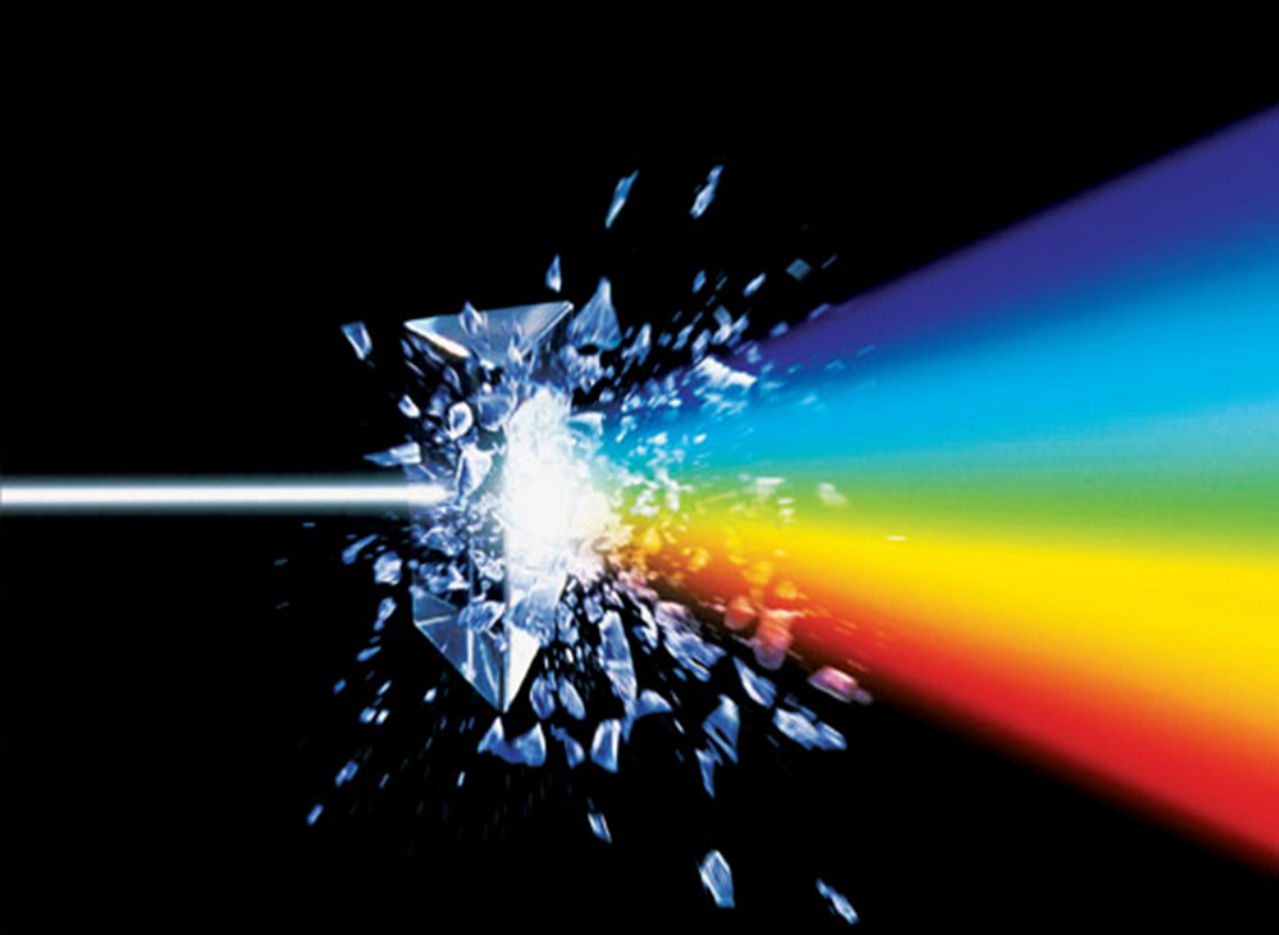


UNDERSTANDING PSYCHOLOGY AS A SCIENCE

AN INTRODUCTION TO SCIENTIFIC AND STATISTICAL INFERENCE

ZOLTAN DIENES

Understanding psychology as a science

Also by Zoltán Dienes

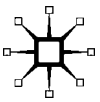
IMPLICIT LEARNING: Theoretical and Empirical Issues (*Berry, D. C.* & *Dienes, Z. (1993), Psychology Press*).

Understanding psychology as a science

An introduction to scientific and statistical inference

Zoltán Dienes
University of Sussex

palgrave
macmillan



First published 2008 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN-13: 978–0–230–54231–0 paperback
ISBN-10: 0–230–54231–X paperback
ISBN-13: 978–0–230–54230–3 hardback
ISBN-10: 0–230–54230–1 hardback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
17 16 15 14 13 12 11 10 09 08

Printed and bound in China

To Mr. Mallenson, who lent me Magee's 'Popper' when I was in Imberhorne's Sixth Form, and to Jonck, whose witty and profound comments on this manuscript I could sadly never hear.

This page intentionally left blank

Contents

Preface

The aim of the book is to address those issues in the philosophy of science, including the conceptual basis of statistical inference, that have a direct bearing on the practice of psychological research. The book will enable lecturers teaching critical thinking, research methods or the new British Psychological Society's core area 'conceptual and historical issues' to cover material that every psychology undergraduate should know, but does not. The book will also be valuable for masters students, PhD students and experienced researchers.

The book is organized around influential thinkers whose admonitions and urgings are heard in the head of every research psychologist. The core arguments surrounding Popper, Kuhn, Lakatos, Fisher, Neyman and Pearson and Bayes are still live, heated, important and potentially within the grasp of every psychology undergraduate. Further, key points of the Neyman–Pearson approach are deeply misunderstood even by seasoned researchers in ways that can reflect badly on their research decisions. The best place to uproot these misconceptions is right at the undergraduate level, but I found no other existing book suitable for this purpose. Further, there is a live debate on whether psychologists even should be following the orthodoxy of Neyman–Pearson. Few psychologists have an opinion on this matter (which has important consequences for how research is conducted and evaluated) because they have not been exposed to the issues (and mostly do not realize there are any issues). There is a large and growing literature on the confusions and misuse of orthodox statistics, and what the alternatives might be. But the clarifications have not percolated through to ground level. The reason is that the literature is largely technical and read only by the aficionado. The very people who need to know it are just the people who do not read it. What has been missing is a simple (though not simplistic) introduction showing conceptually how the characteristics of the different approaches arise from first principles. There has been a gap in our education that has existed far too long. To paraphrase Phil Johnson Laird's famous quip, I hope the current book helps fill this much unwanted gap. (I leave the task of my writing books filling much *needed gaps* to other occasions.) In any case, if my role in all this is to start corrupting undergraduates at a young age, the cat will be let out of the bag. I think the sooner the conceptual issues underlying inference form a part of the undergraduate's education in research methodology, the better.

The first two chapters cover classic philosophy of science (Popper, Kuhn, Lakatos) in a way accessible to psychologists while avoiding the normal textbook caricatures. The aim is to appreciate the depth of these authors so that their ideas provide real tools for thinking about research. Examples are drawn from psychology and practical advice is given.

The next three chapters motivate the reasoning behind statistical thinking. Chapter 3 covers the Neyman–Pearson approach, that is the logic meant to underlie the statistics in every textbook for psychology students. It is a sad fact – indeed, a scandal – that few undergraduates, PhD students or lecturers actually understand the logic of hypothesis testing (with or without Fisherian twists). The Neyman–Pearson logic will not be fully grasped until alternatives are also presented. Many people instinctively believe that they are getting Bayesian or likelihood answers to Neyman–Pearson statistics. This basic confusion more than any other probably underlies the widespread misuse of significance testing. So Chapter 4 shows the logic of Bayesian inference, the opposite of Neyman–Pearson in fundamental

ways. The book's website (http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/) includes a program that can be used for analysing data in a Bayesian way. Finally, the logic of the third major school of statistical inference, likelihood inference, is presented and motivated in Chapter 5. Chapter 5 shows how to use likelihood inference in practice and how it leads to different research decisions than the other schools. The book's website also includes programs for conducting likelihood analyses.

None of the chapters on statistics assume mathematics beyond that necessary for an undergraduate course in statistics as run in an average psychology department. The chapters aim only to provide foundational concepts and link them to practical research decisions. The arguments for each school are presented conceptually, so for the first time the average user of statistics can start making informed decisions – and accept or reject orthodoxy on a rational basis.

I wish to warn the psychology undergraduate reader that the material may at first appear daunting, as it will require thinking in new ways. Persistence will bring reward. Do not deride your own intellectual reactions; the interesting thing about philosophy is that whatever view you hold on a topic, there is bound to be some very eminent philosopher who holds a similar view. In many cases, I am not giving you settled answers. Be confident in thinking through your own arguments. However, the multiplicity of views does not mean that nothing has been achieved in philosophy nor that all views are acceptable. On the contrary, having understood the issues discussed in this book, you will evaluate and practise research in ways you could not have done before. Whatever decisions you come to on the open issues, you will become a far better researcher and evaluator of research having thought about the issues in this book.

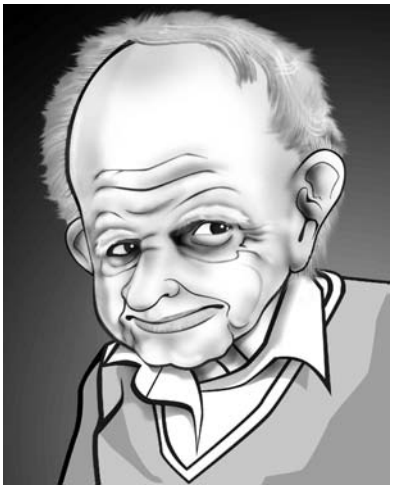
If in reading this book you feel confused at times, that is a very good sign. Confusion means you have found a way to arrive at a deeper understanding. Value that feeling. Confusion is not the end goal of course, it is a sign-post for what to think about and a spur to think it through now because your mind is ripe. While I hope to confuse you, I also hope I have given you what you need to subsequently gain clarity. Any confusion caused by errors or omissions on my part is not good, and I will endeavour to correct them in any future editions. You can contact me on dienes+inference@sussex.ac.uk.

Thanks to the Cognitive Science Society for allowing reproduction of Figures 1.4 and 1.5, for which they hold the copyright. Many thanks to Robert Clowes, Bruce Dienes, Alan Garnham, Nomi Olsthoorn, Ryan Scott and Dan Wright, and also the four anonymous reviewers, for their valuable comments on previous versions of the manuscript. Many thanks also to the students I taught this material to over the last two years without whom truly this book would not have been possible. I am indebted to Leiming He and Xiaolong Zhang for the excellent cartoons. Finally, I very much appreciate the patience of my wife Mina and my son Harry over the past year.

Zoltán Dienes
Brighton, UK, 2007

This page intentionally left blank

1 Karl Popper and demarcation



Philosophy, they say, cannot by its very nature have any significant consequences, and so it can influence neither science nor politics. But I think that ideas are dangerous and powerful things, and that even philosophers have sometimes produced ideas.

Popper (1963, p. 6)

Box 1.1 Questions

What is science?

What is the difference between science and pseudo-science?

What is the difference between good science and bad science?

On what grounds should papers submitted to scientific journals be rejected or accepted?

Are Christian Science, Creation Science, Scientology, astrology, traditional Chinese medicine, or chiropractic sciences? Why or why not and why does it matter?

Is psychology a science? Good science or bad science?

How does knowledge grow?

We are constantly faced by choices concerning what to believe, choices of great practical and personal importance (Box 1.1). Will taking an anti-depressant pill help me if I am depressed? Should I prefer the use of these herbs or those drugs for my hay fever? Are the economic policies for controlling inflation of one political party more effective than those of another? Are some forms of exercise rather than others more effective for losing weight, gaining muscle or excelling in sporting performance? What methods might help me overcome anxiety? What are the best methods for learning a second language? What principles might enable different groups of people to live together more harmoniously, or a given group to work more smoothly, or a relationship to run better?

The choice of what to believe is not just a practical matter. A desire to understand the world seems intrinsic to homo sapiens. I remember as a teenager being fascinated to find

I needed only my high-school algebra to follow arguments concerning the fundamental nature of space and time as portrayed by special relativity. The sheer fact that such elegant and counter-intuitive laws had survived severe tests counted for me as amongst mankind's greatest achievements. (It still does.) Further, the details of how evolution works, of how genetics work, of how the brain works, of how the mind works are surely worth knowing for no other reason than we gain understanding of ourselves and the universe. The burgeoning popular science section of bookshops attests to people's simple desire to understand, that is, to know the best explanations around. But what counts as a satisfactory explanation of phenomena in the world? When is one such explanation better than another? This is a philosophical problem, but not an empty one, it is a problem of real substance regardless of whether our aim is practical or it is understanding itself.

The Austrian Karl Popper (1902–1994) formulated the problem in terms of what distinguishes science from non-science (including pseudo-science). He called it the problem of demarcation. He was not interested in merely categorizing or labelling something as science or not, he was interested in the substantial problem of how we can best pursue knowledge of the world. His works inspired many scientists, including various Nobel Laureates, who publicly declared how beneficial Popper's philosophy had been to them as scientists. For most of his academic career Popper held a chair in logic and scientific method at the London School of Economics. He was also a highly influential political and social philosopher. He continued producing work right up until his death at age 92. For his achievements, Sir Karl was knighted in 1965.

Background: logical positivism

Popper was born in Vienna in 1902, then an intellectual and cultural centre of the Western world. A prominent philosophical movement during Popper's time in Vienna was the so-called 'Vienna Circle' (Wiener Kreis), a group of scientists and philosophers who met regularly to discuss the implications of the major revolutions in science happening at that time, especially those triggered by Einstein's work (see Edmonds and Eidinow, 2001, for a very readable biographical account). The discussions were led by Moritz Schlick (1882–1936); attendance required his personal invitation.[1] Among the members were mathematician Kurt Gödel (of Gödel's theorem fame) and philosophers like Rudolf Carnap and Carl Hempel, and the economist Otto Neurath. They formulated a philosophical approach called *logical positivism*. It is worthwhile knowing a little bit about logical positivism to understand what Sir Karl was partially opposing.

Logical positivism was itself a reaction against a style of philosophy then popular in the German world which emphasized pompous, difficult and obscure writing. In fact, members of the Vienna Circle wondered whether such writing actually said anything at all. In order to sort meaningful statements from meaningless nonsense, the logical positivists proposed that meaningful statements were either definitions, and thus necessarily true (like 'a triangle has three sides'); or else *verifiable* empirical statements (statements about the world). Empirical statements were only meaningful if they satisfied the *verification criterion*, that is, if one could specify the steps that would verify whether the statement was true.

1. Schlick had an early death because of a vengeful and mentally unstable student. Schlick had failed the student and later slept with his wife. The student confronted Schlick on the steps of the University and shot him in the chest. Although Schlick was not a Jew, the killer was subsequently hailed as a Jew killer in the popular press.

For example, the statement 'my desk is three foot tall' can be verified by measuring its height with a ruler, so the statement is meaningful. But how could one verify that 'The essence of Spirit is freedom'? Or 'The world does not really exist, it is just an idea'? Or 'The world really does exist'? Or 'Free will is an illusion'? Or 'God exists'? Or 'God is loving kindness'? For the logical positivists such unverifiable statements were just *metaphysical* nonsense. Much conventional philosophy could be swept away as dealing with pseudo-problems! One could get on with the real work of developing mathematics and logic (which deal with statements which are necessarily true) and natural science. Science could be construed as sets of empirical statements, dealing only with possible observations that could be directly verified, and theoretical statements which acted as definitions linking theoretical terms (like 'electron') to observations. If you have ever read a lot of old German texts (or indeed many things written today in English – e.g. see Sokal and Bricmont, 1998), you may have some sympathy with the urge to throw much in the dustbin labelled 'gibberish' (Box 1.2).

Box 1.2 Operational definitions

A notion inspired by logical positivism, and used extensively in modern psychology, is the idea of an *operational definition*, introduced by the Nobel Laureate physicist Percy Bridgeman in 1927. An operational definition defines the meaning of a concept in terms of the precise procedures used to determine its presence and quantity. For example, an operational definition of intelligence could be the score obtained on a certain IQ test. An operational definition of unconscious knowledge could be above baseline performance on a knowledge test when the person verbally claims they are guessing. An operational definition of the degree of penis envy in women could be the number of pencils returned after an exam. Experimental psychologists habitually produce operational definitions. They are invaluable but often the 'definition' is actually not a definition at all but rather a way to measure more-or-less imperfectly the thing we want to measure. More-or-less imperfectly measuring is different from defining; if the measurement is a definition the outcome is never imperfect. Defining being 'successfully anaesthetized' as 'having received a standard dose of anaesthetic' might strike some unfortunate people as having completely missed the point. When a psychologist has his back against the wall in an argument to the effect he is not measuring what he claims, there is the temptation for him to say, for example, 'But what I MEAN by emotional intelligence is a score on this test'. (The word 'mean' is always emphasized in this tactic.) The temptation should normally be resisted in psychology. It is rarely the case that what we *mean* by a concept is exhausted by some particular way of measuring it. An operational definition should not be an excuse to stop thinking critically about the validity of one's measure.

The logical positivists had two problems in determining whether a sentence was verifiable. There is the problem of how to verify statements about specific individuals and their observable properties, like 'Emma the swan is white'; and there is the problem of verifying generalizations like 'All swans are white'. The first problem was meant to be solved by direct observation, and the second by a putative logical process called *induction*. Induction can be contrasted with *deduction*. Deduction is the process of drawing inferences such that if the premises are true the conclusion is guaranteed to be true. For example:

All swans are white
Sam is a swan

Conclusion: Sam is white

Induction is the process of inferring universal rules given only particular observations:

Sam the swan is white;
Georgina the swan is white;
Fred the swan is white;
. . .
Emma the swan is white

_____
Conclusion: All swans are white (?)

The conclusion here cannot be guaranteed to be true, hence the inference is not deductive. But still one instinctively feels the conclusion has increased in plausibility by the repeated particular observations. We feel very sure the sun will rise tomorrow for no other reason than it has done so repeatedly in the past. You might say, we also have a theory of gravitation that predicts the sun will continue to rise. But why do we feel confident in that theory other than for the fact that it has worked on so many particular occasions in the past? Inductivists, including logical positivists, believe that science proceeds by induction. Science is objective because it is based on actual observations rather than just speculation, and it goes from those particular observations logically – inductively – to general rules.

With the Second World War and Hitler's dismantling of the university system in Germany and neighbouring countries, many Jewish intellectuals, including members of the Vienna Circle, sought refuge in the United States. In fact, logical positivism went from being a minority philosophical position in Europe before the war to being the dominant force in American philosophy by about 1960 (see Giere, 1999, Chapter 11). Since then many people have defined their philosophy of science either as being in the same tradition or as reacting against logical positivism. Popper, however, is the person who credits himself with killing it off (Box 1.3).

Box 1.3

Consider the following theory: In human beings, the hippocampus is required for spatial navigation.

You find Sam who, due to a recent unusual viral illness, has destruction of all his hippocampus and no other brain structure. His spatial navigation is very bad.

Have you established the theory is true?

What about if you found eight teenage drivers who all destroyed their hippocampi (and no other brain structure) in car accidents. Their spatial navigation is very bad.

Have you established the theory?

Your ninth teenage driver with complete destruction of the hippocampus has excellent spatial navigation.

What can you conclude about the theory?

The problem of induction

Popper denied all aspects of logical positivism just described. The positivists rejected metaphysics as meaningless; Popper argued that metaphysics could be not only meaningful but also important. The positivists wished to view science as a method of moving towards certain knowledge: knowledge based on a firm foundation of observation and induction. Popper

denied that either observation or induction provided means for moving towards certain knowledge. He denied further that certainty was even the goal of science. The core feature of Popper's philosophy was fallibilism: we may be wrong in anything we believe. We will see how this principle can be turned into a practical philosophy of science that allows knowledge to progress!

Induction had received a crippling, if not fatal, attack earlier by the Scottish philosopher David Hume (1711–1776). Hume argued that we are never justified in reasoning from repeated instances of which we have experiences (e.g. different swans being white) to other instances of which we have had no experience yet (other swans we are yet to observe the colour of). Induction is never justified. No matter how often we have seen a white swan, it never follows that the next swan we see will be white, yet alone that all swans are white. (A famous example because when the British went to Australia they finally found swans that were black!) A common response is to accept that, yes, from particular observations no generalization follows with certainty – but surely the *probability* of the generalization is increased with each supporting observation. Each time we see a white swan, is it not more likely that all swans are white? Hume pointed out that this does not follow. No matter what observations up to now support a generalization, it may be the case that every instance after *now* fails to support it. No matter how often your car has started early in the morning, one day it will not, and one day will mark the point that the car never starts again. Indeed, you might feel that the more days your car has started, the *less* likely it is it will start the next day, because that is the effect of age on cars! One morning the chicken at the Colonel's ranch is greeted by someone who is not offering breakfast. But surely, you might reply, induction has empirically proved itself because using induction in the past to infer generalizations has often been successful. This argument, Hume pointed out, assumes that induction is true in order to argue for it: No matter how often induction has worked in the past, there is no reason to think it will work ever again. Not unless you already assume induction, that is.

Popper (e.g. 1934, 1963, 1972) accepted all these arguments. The claim that an explanatory universal theory is true is never justified by assuming the truth of any number of observation statements. One stunning intellectual event happening during Popper's formative years was the replacement of Newtonian physics by relativity and quantum physics. No theory had received more 'inductive support' – over several hundred years – than Newtonian physics. But in the space of a few years it went from a theory widely regarded as being the established truth to one recognized as literally false. The moral for Popper was clear: We can never actually establish the truth of our theories. Establishing truth is not what science does. No matter how strongly we or others believe a theory, the theory is and always remains just a guess, even if it is our best guess. Induction does not exist.

What is the problem that people wished to explain by postulating induction? Typically, Popper points out, the situation is that we have several theories which compete as explanations and we must choose between them. Bertrand Russell (1872–1970), the grandfather of the logical positivists, thought that without working out how induction could be valid we could not decide between a good scientific theory and the 'mere obsession of a madman'. Popper argued that we do not need induction to solve the problem of theory choice. He pointed out that while assuming the truth of a particular observation or test statement never allows us to justify the claim that a universal theory is *true*, it sometimes allows us to justify the claim that the universal theory is *false*. For example, accepting that 'Sam the Swan is black' allows us to conclude that the universal claim 'All swans are white' is false. Although we cannot *establish* our theories, we may be able to successfully criticize them. There is a genuine asymmetry here in our ability to falsify versus establish a theory, given we accept

singular observations. Popper exploits that asymmetry to the full in developing his philosophy of science. Rationality consists in critical discussion, trying to find the weaknesses in a theory or argument, weaknesses that may be shown by, for example, observations – observations used to criticize the theory.

The role of critical discussion

How does knowledge grow? According to Popper, there is only one practicable method: critical discussion. Knowledge does not start from unprejudiced pure observation. There is no such thing: All observation involves some theory, some prejudice. Consider the ridiculous nature of the task of just telling someone: 'Observe!'. Observe what? For what purpose? Science is not built from naked observations. One starts with a theory, a conjecture. Then one tries to refute it. If the theory resists refutation, new lines of criticism can be attempted. In the light of a successful refutation, a new conjecture can be proposed to deal with the new problem situation.

Popper points out that in most societies in human history and around the world we find schools of thought which have the function not of critical discussion but to impart a definite doctrine and preserve it, pure and unchanged. Attempting to change a doctrine is heresy and will probably lead to expulsion from the school. In such a tradition, the successful innovator insists that he is just presenting what the master's original doctrine really was before it was lost or perverted! The critical tradition, in contrast, was founded by explicitly establishing the method of proposing a conjecture then asking your students, after having understood it, to try to do better by finding the weaknesses in the proposal. In the critical tradition, the aim is not to preserve a doctrine but to improve it. Popper (1994) suggests that remarkably this method was invented only once in human history. Whether or not Popper's historical conjecture is true is irrelevant for his main point, though it remains an intriguing speculation. Popper suggests that it was Thales (c.636–c.546 BCE) who established the new tradition of free thought in ancient Greece. The students of Thales openly criticized him and presented bold new ideas to overcome those criticisms (for details, see Popper, 1963, Chapter 5). The growth in knowledge in the space of a few years was breathtaking. Xenophanes (570–480 BCE) spread the critical tradition, expressing the philosophy that all our knowledge is guess work in this poem (Popper's translation, in his e.g. 1963):

> But as for certain truth, no man has known it
> Nor will he know it; neither of the gods,
> Nor yet of all the things of which I speak.
> And even if by chance he were to utter
> The perfect truth, he would himself not know it;
> For all is but a woven web of guesses

Though we can never know if we have the truth, we can always try to improve on what we have. This critical tradition died in the West, according to Popper, when an intolerant Christianity suppressed it. It smoldered in the Arab East, from where it finally migrated and ignited the Renaissance and modern science. Once again there was an explosion in knowledge. The scientific tradition just is this critical tradition.[2]

2. In developing your own critical skills, bear in mind the advice of Donald Broadbent: Stand on the shoulders of those who have gone before you and not on their faces.

Popper's historical conjecture about the critical tradition arising only once is a reminder not to take the tradition for granted. It would be easy to see claims of the importance of criticism as platitudes. But participating in a tradition of one's students and colleagues criticizing oneself is not psychologically easy. Further, regardless of the intrinsic difficulties, external attacks and erosions are all around, from religious, party political and corporate authoritarianism. To take an extreme example, in 1957, Mao said, 'Let a hundred flowers bloom; let the hundred schools of thought contend', meaning let everyone voice their opinions so the best ideas may survive. Those who voiced opinions critical of Mao were silenced.

What is science?

Popper rejected the logical positivists' attempt to distinguish meaningful statements from nonsense but instead sought to distinguish science from non-science, or metaphysics. But that does not mean that Popper believed in a scientific method as a specifiable formula to be followed for developing knowledge. 'As a rule, I begin my lectures on Scientific Method by telling my students that scientific method does not exist. I add that I ought to know, having been, for a time at least, the one and only professor of this non-existent subject within the British Commonwealth' (1983, p. 5). Popper despised compartmentalizing knowledge into subjects in any case. What is really important is just interesting problems and attempts to solve them. For example, there is the problem (which I find interesting) of explaining why hypnotized people behave as they do. Solving that problem may involve thinking about it in ways typical of the subjects of philosophy, cognitive and social psychology, or neuroscience. But none of *those* 'subjects' has any separate reality (beyond being useful administrative divisions for organizing a university); I may blend their ideas in my own conjectures concerning hypnosis. All relevant knowledge should be brought to bear on interesting problems. Nonetheless, 'Scientific Method holds a somewhat peculiar position in being even less existent than some other non-existent subjects' (p. 6).

There is according to Popper no method of discovering a scientific theory, no method of inferring a theory from 'pure' observation. Science consists of freely, creatively *inventing* theories: Science is made by people. There is also according to Popper no method – like induction – of determining the truth of a theory, no method of verification. There is not even any method of determining whether a theory is probably true. This follows from Hume's critique of induction. Many philosophers refuse to accept this conclusion (see e.g. Salmon, 2005, for a defense of the Popperian thesis, see Miller, 1994, Chapter 2. For an accessible introduction to arguments concerning induction, see Chalmers, 1999, Chapter 4). We will see later in the book that there is an approach, the Bayesian approach, aimed precisely at determining the probability of a hypothesis. But the Bayesian approach never actually answered Hume's critique as such; instead, the argument is *assuming that the world follows a certain specified type of model* (i.e., *assuming* generalizations of a certain type hold and will continue to hold), *then the probabilities of different versions of the model can be calculated*. Popper rejected such approaches, not only because he accepted Hume's critique but also because he rejected the relevance of the subjective probabilities the Bayesians used. This issue is very much a current live debate and is explained further in Chapter 4.

Popper thought how people invent their theories is not relevant to the logic of science. The distinction between the process of conceiving a new idea (which Popper, 1934, called

the psychology of knowledge) and the process of examining it logically (the logic of knowledge) is more commonly known as the distinction between the *context of discovery* and the *context of justification*. These latter terms were introduced by Hans Reichenbach (1891–1953) in 1938. Reichenbach had founded the 'Berlin Circle' and a type of logical positivism. He had been dismissed from the University of Berlin in 1933 due to Nazi racial laws and eventually established himself in the United States. Reichenbach illustrated the distinction between the context of discovery and the context of justification by the distinction between psychological and historical facts concerning Einstein the man, on the one hand, and the logical relation of his theory of general relativity to relevant evidence, on the other (see Giere, 1999). Giere suggests that Reichenbach was motivated by the cultural climate at the time to deny that the characteristics of a person, such as being a Jew, has anything to do with the scientific validity of a hypothesis proposed by the person. The hypothesis stands or falls on its own logical merits independently of who thought it up or how they thought it up (see Box 1.4).

Box 1.4 Is there a distinction between the contexts of discovery and justification?

The distinction between the context of discovery and the context of justification has good face validity. For example, Kekule is said to have thought of the idea of a ring of carbon atoms as the structure of benzene by dreaming of snakes biting their tails. The process of dreaming is part of the context of discovery or the psychology of knowledge and irrelevant to the evaluation of his hypothesis. It is the hypothesis itself – of a ring of atoms – and its relation to laboratory evidence that belongs to the logic of knowledge or the context of justification. Nonetheless, the usefulness of the distinction between the contexts of discovery and justification has been controversial. Thomas Kuhn (1969), for example, rejected the distinction. He did not believe that there was a special logic of scientific knowledge to make a distinctive context of justification. Kuhn also thought the processes by which a particular scientist within a discipline comes to discover new knowledge is an integral part of what makes the practice scientific. We will discuss Kuhn's position in more detail in the next chapter. Even within Popper's philosophy the distinction can sometimes be blurred. As we will see, central to Popper's account of science is that scientists adopt a 'falsificationist attitude'. Such an attitude is surely as much part of the psychology of knowledge as its logic. The psychological problem of what factors facilitate scientific discovery is a theoretically interesting and educationally important problem on which progress has been made; for example, in the work of Peter Cheng on the role of diagrams (e.g. Cheng & Simon, 1995) and of Roger Shepard on thought experiments (Shepard, 2001). Diagrams and images aid *psychologically* by embodying the underlying *logic* of the scientific problem. However, while the contexts of discovery and justification can be intertwined, Salmon (2005, Chapter five) believes that the distinction is still 'viable, significant, and fundamental to the philosophy of science' (p. 85) if not to scientists themselves. As scientists, we are concerned with the logical relations between theory and evidence, stripped of accidental irrelevancies concerning the discovery of the theory or of the evidence. Just what is relevant and what is irrelevant to these logical relations is an issue every chapter of this book will bear on.

There are no set methods for creating theories and there are no methods at all for showing a theory to be true. According to Popper, theory testing is not inductive, but deductive: Accepting certain observation statements can show a theory is false (the one black swan showing that not all swans are white). This is how observations make contact with theories; this is how, therefore, our knowledge acquires its empirical character. Science can only work in this way if a theory is falsifiable to begin with: The theory says certain things cannot happen. Non-science or metaphysics is, by contrast, non-falsifiable. That is a logical property distinguishing science from metaphysics. This distinction does not render metaphysics