between science and other social activities: its alleged objectivity is just self-congratulation on the part of the establishment – whatever is accepted as truth is determined by power and propaganda. These philosophers consider the practice of scientific inquiry the subject matter of sociology (see Chapter 5), to be explained in the same way as one might study primitive tribes or groups like Hell's Angels: acceptance of theories is governed by 'mob psychology' rather than by objective 'scientific' criteria (see Chapters 4.3 and 4.5). These social and anarchistic approaches tend towards **relativism**; one theory is as good as the next one, and preferences towards any scientific approach are due to arbitrary, irrational factors.

We will defend the view that scientific practice is *not* arbitrary and that scientific knowledge has a legitimate claim to **truth**; that it, in a way, corresponds to an external reality, while at the same time we would recognize that it is subject to a host of social, pragmatic and sometimes irrational influences, and that scientific truth is not something separate from human concerns.

Unification and underlying causes

An impressive feature of science is that it can explain disconnected phenomena as the effects of underlying causal structures. A bewildering variety of chemical reactions, for instance, can elegantly and parsimoniously be explained within the framework of Mendeleyev's table of elements, which in turn is explained by the composition of chemical elements (atoms, consisting of electrons, neutrons and protons) governing binding and so on. A good example in psychology (although a controversial theory) would be psychoanalysis, which shows how underlying traumas produce neurotic behaviour. Theories unify and systematize knowledge. Everyday phenomena are *reduced* to something more basic; they can be fitted into a comprehensive theory, and in that way can be explained and predicted (and manipulated in laboratories). That, of course, is a major triumph of science. Reduction also has the somewhat disturbing consequence that these everyday phenomena are 'really' nothing but atoms and molecules, that thoughts are nothing but physiological mechanisms, etc. Explaining, for instance, the physiological mechanisms of consciousness or memory seems tantamount to eliminating the interesting aspects of mental life, and reducing real people like you and me to drab machines (for the problem of reduction, see Chapter 2.6).

'Criticism': keeping an open mind

Another meeting point for science and culture is that science has been associated with a critical attitude, open-mindedness, and Western liberal democracy (e.g., Popper, 1966). Historically, the rise of empirical critical investigation and the rejection of authority have gone together on several occasions. It was characteristic of Protestantism and the associated New Learning movement in England in the middle of the seventeenth century, which combined a politically progressive (if not subversive) demand for freedom of speech and the press with the development of science, mathematics and medicine (see Schafer, 1983). It could be argued that science is a characteristic of modern society: a rejection of dogma, a critical attitude towards authority, the feeling that in thinking for oneself an individual can find the truth. Others, however, see science as the stronghold of political oppression: it has been identified (especially in psychology and the human sciences) by, among others, Marxists and feminists with repressing human concerns in a methodological straitjacket, providing the establishment with ideological legitimations and/or the technological means for maintaining the status quo of capitalist exploitation and unthinking technocratic dominance (Marcuse, 1964; Weizenbaum, 1976; see also Chapter 5).

More generally, science has been part of the project of modernity (e.g., Toulmin, 1990), seeking rational criteria for conduct in a wide range of human activities. Thus, it has received its share of postmodernist criticism, which rejects the idea of universal criteria for rationality (Feyerabend, 1975; Rorty, 1979).

This kind of debate on the proper place of science, its limitations and strengths, may be elucidated (if not decided) against the background of a principled account of the nature and limitations of knowledge, and of scientific knowledge in particular. Philosophers call the special branch of their trade that deals with evaluating the claims of knowledge **epistemology**.

1.2 KNOWLEDGE: REALISM AND IDEALISM (RELATIVISM), COMMON SENSE AND SCIENCE

Realism

Knowledge is, according to most authors, justified true belief. Of course, one may ask how to fixate beliefs, and what constitutes justification. In broad outline, two possible grounds for justification have been proposed: one, idealism, focusing on the knower, the individual or social processes leading to knowledge claims; the other, realism, focusing on the known, the object of knowledge.

Realism says that knowledge corresponds to reality; more precisely, that the terms for our theories refer to, 'correspond' with, real things in the world. Scientific realism is probably the (largely implicit) image most working scientists would have of what empirical investigation really is (see Chapter 3). It is sometimes assumed that reality consists of elementary atomic facts, which are reflected in observation statements, plus the logical connections between them. Such observation statements thus represent elementary states of affairs ('facts') in the world, and they are connected by tautological logical rules, so that the build-up of knowledge makes a kind of mental blueprint of the world – a **theory**. Language reflects reality in a mirror-like fashion – like a picture, or perhaps more accurately, like a blueprint (see Chapter 3.3 for the early Wittgenstein's picture theory of truth). The common view of **truth** in realism is **correspondence**: theories are true if they correspond with nature. The unsolved (and unsolvable) problem, however, is that there is no measure of agreement between language and reality, if only because it would have to be put into language, in the form of a theory. Hence objectivity, in the sense of letting the world speak for itself, and objective knowledge, as gathering its reflections in the mirror of our theoretical representations, are an illusion. When we ask whether some theoretical term is objectively real – for example, whether personality traits really exist – we can only give the answer in the form of a statement.

Idealism and relativism

The alternative, **idealism**, holds that the world as we know it is somehow a creation of the mind. Our knowledge is a subjective product, and does not necessarily correspond to an outside world; it is not even clear what the concept of an outside world exactly means, if it is construed as independent of a knowing subject. Idealism tends towards **relativism**: if knowledge is a subjective construction, then every subject or every group, every historical period, or socio-economic class may have its own truth.

Idealism is the classical alternative to realism. If all knowledge is a subjective construction, there is no rational, objective way to choose between different points of view. If theories are completely in the eye of the beholder, and have no relation with reality, then anything goes. Idealists like Berkeley, Kant and Descartes were forced to introduce God or Universal Human Nature in order to arrange for some correspondence between representations in individual minds and the represented things in the world.

The common view of **truth** in idealism is **coherence**: theories are true if they are consistent with the rest of our knowledge. The idea has some plausibility in, for example, mathematics: mathematical proofs are true when derived from a theorem's axioms. Mathematics is a self-contained construction of the mind: its truth cannot be checked by empirical means – it makes no sense to start to measure actual triangles in the world to see whether their angles always add up to 180 degrees. Rather, we can deduce this result from a web of other internally cohering statements.

Relativism is a more modern term that emphasizes the collective nature and social determinants of ideas, and the impossibility of universal, objective knowledge. A position close to idealism and relativism is social constructionism: it is believed that much of science is a human construction, a reflection of social interactions in a collective of researchers and society at large, more than a reflection of the world. In Chapter 4 relativism (anti-realism) is discussed, and Chapter 5 looks at the social and psychological influences on theory choice.

The dilemma: the impossibility of 'objective' knowledge

So, we seem to have two equally unattractive options: *idealism* – where the mind makes up the world, or perhaps entirely confabulates it; or *realism* – assuming that the world, as it is in itself, independent of human exploration and theorizing, is accessible to us. The latter option, which assumes that there is some criterion for matching a 'God's eye point of view' (Putnam, 1981) with our own view, is of course paradoxical. As Rorty (1979: 298) puts it, it involves the notion 'that we are successfully representing according to Nature's own conventions of representation', rather than 'that we are successfully representing according to our own'. In other words, there is no criterion for comparing our theories directly with the world, since any such comparison must be, it seems, a theory, so that there is no way of getting beyond, or stepping outside of theory. Thus, realism in the epistemological literal sense is impossible.

Indirect support for *scientific realism* is sometimes sought in the empirical success of empirical investigations, especially in physics (Boyd, 1984). If new findings fit with and extend existing theories, and our world image seems to converge towards a final theory, then these theoretical terms (atoms, quarks, etc.) will probably correspond to something real; hence the name for this position, *convergent realism*. However, theories that are patently wrong can be quite successful predictors (Laudan, 1991). It seems that a theory's empirical success is no guarantee for truth-as-correspondence-with-the world. But, of course, success is important, as the pragmatic view emphasizes. We shall postpone a more elaborate discussion of realism and relativism until Chapter 4, but here we will offer the reader a preview of our own position.

The pragmatic view: functional knowledge

Our position in the realism–idealism dichotomy is that knowledge is interactive, is the product of actively exploring the world, and reveals reality by acting on it. In a sense this is an intermediate position between realism and idealism or relativism (Bem, 1989), between **objectivism** and subjectivism. The idea is that 'the mind and the world jointly make up the mind and the world' (Putnam, 1981: xi). The product of this conjunction is subject-relative but not subjective or relativistic in the sense of arbitrary. We call this a functional view of knowledge. It holds that knowledge is a kind of interaction of subject and object, rather than being either passive picturing or subjective constructing. Knowledge is a methodologically regulated, constrained form of human action (praxis), and therefore is evaluative and value-laden. In Rorty's (1979) words, knowing is coping with the world rather than mirroring it. Therefore, we should expect that the meaning of theoretical terms derives from their practical use, and that manipulation is a determinant in the structure of knowledge. This theory of truth is called pragmatism. More on this in Chapter 4.7 and 4.8.

To sum up, pragmatism and the notion of functional knowledge designate an interactional view of the nature of knowledge, which avoids the extremes of realism and idealism or relativism, of focusing exclusively either on the objective or on the subjective pole.

BOX 1.1 Realism, idealism and pragmatism

Realism:	Knowledge pictures the objective world. Truth is a correspondence between knowledge and the world.
Idealism:	Knowledge is a subjective (or social) construction. Truth is a coherence with the rest of knowledge.
Pragmatism:	Knowledge is functional and interactive, coping with the world. Truth is success.

Everyday knowledge and scientific knowledge

The difference between everyday knowledge and scientific knowledge is loosely related to questions concerning the nature of scientific methodology and scientific explanation, to the tension between methodological reduction and phenomenological experience, to the relation between explanation and understanding. These questions will be set out later in this chapter.

The philosopher Wilfred Sellars (1963) made the classic distinction between the manifest and the scientific image. 'Image' refers to the concept of man in the world, the framework in terms of which man views himself. The manifest image is the world of objects and persons of common sense. Common sense can be simple unquestioning acceptance of everyday things and events, but also sophisticated reflection on everyday life as in literature and philosophy. The scientific image is the world of particles and forces posited by advanced science. Thus, the difference is 'not that between an unscientific conception of man-in-the-world and a scientific one, but between that conception which limits itself to what correlational techniques can tell us about perceptible and introspectible events, and that which postulates imperceptible objects and events for the purpose of explaining correlations between perceptibles' (Sellars, 1963: 19).

So, on the one hand, there is the image of refined categories of common sense and, on the other hand, the image in terms of postulated underlying reality – and these often seem in conflict, each claiming to be the true and complete account of man in the world. The scientific image aims at replacing the manifest one; it holds that water, for instance, is really H_2O; that only the scientific table as described in physical terms is real and the common-sense table is an illusion; that a person and her or his thoughts and feelings are really neurophysiological processes.

Three ways of confronting both images suggest themselves: (1) we may assume that they are identical – this is obviously wrong, since, strictly speaking, molecules are not wet or coloured; (2) the manifest image is real, and the scientific image is only an abstract or condensed way of describing it; (3) the scientific image is real, and the manifest image is only an appearance. Sellars assumes that the scientific image is in principle adequate and true.

Sellars however goes for a fourth option: (4) – that both are real. He wants to unite two images in 'stereoscopic' vision: we should realize that science is not finished, but might progress and recreate in its own terms the concepts of the manifest image. A fine example of this approach is Dennett's (1991a) theory of consciousness, which tries to incorporate consciousness in a state-of-the-art cognitive-neurophysiological theory (see Chapter 10). However, Sellars also held a more utopian view of integrating science with the goals of a community, appropriating the world as conceived by science into a rational and meaningful way of life.

To sum up, the relation of science and common sense is often conceived as a border dispute, with science in the role of the invader. The view taken in this book is that the relation between science and common sense is a continuum, in the sense that scientific methods are a restricted and regimented outgrowth of human praxis. In our view, science can be best understood against the background of practice. A large part of Chapter 4 will be devoted to discussing a pragmatic view in the philosophy of science, as contrasted with a theory-centred view.

Some characteristics of scientific knowledge

Historically, science is no doubt continuous with the knowledge and concerns of daily life. In Western society practical problems – such as optimizing fertility in agriculture, measuring land, traditional healthcare, etc. – have more or less smoothly merged with chemistry, geometry, biochemistry. There is apparently no sharp division between pre-scientific and scientific knowledge. Science is organized common sense (Nagel, 1961).

The methodically definite form of science as we know it began at the end of the Middle Ages. Later, in seventeenth-century England and the Netherlands, the demand for practical knowledge in artillery, fortress building, irrigation, and canalization boosted the study of mathematics and physics. What distinguishes this new scientific method from previous common-sense solutions is its systematic nature, and its endeavour to provide explanations for the phenomena observed. In a nutshell, science is systematic in the sense that it tries to formulate laws that apply everywhere, not just in traditionally established habits, and is explanatory in the sense that it tries to answer 'why' questions, providing an answer to the question of why the phenomena are as observed. Such explanations are both systematic and controllable by factual evidence. As Nagel (1961: 4) puts it: '[I]t is the organization and classification of knowledge on the basis of explanatory principles that is the distinctive goal of the sciences.'

The following list of characteristics may be used to get the gist of scientific method (Nagel, 1961; Sanders et al., 1976):

1 *Systematicity*. Theories must be applicable across the board, the theoretical edifice must be coherent and if possible hierarchical; the domain of application is specified at the outset, and no ad hoc exceptions are allowed.
2 *Well-defined methods* (Kuhn, 1962). Methods also specify what will count as legitimate subject matter, facts and explananda. Psychologists, for instance, will be reluctant to investigate 'poltergeists' as phenomena in their own right; chemists will disown the philosopher's stone: they fall outside the framework, and do not count as observation.
3 *Reduction*. Both in the sense of ignoring certain aspects of reality (which are supposedly accidental) at the descriptive level, and in the sense of reducing phenomena to underlying principles at the explanatory level. As a simple example of the latter: water, steam and ice are explained as the same chemical substance under different conditions. A more complex example is that all matter may ultimately be explained by the final laws of a (future) complete physical theory in terms of elementary particles or fields.
4 *Objectivity*. In the sense of being controllable, reliable and intersubjectively observable. For instance, so-called slow schizophrenia, which could only be observed in Soviet dissidents by Soviet psychiatrists trained by Professor Snezjnevskij in KGB clinics, and nowhere and by nobody else (Joravski, 1989), is not a scientific concept: it is not replicable by others.
5 *Clarity*. Scientific statements are phrased unambiguously, in principle addressed to the public domain.
6 *Revisable*. Scientific knowledge is open, at all times revisable, and never definitive.

From this list, it will be clear that the distinction between scientific and common-sense thinking is a matter of degree, not of principle: science is just more systematic, general, methodical, open, etc. than common sense. So the dictum that science is organized common sense is reasonable, but something remains to be said about the specifics of the mode of organization. *Reduction* is the most distinguishing feature of science – more than in common sense, the aim is to find the hidden springs behind the phenomena. We devote Chapter 2.6 to this issue.

Notions of *classification* and taxonomy play an important role in reflections on scientific method. 'Cutting nature at the joints' is essential for the organization of knowledge in a systematic way. The suggestion that some term is 'merely' descriptive, and therefore unimportant and arbitrary, is certainly wrong. In biology the choice of taxonomy (mammals, reptiles, fish, insects) is a *sine qua non* for a viable science; the classification of whales as mammals, rather than fish, is no trivial or linguistic matter, but an essential feature of the systematic nature of science.

Classical accounts emphasize explanation as the hallmark of science (e.g., Nagel, 1961; Rosenberg, 2005; see also Chapter 2), describing the underlying mechanisms that account for or cause surface phenomena. The explanatory aspect of science can be seen in its extensive use of unobservables – underlying explanatory entities (like atoms, or the Freudian unconscious) that try to explain the observed phenomena. This may sometimes have the unpleasant consequence of parting company with the layman's view, as discussed above. Furthermore, systematicity implies that phenomena are isolated into small and unambiguously observable units, with the aim of subsequently integrating them into a larger system of facts. Science attempts to provide a logically unified body of knowledge, ideally in the form of a closed, axiomatic, deductive system (see Chapter 3 for the logical positivists, who pioneered this view of theories) in which propositions can be derived from theories describing empirical facts. A good example of this would be Mendeleyev's system (the table of chemical elements).

In its *testability* science also goes beyond common-sense knowledge; common sense employs broad and relatively fuzzy concepts, whereas science refines these into precise notions. The greater determinacy of scientific concepts contrasts with the loose generalizations elsewhere, and allows for more rigorous testing. It makes knowledge claims more vulnerable, but also provides more opportunities for neatly fitting these into a larger, clearly articulated, coherent theory. Thus, previously unconnected facts can be related and systematized. Common sense is relatively dependent upon unchanging conditions and a number of unarticulated background assumptions, whereas science is explicit as to its assumptions. Scientific knowledge is systematic and coherent in the way that everyday knowledge is not and, unlike common sense, it is explicit about the range of applications for its concepts. Science avoids the inconsistencies common sense is not concerned about and tries to build an homogeneous network of concepts.

In our view, then, science is to be considered from a pragmatic perspective. Its methods are to be evaluated with respect to its central aim of producing knowledge about the world and finding generalizations ('laws') that apply to it (Chalmers, 1990). Thus we can circumvent, at least for practical purposes, the unsettling problems of relativism. Briefly, the view that pragmatism is relativist and irrationalist, and cannot distinguish between accidental success and genuine scientific rationality, only follows on from the hidden assumption that a philosophical account of how knowledge is anchored in some form of contact with the world is the only defence against absurdity and fraud. We think that there is no need for a single, fixed, ahistorical canon of scientific method. Knowledge about the world comes in many varieties, and should be evaluated pragmatically, in the light of practice.

1.3 ARGUMENTS: DEDUCTION, INDUCTION, ABDUCTION

The objective of knowledge is to understand (parts of) the world in order to get on in it. Science is a special branch of knowledge; it is usually not content with the immediate environment, and probes deeper than common, everyday knowledge – and most importantly, science is a controlled enterprise. Scientists want to comprehend *why* things happen, what are the mechanisms or processes *behind* the phenomena. To form their opinions, to convince others, to provide evidence and to predict events they will use different means – assumptions, observations, arguments, explanations, predictions, descriptions, theories, models. In this and the following sections we will discuss some of these basic scientific concepts.

Deductive arguments

Arguments are sets of statements (the premises) connected in such a way that a conclusion results from them. In some arguments, the conclusion will be definitively supported. An example here is:

Men are bigger than mice

Mice are bigger than ants

Thus: Men are bigger than ants

Because the premises ‘contain’, so to speak, the conclusion, or the conclusion can be ‘extracted’ or deduced from the premises, these conclusive arguments or inferences are also called *deductive.* If you accept the premises of a deductive argument then you also must buy into the conclusion; otherwise you will produce a contradiction. The soundness of conclusive arguments is a consequence of the meaning of the relations (in the example, 'are bigger than') and the arrangement of the terms (the names which stand for the subjects). This pattern can be abstracted from inferences about specific states of affairs, and can be *formalized* as follows:

$$\begin{array}{l} x\ R_t\ y \\ y\ R_t\ z \\ \hline \text{Thus: } x\ R_t\ z \end{array}$$

where R_t stands for a transitive relation, such as 'bigger than', 'smaller than', 'older than'. So, we have here a valid inference-pattern which can be interpreted by any transitive relation between two different entities. This is what logicians do (among other things): they abstract from or generalize about specific arguments and study under what conditions arguments are valid, in what respect they are similar or different, etc. Unlike other scientists and people in their everyday discourse, a formal logician is not interested in the subject matter or content of arguments, only in their formal structure.

Among conclusive arguments *syllogisms* are well known. Here is an example:

All politicians are liars

All members of parliament are politicians

Thus: All members of parliament are liars

In this example it is easy to see that a conclusive (deductive) argument that is perfectly valid can be doubtful or even false, because the first of the premises is. The truth has something to do with the content of the inference; the soundness with the pattern or form. So, the conclusion of a deductive inference is true under two conditions: (1) the argument must be valid or sound; and (2) the premises must be true. In other words, if an argument is deductively valid, it is impossible for the premises to be true while the conclusion is false.

This promise of absolute certainty constitutes the appeal of the deductive method. Ideally, one could start with a few unquestionable truths or axioms and then deduce other statements or theorems from them. The geometry devised by Euclid, the Greek geometer who lived in the third century BC, was such an *axiomatic system.* Very much impressed by the elegance of this system, the French philosopher René Descartes (1596–1650) thought it possible to deduce all scientific statements from some axioms, which were placed within us as innate ideas by God. Many other scientists in the sixteenth and seventeenth centuries also thought that nature was mathematically structured, that the world was a machine, or clockwork, working according to precise mathematical laws, that the human mind was designed in accordance with that system and could comprehend it, and that knowledge reflected that system. In our century the behaviourist Clark Hull (1884–1952) had a similar ideal in mind for psychology. Psychology, he thought, was a natural science and since nature was a mathematical and mechanical system the mental was nothing but physical and behavioural, and psychology could be formalized into one single deductive system. In the end behaviour was the complex result of basic physical entities like electrons and protons. Since psychology made up a chapter of the entire scientific system, theories and predictions about behaviour could be deduced from clearly stated principles (Leahey, 2001).

However, Hull's system didn't work out. Apparently, we cannot put our knowledge so rigorously and absolutely into a comprehensive, unified and fixed system. Scientific theories happen to be fallible and changeable. As we said before, science is never closed.

Inductive arguments

We cannot rely on deductive arguments exclusively; in science as well as in everyday discourse we will mostly apply non-conclusive arguments. While the premises of a conclusive argument already logically 'contain' the conclusion, which therefore must be accepted, the conclusion of a non-conclusive argument is only, more or less, supported by the premises. If you do accept the premises, but still doubt the conclusion, you could be reproached for being stubborn or an arch-sceptic, but you cannot be reproached for contradicting yourself. Among the non-conclusive arguments are *inductive* arguments which are generalizations from statements of lesser scope: what is true of a number of members of a class is likely to be true of all members. Here is an example:

> I know five psychologists and, boy, are they arrogant!
>
> Therefore, I think that all psychologists are arrogant.

If someone said this he had perhaps not been very fortunate in his meetings with psychologists, and was rather hasty in drawing this general conclusion from the poor sample, because there are thousands of psychologists. Suppose, however, that the sample of arrogant psychologists is not five but 100, or 1,000: wouldn't we be moving from less to more evidential support? Thus, inductive support for the conclusion comes by degrees: it depends on the amount of evidence in relation to the extent of the conclusion, and it varies with different types of subject matter. It is reasonable therefore that people who are confronted with inductive conclusions should want to know the weight of the evidence. In order to accept, for instance, the assertion that frequent use of marijuana or hashish every day will impair your memory, one wants to know how the study which constitutes the evidence has been done, how many subjects have been examined (the sample), what data have been gathered, etc. A subclass of inductive arguments are *statistical* arguments in which the degree of probability is given in numbers or percentages; often you will find non-numerical terms such as 'many', 'nearly all', or 'never' in the conclusion.

The 'problem' of induction

In an inductively strong argument, then, if the premises are true, it is only probable that the conclusion is true. Some logicians think it better to speak of a successful induction not as a valid but as a strong argument and they will reserve the notion of validity for deductive arguments. No matter how strong the inductive reasoning, it will always be an inconclusive argument: the conclusion will always go beyond the evidence. For this reason philosophers of science have had a love–hate relationship with induction. On the one hand, it is acknowledged that inductive arguments provide new empirical knowledge, that is, knowledge that is not already contained in the premises as with deduction. Science is, to a large extent, empirical and inductive: it generalizes from observed instances and it predicts by inferring what will happen from what has happened. This, one could say, has contributed to scientific successes. But on the other hand, it does not provide the ardently desired certainty, one cannot anticipate future cases or predict with certainty, and one has seldom witnessed all the cases in the past. There is always room for *scepticism,* as the empiricist David Hume wrote:

> *That the sun will not rise tomorrow* is no less intelligible a proposition, and implies no more contradiction than the affirmation, *that it will rise*. (1963 [1748]: section iv, 25–6, original emphasis)

On reflection you would perhaps point to the general presupposition upon which expectations of natural events are based – that the course of nature is uniform and continuous. This means that, other things being equal, if nothing interferes, nature will operate in the same way. But, replies Hume, that again gives you an inductive inference: up until now nature behaved … etc. And then you realize that you are merely begging the question and you are going in circles.

Hume's conclusion of his discussion of induction is sceptical and negative: inductive arguments cannot be justified by (logical) reasoning; there is no rational foundation for them. There is no cogent line of reasoning that leads from premises to a conclusion, no absolute certainty in the manner of deduction. We arrive at inductive conclusions via a non-rational process: by habit. The process of inference is not logical thinking but a psychological step. We are used to the fact that the sun rises every morning, and the prediction that it will rise tomorrow is not the conclusion of a rational – read: logical – argument but a psychologically understandable expectation. For Hume 'rational' is deductive certainty and, except in mathematics, most scientific reasoning is 'merely' inductive.

This lack of certainty or, more precisely, the suspicion that scientific inference is not justifiable, and consequently that science is unfounded, has been called *the problem of induction*. Philosophers have been trying to find a logic of inductive justification. A classic example illustrates why this never worked out: the 'Raven paradox'. If we are to inductively confirm the hypothesis that all ravens are black, we must list all ravens and check whether these are all black. However, logically $(x)\ (rx \rightarrow bx)$ (for every x, if it is a raven, it is black) is equivalent with $(x)\ (-bx \rightarrow -rx)$ (for every x, if it is not a raven, it is not black). So, observing non-black things that are not ravens confirms that ravens are black; seeing a pair of white sneakers corroborates the blackness of ravens. That result is bizarre of course, but logically impeccable, and the conclusion must be that the logic is not working for induction – induction cannot be logically justified in the way deduction can.

There is an even deeper problem with induction: we have to start with concepts and criteria to gather observations, and in particular, criteria for similarity. In order to generalize, we need to know what counts as instances of the same and what does not: we should be able to tell a swan from a flamingo, and a mammal from a fish. Whether a flamingo is a pink swan or a separate species, and whether a whale is a fish or a mammal, will depend on one's assumptions at the start. And where could these come from? If from observation, then we will go round in a circle. In addition, we will have to know in advance what is relevant: just listing everything we can see (the colour of the clouds, the balance on our bank account) won't do. On the other hand, in principle anything might bear upon anything (maybe details about cell biology can help decide whether there is life on Mars), but unfortunately there seems to be no rule for deciding in advance what does (Quine, 1969b).

Some scientists think the problem can be ignored, saying that only if you think that science has to search for absolute certainty, only if you think that truth has to be absolute, do you have a problem because you are asking too much: it is enough that science scores its successes without strong logical justification. This is the pragmatic response (see Quine, 1969b). Others will also deny that there is a problem, saying that not only science but also life would be impossible if a zillion everyday expectations were in need of a strong foundation, such as if I collide with that tree my car gets smashed up. Some philosophers of science, however, are not satisfied with the idea that the factual success of science and scientific reasoning lacks justification, and hence is simply fortuitous.

Inference to the best explanation

Inconclusive reasonings are often used in explanations. Arriving home you find the fridge door open. Since your partner, the only other inhabitant, has lately developed the bad habit of showing some negligence in this matter, you conclude that once again he has not closed it. Of all the possible other explanations this is the best. But is this conclusion more logically legitimate than others? What you did was arrive at a hypothesis. The question is: how do we do this, and on what basis?

This kind of probably reliable explanation, inference to the best **explanation** as it is called today, the pragmatist C.S. Peirce once christened **abduction**. It is a kind of reasoning in which an explanatory hypothesis is derived from a set of facts and has the following structure:

If S is the case then R

R is the case

Therefore, it is possible that S was the case

(If it has rained, the streets are wet

The streets are wet

It may have rained)

Note that this is not a logical certainty, just a possibility (a water pipe may have burst). Peirce was very much interested in the testing of these hypotheses and he tried to construct a logic for it. How do we arrive at hypotheses? And what criteria can we legitimately use for testing them? This logic of testing and of finding rules and criteria for hypotheses is also called the *logic of discovery*.

Context of justification and context of discovery

Some philosophers, however, denied that there can be a logic of discovery because discoveries are too different and complex to be captured in logical and methodological rules. There is no algorithm for discovery, or a recipe that inevitably and mechanically leads to new facts and generalizations – as the induction problem shows. And they contended that the acquisition of scientific facts or theories is not the business

BOX 1.2 Induction, deduction and abduction

Induction: from individual observations to general statements.
No logical certainty, but new knowledge.
Example: Lots of swans were observed, all were white
Maybe all swans are white

Deduction: from general statements to individual observations.
Logical certainty, because the conclusion is contained in premises: no new knowledge.
Example: All humans are mortal
Socrates is human
Therefore, Socrates is mortal

Abduction: inference to the best explanation.
No logical certainty, new hypothetical knowledge about causes.
Example: All CJD patients ate beef
Beef may be the cause of CJD

of philosophy but of psychology. All kinds of extrascientific factors induce discoveries, such as Archimedes sitting in his bath and discovering the way to calculate the volume of solid objects, and Newton guessing the law of gravitation by observing an apple falling from a tree. These philosophers consider it the task of the philosophy and methodology of science to guard the rationality of science and analyse whether the scientific products, the finished theories, can be justified, abstracting from the messy ways in which scientists arrive at their conjectures: What is the argumentative basis? What are the empirical data? How strong are the logical connections between the statements? What are the norms for good theories? Romantic flashes of insight and other personal histories are non-rational or irrational and irrelevant to the task of justification.

This led to a distinction between the **context of justification** and the **context of discovery**, introduced by traditional empiricist philosophers to demarcate the domain of scientific rationality (for which only the context of justification is relevant). Others, in direct opposition, demonstrate the importance of the historical, social and psychological contexts of scientific discoveries (Thomas Kuhn, 1970; see Chapter 3.7;

see also Chapters 4 and 5). Apart from the question of whether there is a discovery algorithm, they argue that:

> ... to ignore discovery, innovation, and problem solving in general is to ignore most of the scientists' activities and concerns, in many cases not only the most interesting phases of scientific research but also (more importantly) phases highly relevant to epistemology, e.g., to the theory of rationality and the understanding of conceptual change and progress in science. (Nickles, 1980: 2)

In fact, sociologists and psychologists took over segments of the epistemological domain that philosophers traditionally claimed for themselves – guarding rationality and setting the rules for scientific method. This takeover is part of what has been called *naturalistic epistemology*. In this project, initiated by the philosopher W.V.O. Quine, epistemology is seen as a part of natural science because it 'simply falls into place as a chapter of psychology' (Quine, 1969a: 82). It can be contested whether psychology is entirely a natural science, but if one takes 'naturalistic' in a broader sense, meaning 'continuous with science', one might perhaps agree with our suggestion that the sociology of science, psychology of science and/or psychology of cognition are legitimate chapters in the programme of naturalizing epistemology. We consider this project as an inquiry into the processes by which scientists tend to arrive at their scientific beliefs. In the various chapters that follow we will pursue this line of thinking.

BOX 1.3 Context of discovery and context of justification

Context of discovery

In this context the focus is on a description of the historical, social and psychological circumstances and influences that were relevant to the invention or discovery of scientific theories. Historians and sociologists of science try to find out under what conditions science works.

Context of justification

In this context the focus is on normative criteria for holding a theory to be true, or acceptable, or justified. Philosophers of science will try to develop general methodological requirements for a scientific theory, for example, the degree to which the conclusions are empirically or logically supported (induction, deduction).

In the traditional view, philosophy of science is only about justification, not about the social or psychological circumstances of the problem-solving situation.

1.4 LAWS, THEORIES, MODELS AND CAUSES

Empiricism: pure observation?

It is sometimes said that the job of science is to discover facts. This has to be qualified, however. The **empiricist** Francis Bacon (1561–1626) thought that collecting facts like a bee gathers honey was the right method for doing science: doing research is systematically collecting observations and compiling lists of data, and if scientists do that carefully the scientific laws will be discovered automatically. However, it is highly implausible that science has ever been undertaken in such a way because it is not an automatic process at all. One always departs from preconceived ideas when gathering data. We cannot do science without some power of imagination, without some idea of what to look for. For Bacon, however, imagination and fantasy constituted dangers for science, which should eschew prejudices ('idols'), and he put all his money on 'pure' empirical facts.

Thus direct, 'pure' observations are a problem. There is a tension between observation and theory (meaning here, going beyond direct observations) that has always haunted philosophers of science. It is, of course, a major concern of science to understand what happens and what will happen. To this end scientists have to generalize about relations between different facts. We saw earlier that empiricists reached the view that inductive reasoning was highly problematic. The philosophers of science whose idea of science was strictly empirical (observation as the foundation of inquiry) had to think hard about these problems, and also accept a certain amount of uncertainty (see Chapter 3 on the Logical Positivists). Observations cannot be strictly objective but perhaps intersubjective agreement is possible, and the ideal of exact observation statements as the foundation of theories has remained.

Observation and unobservables

However, in order to comprehend underlying structures, to formulate 'laws' of nature, scientists have to venture beyond the mere inspection, enumeration and description of what can be observed. This very often makes them decide to conjecture *unobservable entities* and relations, such as protons, gravitation, energy, attitudes, motives, personality traits, or the cognitive map, as none of these is directly observable.

You can imagine that the empiricists' focus on observable facts made them suspicious about theoretical imagination. Strict empiricists do not want to have anything to do with unobservables. However, criticism of strict empiricism has become so loud since the 1950s and 1960s that almost no one thinks any more that science can be exciting, or can be done at all, without making conjectures about unobservables. A philosopher who calls himself an 'empiricist' (Van Fraassen, 2002) now just advocates a critical attitude towards speculation about a world 'behind' the phenomena, and urges us to 'save the phenomena', and to be wary of **metaphysics**. Van Fraassen has, however, abandoned the attempt to lay a firm foundation for science in objective empirical observation.

Theory-ladenness

The notion of theory-neutral data that may be 'read off' from the world has been severely attacked and has given way to the notion of the '**theory-ladenness**' of observations: observation is always partly determined by one's theoretical assumptions (see Chapter 2). Observations are not neutral and facts are not directly given events in the world – facts are statements about those events one holds to be true. A fact is a conviction or a belief that something is the case and is never independent of other notions one happens to believe.

Hence, the one-time demand that every scientific statement should be reducible to (an) observation statement(s) has been replaced by the notion that theories can be 'underdetermined by data' (see Chapter 3.5): several different theories may be compatible with the same dataset, or you may simply have not enough data. Cognitive science, for instance, would be impossible if we had to stick to direct observations.

This is not to say that empirical tests can be dismissed. On the contrary, to use our imagination in order to construct bold theories is one thing, to stay open-minded and revise or even refute your theory in the light of evidence to the contrary is another. This is a far cry from mere speculation, superstitious explanation and prejudice. Unscientific explanations like these tend to be final and dogmatic, invoking revelation or authority without giving reasons and seeking evidence. Science, on the contrary, should be open to tests and arguments and sensitive to evidence, including empirical evidence. For this reason, the logician Irving Copi wrote: 'The vocabulary of "hypothesis", "theory", and "law" is unfortunate, since it obscures the important fact that *all* of the general propositions of science are regarded as hypotheses, never as dogmas' (1961: 423, original emphasis).

As we said before, scientific knowledge is at all times revisable and never definitive. Though scientists may have good reasons and good evidence for thinking that their theories are true, they can never be certain in an absolute sense.

Theories

Informally speaking, a theory is a set of statements that organizes, predicts and explains observations; it tells you how phenomena relate to each other, and what you can expect under as yet unknown conditions. De Groot gives the following more formal definition:

> [A theory is] a system of logically interrelated, specifically non-contradictory, statements, ideas, and concepts relating to an area of reality, formulated in such a way that testable hypotheses can be derived from them. (1969: 40)

A theory, to some extent, fixes the vocabulary in which observations are phrased. A feature of natural science is that its vocabulary consists of a limited set of unambiguously defined terms; mathematical symbols are the most telling example, but also in physics the description of what is observed is limited to what can be expressed in terms of force, mass, velocity, etc.

From a theory predictions can be derived, and predicting is tantamount to explaining. When your theory predicts the position of the planets – that is, when a prediction can be derived in an unambiguous way from the theory – you can be said to have a model that explains (a relevant part of) the movement of the planets.

What exactly 'deriving' predictions means is tricky. As discussed elsewhere (see Chapter 3), the original idea (with the logical positivists) was that theories have a formal structure, like an abstract calculus, and deriving predictions is considered an exercise in formal logic. In physics, mathematical theories do indeed permit such quantitative predictions. In the history of psychology, however, Hull's attempt to build a formal deductive system for the prediction of behaviour was a failure (see above in Chapter 1.3). In most cases we have to rely on informal but still reasonably uncontroversial ways of deriving predictions from a theory. Usually, additional assumptions are therefore required, and some kind of translation of theoretical terms in empirical phenomena is also needed.

Laws and theories

A law can be defined as an empirical generalization. Ideally, it has the form: $(x)(Px > Qx)$ for all x at any time and place, if x is P, then it is Q – for example, frustration leads to aggression: all individuals, if frustrated (P), will exhibit aggression (Q). Of course, all kinds of exceptions and conditions will usually have to be specified.

Laws then are generalizations, but not all generalizations are laws. A nasty problem in the theory of science is how to distinguish between real laws and accidental generalizations. There is a genuine difference between the law that (all) copper (always) expands when heated, and the fact that all the coins in my pocket are silver: the former is a law of nature, the latter is an accidental generalization. The difference is usually expressed in terms of necessity or counterfactuals.

A law must necessarily hold, even in circumstances which do not now obtain – which is known as *counterfactual*. If we took a piece of copper to the moon and heated it, it would expand, but if I put a copper coin in my pocket, it would not turn into silver. The former generalization is counterfactual-supporting and thus a real law, the latter is not.

Put slightly differently, we may require that theories exceed the known evidence for them, that is, tell us more than we already knew. A genuine theory is also a commitment about what might happen, under conditions that are as yet unobserved.

The philosopher of science Karl Popper made refutability (the possibility that future situations will prove a theory wrong) of predictions the hallmark of real science, as we will see in the next chapter.

Furthermore, being part of a network of theories and concepts, and as strictly as possible (logically) connected to other laws, is a highly desirable property for laws in science. The system of classical mechanics is a case in point. Hooking up with good theories in other domains enhances the credibility of a theory. In psychology, for example, good working relations with functional neurophysiology are an asset for a model in cognitive or clinical psychology.

Empirical/experimental and theoretical laws

A distinction can be made between empirical and theoretical laws (Nagel, 1961; De Groot, 1969: 76–7). Laws in which *observables* occur are *empirical* generalizations, laws with *unobservables* can be defined as *theoretical* laws. In genetics, Mendelian laws which capture regularities in the inheritance of certain traits (e.g., hair colour, eye colour) are empirical: in contrast, whenever genes or chromosomes are mentioned in a law, this assumes a theoretical character.

As you can imagine, reverting to abstract and/or unobservable parameters goes hand in hand with larger and more theoretical networks. This suggests that there is virtue in constructing theoretical laws. First, it enhances the scope and anchoring of a theory – empirical laws are in danger of just enumerating familiar and trite facts. Second, theories ideally bring together qualitatively different phenomena within a single framework. Empirical laws capture commonalties at a phenomenal level, but theoretical laws suggest a deeper insight into underlying mechanisms and, consequently, the possibility of bringing together disparate observations under the same conceptual umbrella. Unification, the subsumption of many domains of empirical observation under a single conceptual framework, is an important goal for scientific inquiry. A classic example is Newtonian mechanics, which applies to falling apples as well as to the movement of the planets, and more recently to launching missiles, accelerating motorcycles and the like.

The difference between empirical and theoretical laws is not absolute, but gradual: observations in empirical laws are theory-laden (the theory to some extent determines what counts as a phenomenon), and unobservables in theoretical laws must also be verified by observation – only indirectly (see the paragraph on operationalization below, 1.4). Nevertheless, it is useful to be clear about the difference. An example in psychology where this tends to be ignored is Freud's psychoanalysis, where theoretical constructs are easily confused with clinical observations. Presumed 'observations' were only understandable and verifiable to trained observers who had already subscribed to the whole theoretical framework.

Box 1.4 shows a classical way of ordering observations, laws and theories.

Models

A **model** is a kind of mini-theory: it provides a more or less visualizable representation of the theory, as in some kind of analogy. A classic example is the model of the atom as a collection of coloured balls (electrons) circling around a core composed of differently coloured balls (protons and neutrons).

The term 'model' is also used for a more or less abstract picture of a part of reality in a field of inquiry where no fully fledged theory is (as yet) available. Psychology is rich in models: in any textbook of cognitive psychology we will find pictures of boxes and arrows that purport to model things like the working of memory (say, different kinds of storage from which information is retrieved) or attention (which may be modelled as a searchlight focusing on selected objects or as glue integrating features with objects). Sometimes, the model will take a mathematical form. Such models are, for example, used in economics to express relations between economic parameters (say, between average wage and unemployment) and can be utilized even if the underlying causes of such relationships are still unknown, that is, when a real theory is not available. In psychology, computer programs that simulate cognitive processes, like learning or problem-solving, can be regarded as models in the above sense. Whether such simulations qualify as a genuine *theory* of the domain they model is a moot point (De Groot, 1969: 335–42; see also Bailer-Jones, 2009).

Philosophers of science have started to put more emphasis on models and less on laws in their analysis of explanation on science. It is increasingly clear that laws cannot simply be applied to phenomena (explananda), as the classic story of explanation has it, but that some sort of redescription is needed. For example, in applying Newtonian mechanics ordinary things are described in terms of forces, mass, solids, points, etc. and certain idealizations and abstractions are made (e.g., a point as having no dimensions) (Cartwright, 1983). The description a model provides will always be partial, ignoring some and emphasizing other aspects. Models can be directly applied to concrete phenomena, unlike theories. In fact, applying a law to a phenomenon, as in explaining a high tide as an instance of gravity, one must use a model to connect them, redescribing water as mass attracted by the mass of the moon. Thus, models mediate between laws and phenomena.

A model therefore helps to apply laws to phenomena, but models can also provide an explanation or some kind of understanding even without theory. Daniela Bailer-Jones (2009) defines a model in science as an interpretive description of a phenomenon that facilitates access to the phenomenon. That is very broad, and includes not only all kinds of descriptions, analogies and metaphors, but also visualizations, drawings and scale models (like a model aeroplane, in a wind tunnel). The model of an atom as a kind of solar system with electrons circling the protons and neutrons in the centre, or of the heart as a double pump with four chambers, is another example. Models are not always visualizations: mathematical equations can also be a kind of model under this definition, for example dynamical systems modeling systems that change over time (see Chapter 8). Game theory is a model of human decision-making where mathematical functions model the utility (the value humans attach to outcomes).

Modelling in psychology

In psychology, laws and complete theories as in Newtonian mechanics are very scarce. Most explanations in psychology are low-level generalizations or models. In cognitive science, constructing models often means making computer simulations or flow diagrams. In Chapters 7 and 8 we will see examples of symbolic and connectionist models. The successes of such models in simulating real phenomena (e.g., in memory or visual cognition) are interpreted as evidence for their adequacy. For example, so-called graceful degradation where the performance of a network declines gradually, not abruptly when the network is damaged, is similar to the deterioration that occurs in the aging or damaged brain: this is interpreted as evidence in favour of connectionist models and against discrete symbolic models that collapse entirely when damaged.

The reasoning behind this is roughly that a successful working model of a phenomenon must somehow capture some aspects in a more or less correct way. For example, early cognitive psychology modelled attention as a limited capacity channel with filters and a kind of switch mechanism (input in such a model is pictured as little balls that can either enter the channel or be locked out at the entrance by the switch). To the extent that such a mechanism behaves like human attention, it seems that that attention is in some ways actually like a channel where something like information travels from the senses to our consciousness, or can be shut down on the way. The classical symbolic paradigm in cognitive psychology modelled thinking as symbol manipulation and at least some philosophers argued that they thus showed that somehow mind really is the processing of stored mental symbols. In Chapters 7 and 8 we will discuss the evidence for computational models of mind, both classical symbolic and connectionist.

However, not everyone agrees that such models really prove or explain anything about cognitive mechanisms – some experimental psychologists would argue that it is only by empirical investigation, in experiments manipulating the conditions and measuring the performance of a system, that we can find out about the mechanisms of the mind. In a sense, they say, models are cheap (not that they are easy), and having a model that seems to behave like the real thing (e.g., human memory, or vision, or attention) does not mean that the model is really equivalent to the cognitive system.

Modellers in turn argue against this and hold that the systems in psychology are too complex for experiments. Perhaps both sides are right.

BOX 1.4 Theories, laws and data: a hierarchy of language levels

1. *Theories*: a deductive system of related statements, partly unobservable, connected with correspondence rules to observations (e.g., kinetic gas theory).
2. *Experimental laws*: single statements about invariant relations between concepts, inductive generalizations (e.g., Boyle's law $PV = cT$).
3. *Assigning numeric values to concepts* (e.g., $P = 1.4$, $V = 3.2$, where P is pressure and V is the volume of a gas).
4. *Primary data* (observations, e.g., instrument readings).

From bottom to top, as we move from observation to theory, predictive power increases. The lower levels 'interpret' the higher levels, in the sense that they provide the connection between a theory and data, and they also provide visualizable or conceptual models (Nagel, 1961: ch. 5). Correspondence rules connect theoretical notions with measurement operations (e.g., P is the reading of a manometer).

(After Losee, 2001: 159–60)

Causes

A notorious problem in philosophy is the notion of **causality**. Philosophers have spent considerable effort in investigating the metaphysical foundations for the notion of cause (see e.g., Bochenski, 1973; Sosa and Tooley, 1993: Psillos, 2002): do causes really exist as a part of the furniture of the world? The answer to this question is still debated. For our purposes, we will concentrate on the role of causation in the philosophy of science and not on deep metaphysical issues. We are interested in what constitutes causal *explanation* in science.

David Hume thought that causality was no more than constant conjunction: when one event is always preceded by the same event (say, a billiard ball hits another, and the second one starts to move) we experience the first as the cause of the second. We also experience the connection as necessary: the two events should always go together.

Hume argued that causation and the feeling of necessity that goes with it are in our heads – as it were, psychological. We cannot know whether there is anything behind the sensory experience of 'constant conjunction'. (Psychologists have investigated the perception of causation: when two balls move across a screen, we think automatically that one is chasing the other.)

This is sometimes called the regularity view of causation (Psillos, 2002). Followers of Hume have refined his account. As a quite sophisticated example consider the so-called INUS condition: a cause is an insufficient but non-redundant part of an unnecessary but sufficient condition for an effect. When a cause is a necessary condition this means that the effect will not occur without the cause. A sufficient condition means the cause will not occur without the effect. (Being hit by a moving train may be a sufficient cause of death, but not a necessary one, since there are many more causes for our dying; being HIV positive is a necessary condition for developing AIDS, but not sufficient; and smoking may be the cause of lung cancer, but it is neither sufficient nor necessary.)

As an example of a causation explicated in terms of INUS conditions, consider how a short circuit can be the cause of a farm burning down. It is not a sufficient condition (without dry hay, the absence of rain, or if firefighters had arrived earlier, the fire might not have happened); the short circuit is a non-redundant condition, since all things being equal, the hay would not have caught fire without the short circuit; and it is not a necessary condition since the fire might have been caused by something else (e.g., lightning).

For our purposes, what is most crucial is to distinguish between causes and accidental correlations. Across the globe, the number of lampposts is correlated with the incidence of colon cancer. Nevertheless, removing lampposts will probably not occur to most people as a prevention of cancer (in rich countries there are more lampposts as well as a greater consumption of red meat, which seems to be related to colon cancer). In psychology and other social sciences, phenomena may have a common underlying cause, for example smoking causes yellow fingers and lung cancer. Yellow fingers are not the cause of lung cancer or vice versa, but both have a common underlying cause. Furthermore, multiple causality is a very common phenomenon. Violent crime rises with the temperature in inner cities, but of course violent crime has many more causes that have to cooperate in subtle ways to produce the crime effect (see Pigliucci, 2010, and Stanovich, 2010, for examples of causal reasoning and its pitfalls).

In psychology statistical techniques like multiple regression and path analysis are used to disentangle multiple causes. Behavioural genetics is a domain that has long moved beyond the simple question of whether behaviour is caused by nature or nurture.

The way genes cause a trait is hugely complicated and simplistic views of causation have led to naïve claims about genes for just about anything – baldness, buying expensive cars, etc. (the 'gene of the month').

In Chapters 8.3 and 9.4 we will encounter circular causation: what a roaming animal perceives will cause the way it moves, and wherever it moves will cause what enters its perceptual field. Complex causal tangles and circular causation abound in biological and psychological systems. Another complication in life sciences and social sciences is back-up and buffering systems: power generated by a power plant is the cause of our electrical appliances humming away, but when the plant breaks down the grid will draw its power from another plant without any change in the effect.

To sum up, in psychological and biological systems the simple view of causes, sometimes ridiculed as 'billiard ball causality', is simplistic and inadequate. Functional and mechanistic explanation to be discussed in Chapter 2 might be useful here, expanding the explanatory toolkit for the sciences of life and mind.

To introduce yet another complication: in the practice of explanation, what counts as a cause depends on the context and the *explanatory interests* of the investigator. When you ask what caused the death of an assassinated politician, you may say that religious extremism was the cause, but also that it was sloppy security, and for a pathologist it would be the biomechanics of bullets and human tissue – it depends on whether your explanatory interests are in politics, physiology or security tactics. Usually, phenomena are the products of a web of causes – what we would single out as 'the' cause depends on what sort of 'why' question we like to be answered, and what counts as the most relevant or conspicuous factor depends on a point of view. One man's cause is another man's background assumption. To give just one very simplified example, in one context we can say that genes cause depression, in another context that neurotransmitter deficiency causes depression, or that maternal deprivation causes depression. All of these are legitimate answers to the question of why an individual is depressive. Apparently, in scientific explanation there may be several domains and levels of causation.

Causal laws

Now, let us focus on the notion of causal laws. If we take a law to be a generalization connecting several events (as in the example of Box 1.4, increasing the temperature and keeping volume constant will increase pressure), then there is, intuitively, a difference between the mere contiguity of two events and a causal relation. One position is that to be really explanatory, laws must be causal. Recall the models and experimental generalizations: ideally, we want to know about the universal causes of things, such as the laws of gravity that explain falling apples, and soaring rockets and planetary motion. A crucial distinction that is mentioned above is between real laws and accidental generalization: an infection is lawfully caused by germs, but the fact that the entire village caught those germs at the church fair and the vicar has blue eyes is immaterial. The difference is that causes distinguish real laws from accidental generalizations. The intuition is that cause determines or necessitates the effect, and that the necessary connection underwrites the explanation. One might ask whether that solves the problem, since the question now becomes one of how to define causes. If physical necessity is part of that definition we must admit that this is observable, and real empiricists must be wary of such metaphysic constructs.

Nagel (1961: 74) lists four conditions for causal laws. First, there must be an invariable relation between cause and effect: the cause must be both a necessary and a sufficient condition for the effect. Second, cause and effect must be in the same spatial domain, or there must be an intermediate chain of causes connecting them across space. Third, the cause must precede the effect and be temporally close to it. And fourth, the relation must be asymmetrical: sunlight causes shadows, but not vice versa. According to these criteria, many laws of nature are not causal: it is a law that water is H_2O, but this is not a causal relation. Boyle's law does not qualify as causal, since P and V change at the same time. Furthermore, very few if any interesting laws in psychology are necessary: frustration can sometimes not lead to aggression. A partial solution is the notion of *ceteris paribus* laws: the effect follows only when the circumstances do not change. But even then, we must admit that many laws are only statistical: it is pretty certain, and very important to know, that smoking causes cancer, but the latter does not always follow from the former, only more frequently.

It seems then that the notion of causal laws satisfies our intuition that, unlike accidental generalizations, real explanations show how the effect follows with physical necessity from the cause, but that it is unclear how to delineate causes and necessity.

Interventionism

James Woodward (2003) proposed an interesting and influential account of causation, known as interventionism. The idea is that whatever is important about causal explanation can be understood in terms of experimental manipulation.

A cause is a point of intervention by, for example, experimental manipulations. One application of this idea is that evaluating counterfactuals is a matter of finding out what happens when interference is omitted; this amounts to verifying that observed regularities are real laws (causal nomological connections), not just accidental generalizations. So the manipulative view on causation sees an essential connection between human agency and causation. Woodward makes a strong connection with experiments: he proposes to define causes as what would happen if certain experiments (natural or human) were conducted. This idea of causation goes hand in glove with an analysis of experiments as intervention in some causal chain.

So, a causal claim really refers to the outcome of a hypothetical (not necessarily actual or practically possible) experiment, and causal claims that cannot be phrased in terms of (perhaps hypothetical) experiments will probably just lack meaning.

Bringing induction, deduction, laws and observations together: the empirical cycle

The notion of an empirical cycle (De Groot, 1969) nicely captures the interplay of data and theory, deduction and induction, in the practice of science. It consists of the following stages: observation, **induction**, **deduction**, testing and evaluation (De Groot, 1969: 27 ff.).

Observation is the stage where empirical material is collected and ordered. As a first approximation, it is systematic perception (the reader will recall that organized – or systematic – common sense was our 'quick and dirty' definition of science). Tentative or implicit hypothesis formation also occurs in this stage – if only because no perception is possible without (perhaps implicit) concepts and presuppositions, without some point of view. What is selected and observed reflects implicit hypotheses and theories and these are made explicit in the next stage.

Induction (including abduction) then is the phrasing of an explicit hypothesis. 'Explicit' means that the hypothesis yields specific, verifiable predictions that can be empirically tested.

Deduction refers to the derivation of predictions from hypotheses. The logical positivists demanded that all theories have a strictly logical or mathematical form, so that in their view deduction was an exercise in formal logic or mathematics. Such strictly formal theories, however, are very rare in psychology, if they exist at all, and as mentioned before, attempts by, for instance, Hull to force such an abstract calculus on psychology were unsuccessful – some would say, just silly. However, even in a less formalistic conception of hypothesis, the requirement that empirical consequences of a theory must be specified (and subsequently tested) remains. One of the ways to derive testable predictions from theoretical concepts is *operationalization*. This means that a concept is defined in terms of measurement operations. A good example is intelligence, defined as the score on an intelligence test. The choice of quantifiable behavioural indicators for psychological constructs is an important aspect of psychological experimentation.

The aim of the deduction stage is to formulate predictions, in such an explicit, precise and unambiguous way that they can be tested against empirical data.

The *testing* stage is about the confrontation of these predictions with empirical data. It must be emphasized that a hypothesis is a generalizing statement: it refers to a class of events, not to single facts (it is, possibly, a law that stress is conducive to premature ageing; it is not a law that John has grey hair). This implies that predictions must contain references to new situations, which are not already observed. For example, the law that frustration leads to aggression should be tested by comparing the prediction with the behaviour of a new population.

Finally, in the *evaluation* stage the results of the test are used as feedback for the more general theory from which the hypotheses are derived. Depending on the situation, one of two competing theories might have to be rejected in favour of the other, but more frequently, no such choice is available, and the theory will be expanded, qualified or amended; for instance, frustration leads to aggression only in certain circumstances, or in certain populations. There are no hard and fast rules for interpreting the results: the decision about what to change in one's theory will to some extent remain subjective, influenced by prejudices and opportunism. For example, the investigator may blame contradictory results on artefacts, or nuisance variables or whatever, or may invent ad hoc hypotheses to save his or her favourite theory. Alternatively, unexpected results may lead to new discoveries completely beyond the hypothesis tested (so-called serendipity). This stage then can at least partially be situated in the context of discovery.

In any case, the new theory will again spawn new hypotheses, to be tested on new data, leading to new tests and interpretations – and so the empirical cycle starts all over again. Bad ideas will fade away when no empirical evidence for them is found. The empirical cycle is thus a never-ending circular process, where subjective decisions will always in principle be formulated in an (at least partially) objectively testable form. Thus, the *context of discovery* and the *context of justification* are both in play. Induction is as indispensable as deduction.

BOX 1.5 The empirical cycle

1 Observation
2 Induction, abduction (hypothesize theory/law)
3 Deduction of prediction from theory
(may require operationalization: define concepts as measuring operation, e.g., IQ)
4 Testing hypothesis
observation, confirmation or disconfirmation
5 Evaluation → 2. hypothesize revised theory

Note that:

- Induction is to some extent guesswork; its results are not objectively certain.
- Explaining is equivalent with (successful) prediction; both consist of deducing from a theory.
- Observations are inductively collected into a theory, and then predictions/explanations are deduced from that theory.
- Testing and evaluation are theory-laden, and depend on interpretation and interest.

1.5 CONCLUSION

In this chapter we outlined the contours of scientific knowledge. Explanation and reduction, referring to underlying causes, are crucial for science. Causes are difficult to define, but intuitively causal explanation (or explanation by laws) marks an intuitive difference between deep necessary explanations and accidental generalizations.

Realism, idealism and pragmatism put forward different views on the origin and justification of knowledge. Pragmatism recognizes both the subjective component and the objective success of science.

Justification of knowledge should be distinguished from the factors influencing discovery, but the distinction here is fluid. Induction, deduction and abduction underpin knowledge claims. Induction and abduction generate (fallible) theories from which testable hypotheses are deduced, and after empirical testing new or amended theories are produced, and so on.

The difference between observations and theories is a matter of degree: observations are theory-laden, theories should be partly translatable into possible observations.

A convenient way to understand how data, laws and theories relate is as a hierarchy of descriptions, from more concrete and observable to more abstract and formalized statements.

FURTHER READING

A textbook on theory construction in psychology:

Kukla, A. (2001) *Methods of Theoretical Psychology*. Cambridge, MA: MIT Press.

A collection of key readings:

Boyd, R., Gasper, P. and Trout, J.D. (eds) (1991) *The Philosophy of Science*. Cambridge, MA: MIT Press.

Two books on scientific method and (un)scientific reasoning:

Pigliucci, M. (2010) *Nonsense on Stilts: How to Tell Science from Bunk.* London: University of Chicago Press.

Stanovich, K.E. (2010) *How To Think Straight About Psychology* (9th edn). Boston, MA: Allyn & Bacon.