

The cognitive basis of science

edited by

Peter Carruthers

Professor and Chair, Department of Philosophy, University of Maryland College Park

Stephen Stich

Board of Governors Professor, Department of Philosophy and Center for Cognitive Science, Rutgers University

and

Michael Siegal

Professor of Psychology, University of Sheffield

Published in association with the Hang Seng Centre for Cognitive Studies, University of Sheffield, and the Evolution and Higher Cognition Research Group, Rutgers University



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa
<http://www.cambridge.org>

© Peter Carruthers, Stephen Stich and Michael Siegal 2002

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2002

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10/12 pt. System L^AT_EX 2_ε [TB]

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

Library of Congress Cataloguing-in-Publication data

The cognitive basis of science/edited by Peter Carruthers, Stephen Stich and
Michael Siegal.

p. cm.

Includes bibliographical references and index.

ISBN 0-521-81229-1 – ISBN 0-521-01177-9 (pb.)

1. Science – Philosophy. 2. Science – Psychological aspects. I. Carruthers,
Peter, 1952– II. Stich, Stephen P. III. Siegal, Michael.

Q175 .C6127 2002

501 – dc21 2001043131

ISBN 0 521 81229 1 hardback

ISBN 0 521 01177 9 paperback

Contents

<i>List of contributors</i>	<i>page</i> ix
<i>Preface</i>	xi

1 Introduction: what makes science possible?	1
PETER CARRUTHERS, STEPHEN STICH AND MICHAEL SIEGAL	

Part one Science and innateness

2 Human evolution and the cognitive basis of science	23
STEVEN MITHEN	
3 Modular and cultural factors in biological understanding: an experimental approach to the cognitive basis of science	41
SCOTT ATRAN	
4 The roots of scientific reasoning: infancy, modularity and the art of tracking	73
PETER CARRUTHERS	

Part two Science and cognition

5 Science without grammar: scientific reasoning in severe agrammatic aphasia	99
ROSEMARY VARLEY	
6 Causal maps and Bayes nets: a cognitive and computational account of theory-formation	117
ALISON GOPNIK AND CLARK GLYMOUR	
7 The cognitive basis of model-based reasoning in science	133
NANCY J. NERSESSIAN	

8	Understanding the role of cognition in science: the <i>Science as Category</i> framework KEVIN N. DUNBAR	154
9	Theorizing is important, and collateral information constrains how well it is done BARBARA KOSLOWSKI AND STEPHANIE THOMPSON	171
10	The influence of prior belief on scientific thinking JONATHAN ST B. T. EVANS	193
11	Thinking about causality: pragmatic, social and scientific rationality DENIS HILTON	211

Part three Science and motivation

12	The passionate scientist: emotion in scientific cognition PAUL THAGARD	235
13	Emotions and epistemic evaluations CHRISTOPHER HOOKWAY	251
14	Social psychology and the theory of science PHILIP KITCHER	263

Part four Science and the social

15	Scientific cognition as distributed cognition RONALD GIERE	285
16	The science of childhood MICHAEL SIEGAL	300
17	What do children learn from testimony? PAUL L. HARRIS	316
18	The baby in the lab-coat: why child development is not an adequate model for understanding the development of science LUC FAUCHER, RON MALLON, DANIEL NAZER, SHAUN NICHOLS, AARON RUBY, STEPHEN STICH AND JONATHAN WEINBERG	335

	<i>References</i>	363
	<i>Author index</i>	396
	<i>Subject index</i>	405

1 Introduction: what makes science possible?

Peter Carruthers, Stephen Stich and Michael Siegal

In this brief opening chapter we briskly review some of the recent debates within philosophy and psychology which set the stage for the present collection of essays. We then introduce the essays themselves, stressing the inter-linking themes and cross-connections between them.

1 Introduction

The central position of science in our contemporary world needs no emphasis. Without science (broadly construed, to include all forms of technical innovation) we would still be roaming the savannahs of Africa like our *Homo habilis* ancestors, digging up tubers and scavenging scraps of meat. And without science (construed narrowly, as involving the application of an experimental method) we would have seen none of the advances in knowledge, technology and accumulation of wealth which have transformed the world and most of its people in just the last four centuries or so. Science now touches every aspect of our lives, from cradle (indeed, sometimes from conception) to grave. Given the manifest importance of science, the search for a scientific understanding of scientific thought and activity itself should need no further motivating. But in fact, the attempt to explain scientific cognition is not only extremely hard, but raises a whole host of fascinating and puzzling questions about the nature, development and operations of the human mind, and its interactions with culture.

This book is about the question: what makes science possible? Specifically, what features of the human mind, of human cognitive development and of human social arrangements permit and facilitate the conduct of science? These questions are inherently inter-disciplinary, requiring co-operation between philosophers, psychologists and others in the social and cognitive sciences. And they are, it should be stressed, questions which are as much about the psychological underpinnings of science as they are about science itself. That is, they concern what it is about our minds and/or mind-guided social interactions which make science possible, and how these factors relate to other things which we can do, either as adults or children. Indeed, one of the important themes of the book is the broad-scale architecture of the mind. For in order to understand

how science is possible we have to understand how our capacity for scientific theorizing fits into the structure of the mind, and what consequences that might have for the practice of science.

Steven Pinker in his well-known book *How the Mind Works* (1997) raises questions to which this volume is a partial answer. Having argued that cognitive science is approaching an understanding of many different aspects of the mind – vision, language, memory, and so on – he then lists various factors about human beings which (he says) we *cannot* yet begin to explain. One of these is consciousness. Another is science. According to Pinker, we don't really have any grip on how human beings can be capable of scientific thinking and reasoning. That is a striking and challenging claim. Taking up and answering that challenge, as this book begins to do, is something which should be of interest, not only to psychologists and philosophers of psychology, but to all those interested in understanding either the nature of the human mind, or the nature of science, or both.

2 Philosophy of science: a very short recent history

As the title of this section suggests, we here take the reader through a *very* brisk tour of recent developments in the philosophy of science.

2.1 *Positivism and beyond*

In the beginning of our story there was logical positivism, which dominated much of the middle part of the twentieth century (Ayer, 1946; Carnap, 1950, 1967; Hempel, 1965). The logical positivists were heirs to the classical empiricist tradition in philosophy of science and theory of knowledge – believing that all enquiry should be grounded in observation, and that the scientific task is essentially that of accommodating existing observations while correctly predicting new ones. They also believed that one central paradigm of scientific enquiry is enumerative induction. A pattern is discerned in our observations (each raven so far observed has been black), and then generalized into a universal law (all ravens are black). Much intellectual effort was expended in attempts to justify our inductive practices, and in discussion of the problem of underdetermination of theory by data (there are always, in principle, infinitely many distinct theories – that is, generalizations of the data – consistent with any finite data-set).

One stalwart and long-standing critic of logical positivism was Popper (1935, 1963, 1972). Popper pointed out that induction can only *generalize from* observation, whereas science characteristically goes *beyond* experience, by postulating theoretical entities such as electrons and X-rays, for example, which might explain it. The method of science, Popper argued, is not observation and

induction, but rather non-algorithmic theory-construction followed by testing. Scientists invent theories to *explain* their data, using imagination, analogy, intuition and any other resources which may happen to come to hand. (And since explanation isn't easy, scientists will generally be satisfied if they can construct just *one* explanatory theory.) But then, having devised a theory, they subject it to rigorous testing – deriving from it predictions concerning the observations which might be expected if the theory were true, and attempting to falsify the theory by making those observations which seem least likely. A theory is justified when it escapes falsification, on this view.

Up to this point, philosophy of science had been conducted in a relatively *a priori* fashion – with some reference to real scientific examples, admittedly, but mostly with philosophers of science just thinking about what scientists *ought* to do, rather than about what they actually *do* do. This all began to change in the 1960s and 1970s, when philosophy of science took its so-called 'historical turn', through the work of Kuhn (1962), Feyerabend (1970, 1975) and Lakatos (1970).

2.2 *Historical and naturalistic turns*

As Kuhn and others noticed, when one studies the history of science one discovers that the behaviour of actual scientists often fails to conform to the norms of scientific method laid down by philosophers of science. In particular, when scientists know of data inconsistent with their theories, they do not immediately abandon those theories and start again, as Popper would have had them do. Sometimes they try to explain away the recalcitrant data while continuing to hold onto their theory; but as often as not they simply ignore it, and get on with the business of developing their favoured theoretical approach. This gives rise to a dilemma for *a priori* philosophers of science. Either they can claim that the actual practice of scientists has been irrational, or at least inappropriate – in which case the immense success of science is rendered utterly mysterious. (How *can* science be so successful if scientists have mostly been doing it all wrong?) Or they can take the historical results as a refutation of their proposed methodologies. Almost all philosophers of science converged on the latter course – and wisely so, surely.

At about the same time as, and not unrelated to, the historical turn in philosophy of science, much of philosophy was undergoing a 'naturalistic turn'. This took place in epistemology and philosophy of mind generally, as well as in the philosophy of science in particular. Most philosophers started to accept as a serious constraint on their theorizing, that both human mental processes and human modes of acquiring knowledge are *natural*, happening in accordance with causal laws as do all other events in nature; and that philosophical attempts to achieve an understanding of the nature of these processes should be seen as continuous with scientific enquiry. This resulted in a plethora of

causal theories of mental and epistemic phenomena – with the development of causal theories of reference (Kripke, 1972; Putnam, 1975), of memory (Martin and Deutscher, 1966; Locke, 1971), of perception (Grice, 1961; Dretske, 1969; Goldman, 1976), of knowledge (Armstrong, 1973; Dretske, 1981; Goldman, 1986), and of justification (Goldman, 1979, 1986). Indeed, the dominant theory of the overall nature of the mind, which rose into ascendancy during this period, was functionalism, which sees mental states and events as individuated by their characteristic causal roles (Putnam, 1960, 1967; Lewis, 1966; Armstrong, 1968 – see section 3 below).

It became important, then, to see science, too, as a natural phenomenon, somehow recruiting a variety of natural processes and mechanisms – both cognitive and social – to achieve its results. Philosophers of science began to look, not just to history, but also to cognitive psychology in their search for an understanding of scientific activity (Nersessian, 1984b, 1992a; Giere, 1988, 1999a; Thagard, 1992). This trend is continued into the present volume, with a number of its philosophical authors appealing to psychological models or data in their chapters (e.g. Carruthers, chapter 4; Faucher *et al.*, chapter 18; Giere, chapter 15; Nersessian, chapter 7).

2.3 *Science and the social*

Our story so far has mostly been one of good news – with philosophy of science in the last century, like science itself, arguably progressing and/or getting somewhat closer to the truth. But one out-growth of the historical turn in philosophy of science was a form of social constructivism and relativism about science (Bloor, 1976; Rorty, 1979; Latour and Woolgar, 1986; Shapin, 1994). On this view, scientific theories are proposed, accepted and rejected in accordance with a variety of political and social forces which needn't have any connection with truth or reality, nor with reliable standards of evidence and rationality. Indeed, on this account the very idea of 'reality' as something set over and against our various socially constructed perspectival representations is unintelligible. The only real sense which can be made out of one theory being better or worse than another is in terms of its political–social influence.

While social constructivism has not found wide acceptance among philosophers of science generally (nor among the contributors to this volume in particular), it has perhaps played a useful role in emphasizing the social dimension of science and scientific activity. And one characteristic of recent work has been to combine elements from the social constructivist position with an overall realism about science – allowing that science progresses through varied social interactions among scientists, while arguing that those interactions can still be such as to facilitate increases in knowledge of the world (Kitcher, 1993, chapter 14 in this volume; Goldman, 1999; Thagard, 1999).

If one had to characterize the current state of play in contemporary philosophy of science – using very broad brush-strokes, of course – it would be that it is naturalistic and broadly realist in orientation; interested in both descriptive and explanatory accounts of scientific activity (historical, psychological and sociological); but also concerned with normative issues concerning the probity and/or reliability of various scientific practices. That picture pretty much represents, too, the orientation of most of the contributions to this volume.

3 Philosophy of mind: another short recent history

Since this book is as much about the mind as it is about science, in this section we provide a brief overview of the main developments to have taken place in the philosophy of mind over the last half-century or so.

3.1 Behaviourism and beyond

If our story of recent philosophy of science began with logical positivism, then our account of recent philosophy of mind has to begin with logical behaviourism, which also dominated much of the middle part of the twentieth century (Ryle, 1949; Wittgenstein, 1953). The leading idea of behaviourism is that it is a mistake to treat talk about the mental as talk about inner causes of overt behaviour. To think in this way, according to Ryle, is to commit a kind of *category-mistake*. Talk about the mental is not talk about mysterious inner causes of behaviour, but is rather just a way of talking about dispositions to behave and patterns of behaviour.

Behaviourism did have some attractions. It allowed humans to be included smoothly within the natural order by avoiding postulation of anything ‘ghostly’ inside the organic machinery of the body. It was thus able to reject any sort of ontological dualism, between non-physical minds and physical bodies. For the main objection to such a dualism has always been the problem of explaining how there can be any sort of causal commerce between the states and events of a non-physical mind and those of a physical brain or body.

The deficiencies of logical behaviourism were even more apparent, however. There were two main problems. One is that it seems quite implausible that knowledge of one’s own mind could consist in knowledge of one’s behavioural dispositions, since this hardly leaves any room for the idea of first-person authority about, or any kind of privileged access to, one’s own thoughts and feelings. (Hence the old joke about the two behaviourists who meet in the street – ‘You’re feeling fine’, says one, ‘But how am I?’)

The other major deficiency is this: logical behaviourism was offered as a piece of conceptual analysis. It was supposed to be an account of what had all along been the import of our psychological discourse. That being the Rylean

stance, a serious criticism of logical behaviourism is that it fails on its own terms, as an exercise in analysis. According to behaviourism what look like imputations of internal mental events or states should actually be construed as 'iffy' or conditional statements about people's actual and possible behaviour. The objection to the pretensions of behaviourist conceptual analysis, then, is that nobody has ever actually produced a single completed example of the behavioural content of such an analysis.

Indeed, there are principled reasons why no such behavioural analysis can be provided. For as Davidson (1970) pointed out, a particular belief or desire only issues in action together with, and under the influence of, other intentional states of the agent. There is no way, therefore, of saying what someone who holds a certain belief will do in a given situation, without also specifying what other beliefs and desires that agent holds. So analysis of a belief or a desire as a behavioural disposition requires invoking other beliefs and desires. This point has convinced practically everyone that Ryle was wrong. A belief or a desire does not just consist in a disposition to certain sorts of behaviour. On the contrary, our common-sense psychology construes these states as internal states of the agent which play a causal role – never singly, but always at least jointly – in *producing* behaviour.

3.2 *Physicalism and functionalism*

With dualism and logical behaviourism firmly rejected, attempts since the 1960s to give a philosophical account of the status of the mind have centred on some combination of physicalist *identity theory* with *functionalism* of one or another sort.

There are two distinct versions of identity theory which have been the focus of philosophical debate – *type-identity* theory and *token-identity* theory. According to the former, each type of mental state is identical with some type of brain state – for example, pain is the firing of C-fibres. According to token-identity theory, in contrast, each particular mental state or event (a 'token' being a datable particular rather than a type) is identical with some brain state or event, but it allows that individual instances of the same mental type may be instances of different physical types.

Type-identity theory was first advocated as a hypothesis about correlations between sensations and brain processes which would be discovered by neuroscience (Place, 1956; Smart, 1959; Armstrong, 1968). Its proponents claimed that the identity of mental states with brain states was supported by correlations which were just starting to be established by neuroscience, and that this constituted a scientific discovery akin to other type-identities, such as *heat is molecular motion*, *lightning is electrical discharge* and *water is H₂O*.

Most philosophers rapidly came to think that the early confidence in such type-correlations was misplaced, however. For consider a sensation type, such as pain. It might be that whenever *humans* feel pain, there is always a certain neurophysiological process going on (for example, C-fibres firing). But creatures of many different Earthly species can feel pain, and it is also possible that there are life-forms on other planets which feel pain, even though they are not closely similar in their physiology to any terrestrial species. So, quite likely, a given type of sensation is correlated with lots of different types of neurophysiological state. Much the same can be argued in the case of beliefs, desires and other mental kinds.

The conclusion drawn from these considerations was that type-identity theory is unsatisfactory, because it is founded on an assumption that there will be one-one correlations between mental state types and physical state types. Rather, we should expect mental state types to be *multiply-realized* in physical states. This is just what the thesis of token-identity affirms: each token mental state is identical with some token physical state; but instances of the same mental state type can be identical with instances of different physical types.

At about the same time, and connected with these debates concerning mind-brain identity, *analytic functionalism* was proposed as an account of the manner in which we conceive of mental states. The guiding idea behind functionalism is that some concepts classify things by what they *do*. So transmitters transmit, while aerials are objects positioned to receive air-borne signals; and wings are limbs for flying with, while eyes are light-sensitive organs for seeing with, and genes are biological structures which control development. Similarly, then, it was proposed that mental concepts are concepts of states or processes with a certain *function*, or distinctive causal role (Putnam, 1960, 1967; Lewis, 1966).

Functionalism seemed to be the answer to several philosophical prayers. It could account for the multiple realizability of mental states, since physiological states of a number of distinct types could nevertheless share the same causal role. And it had obvious advantages over behaviourism, since it accords much better with ordinary intuitions about causal relations – it allows mental states to interact and influence each other, rather than being directly tied to behavioural dispositions. Finally, it remains explicable that dualism should ever have seemed an option. For although (according to functionalists) we conceptualize mental states in terms of causal roles, it can be a contingent matter what actually *occupies* those causal roles; and it was a conceptual possibility that the role-occupiers might have turned out to be composed of some sort of *mind-stuff*.

There were two main problems with analytical functionalism, however. One is that it is committed to the analytic–synthetic distinction, which many philosophers think (after Quine, 1951) to be unviable. And it is certainly hard to decide

quite *which* truisms concerning the causal role of a mental state should count as analytic (true in virtue of meaning), rather than just obviously true. (Consider examples such as: that *belief* is the sort of state which is apt to be induced through perceptual experience and liable to combine with *desire* to generate action; that *pain* is an experience frequently caused by bodily injury or organic malfunction, liable to cause characteristic behavioural manifestations such as groaning, wincing and screaming; and so on.)

Another commonly voiced objection to functionalism was that it is incapable of capturing the felt nature of conscious experience (Block and Fodor, 1972; Nagel 1974). Objectors have urged that one could know everything about the functional role of a mental state and yet still have no inkling as to *what it is like to be in that state* – its so-called *qualia* or subjective *feel*. Moreover, some mental states seem to be conceptualized purely in terms of feel; at any rate, with beliefs about causal role taking a secondary position. For example, it seems to be just the feel of pain which is essential to it (Kripke, 1972). We seem to be able to imagine pains which occupy some other causal role; and we can imagine states having the causal role of pain which are not pains (which lack the appropriate kind of feel).

3.3 *Theory-theory*

In response to such difficulties, many have urged that a better variant of functionalism is *theory-theory* (Lewis, 1970, 1980; Churchland, 1981; Stich, 1983; Fodor 1987). According to this view, mental state concepts (like theoretical concepts in science) get their life and sense from their position in a substantive *theory* of the causal structure and functioning of the mind. To know what a belief is (to grasp the concept of belief) is to know sufficiently much of the theory of mind within which that concept is embedded. All the benefits of analytic functionalism are preserved. But there need be no commitment to the viability of an analytic–synthetic distinction.

What of the point that some mental states can be conceptualized purely or primarily in terms of feel? A theory-theorist can allow that we have *recognitional capacities* for some of the theoretical entities characterized by the theory. (Compare the diagnostician who can recognize a cancer – immediately and without inference – in the blur of an X-ray photograph.) But it can be claimed that the concepts employed in such capacities are also partly characterized by their place in the theory – it is a *recognitional* application of a *theoretical* concept. Moreover, once someone possesses a recognitional concept, there can be nothing to stop them prizing it apart from its surrounding beliefs and theories, to form a concept which is *barely* recognitional. Our hypothesis can be that this is what takes place when people say that it is conceptually possible that there should be pains with quite different causal roles.

Some or other version of theory-theory is now the dominant position in the philosophy of mind (which is not to say that there are no difficulties, and no dissenting voices, of course). And in many of its forms, theory-theory is of-a-piece with the sort of naturalism in philosophy which holds that philosophical and scientific enquiries are continuous with one another. From this perspective, both philosophy of mind and cognitive psychology are engaged in fundamentally the same enterprise – to characterize the nature and operations of the human mind.

4 Developments in developmental psychology

In this section we once again provide a very brisk recent history, this time in respect of developments in developmental psychology.

4.1 The Piagetian account

Piaget claimed that children's initial knowledge of relations of cause and effect is limited to what they see (in this respect his position was close to that of some logical positivists). In his early work he characterized their ideas about causality as restricted by 'syncretism' – in a word, by the tendency to connect everything with everything else (Piaget, 1928, p. 4). If asked to complete the beginning of a sentence such as, 'The man fell off his bicycle because . . .', children under five or six years will respond with, 'Because he broke his arm' rather than, say, 'Because he lost his balance.' On this basis, Piaget denied that young children are able to detect causal relations. In later work with Inhelder (Inhelder and Piaget, 1958), Piaget reiterated his view that children's scientific cognition is muddled and chaotic, and that their beliefs about events are juxtaposed together instead of causally linked.

Not only did Piaget judge young children to be incapable of identifying causal relations clearly, but he contended that they assign internal states and motives to inanimate objects. For example, they believe that inanimate objects – especially those that move – can possess consciousness and have sensations and emotions just as persons do. Piaget (1928, 1929) interpreted children's answers to questions about the movement and feelings of objects to indicate that their notions of causality are primitive, and reflect an inability to reason about the physical world.

According to the Piagetian analysis, moreover, the causal understanding of young children points to a suite of domain-general processes which underpin the nature of early cognitive development. Children's understanding across domains as diverse as geometry, physics and biology constitute a 'structured whole' in the sense that they share common properties in reasoning and problem-solving. What children know is tied to appearances rather than

involving underlying transformations and causal mechanisms which aren't visible to the eye. In conservation experiments, for example, young children typically believe that when water is poured from a short, wide glass into a tall, narrow one, the amount has changed even though nothing has been added or subtracted. Such children are also supposed not to understand the difference between animate and inanimate objects – believing, for example, that a shadow is a substance emanating from an object but participating with the night; and they attribute animistic qualities to leaves in projecting shadows. Only with increasing age do children know how shadows are projected and deny that objects cast shadows at night. Similarly, Piaget proposed that young children misunderstand the nature of dreams; they believe that dreams originate from outside and remain external to the self.

In later childhood, after the age of seven years, children can use transformations and invisible mechanisms in their causal reasoning. However, not until they achieve a formal operational understanding in early adolescence do they systematically test hypotheses, on a Piagetian approach. While much of this theoretical apparatus has been rejected in later work, many developmental psychologists continue to share Piaget's view that children's understanding undergoes radical conceptual change over the course of development (Carey, 1985; Wellman, 1990; Perner, 1991; Gopnik and Meltzoff, 1997).

4.2 *Modern evidence for early causal understanding*

We now know that Piaget significantly underestimated children's capacity for causal understanding. Even on the most sympathetic evidence, children's knowledge should be seen as variable rather than as constrained by domain-general stages (Siegler, 1994). It is now well documented that, although no single factor can fully explain children's inability to conserve, their responses on conservation tasks have much to do with the child's perception of the relevance and purpose of the context in which questions are asked (Donaldson, 1978; Siegal, 1997, 1999). Similarly, young children seem surprised when inanimate objects appear to move by themselves, or when they unaccountably appear or disappear (see the reviews by Carey, 2000b, and Rakison and Poulin-Dubois, 2001). Moreover, children as young as three years old have been shown to be very adept in distinguishing real from pretend objects and events in a simplified testing procedure where they are asked to sort items into those which are real (can be seen, touched and acted upon) and those which are not (Wellman and Estes, 1986; Leslie, 1994a).

Young children can also use causal knowledge in the fundamental process of naming and classifying artefacts. In a recent demonstration, Gelman and Bloom (2000) asked preschool children to name a series of simple objects. In

one condition, the objects were described as purposefully created (for example, ‘Jane went and got a newspaper. Then she carefully bent it and folded it until it was just right. Then she was done. This is what it looked like.’); in another, the objects were described as having been created accidentally (‘Jane was holding a newspaper. Then she dropped it by accident, and it fell under a car. She ran to get it and picked it up.’). Even three year-olds were more likely to provide artefact names (e.g. ‘hat’) when they believed that the objects were intentionally created, and material-based descriptions (e.g. ‘newspaper’) when they believed that the objects were created accidentally. All of this work points to an early capacity for causal understanding underlying children’s ability to classify objects in the physical world.

4.3 *Childhood science?*

How, then, can we characterize children’s causal understanding of the world – that is, children’s ‘science’? One proposal is that the child begins with naïve theories of the world (different in different domains of activity) which undergo transformation when encountering contrary evidence. This ‘theory-theory’ account claims that the growth of children’s scientific understanding is not dissimilar from that of adult scientists who revise their theories in the light of contrary evidence (Gopnik and Meltzoff, 1997). As Gopnik and Glymour (chapter 6 in this volume) put it, ‘The assumption behind this work is that there are common cognitive structures and processes, common representations and rules, that underlie both everyday knowledge and scientific knowledge.’ On this account, children’s scientific knowledge is built out of initial innate theories, which are then altered and fundamentally restructured in the light of observational evidence, by means of processes similar to those employed in scientific reasoning.

Note, however, that this ‘theory-theory’ account of childhood cognitive development is not the same as the ‘theory-theory’ account of our understanding of mental-state concepts, discussed in section 3 above. Philosophical theory-theorists, while agreeing with these developmentalists that the end-point of normal development is a theory of the workings and causal structure of the human mind, are not committed to any particular mechanism for reaching that end-point. On the contrary, they can be modularists (see below), or they can believe that the fundamental developmental process is one of *simulation* of the minds of others (Goldman, 1993 – note, however, Goldman himself isn’t a theory-theorist about the end-point of development), or whatever.

Similarly, ‘modular’ accounts of development, too, propose that children have mechanisms to detect causal relations from the start (Scholl and Tremoulet, 2000). These form the rudiments of scientific understanding. Thus children have core knowledge in areas such as number and syntax which provide a structural

foundation for the development of scientific understanding (Macnamara, 1982; Bloom, 2000; Wynn, 2000). But the process which carries children from their initial knowledge to their later understanding, in these domains, is not (or not fundamentally) one of theorizing, but rather one of biological maturation, on a modularist account.

Related considerations have to do with the unevenness of the process of development. For theory-theorists agree with Piaget, at least to the extent of thinking that the basic engine of cognitive development is domain-general, even if that engine has to work with materials of varying sophistication and complexity in different domains. Modularists, in contrast, will hold that the factors influencing cognitive development may vary widely across domains.

Nevertheless, on either theory-theory or modular accounts, children have a head-start in their causal understanding of the world. The questions are what – if anything – needs to be added in order to get genuine scientific thinking emerging; and how much of the initial cognitive beliefs and reasoning processes which are at work in childhood survive into mature scientific thought.

5 Recent psychology of reasoning

In this section we briefly review some of the recent psychological work on reasoning – and especially scientific reasoning, or reasoning related to science – in adults.

5.1 *Stage one: three traditions*

During the heyday of behaviourism, from the 1930s until about 1960, there was relatively little work done on human reasoning. In the decades after 1960, reasoning and problem-solving became increasingly important topics for experimental investigation, leading to the rise of three main research traditions, as follows.

- (1) Simon, Newell and their many students looked at problem-solving in tasks like the Tower of Hanoi and ‘Crypt-arithmetic’ (Newell and Simon, 1972). (Sample problem: Assign numbers to letters so that SATURN + URANUS = PLANETS.) The normative theory which guided this work was based on Newell and Simon’s ‘General Problem Solver’ (GPS) – a computer program for searching the problem space. The empirical work in this tradition seemed to indicate that people had the basic idea of the GPS, although they could easily get lost in the book-keeping details. Simon labelled the increasingly elaborate normative theory which emerged from this work ‘the logic of discovery’, and in later work he and his collaborators tried to simulate various episodes in the history of science (Langley *et al.*, 1987; Kulkarni and Simon, 1988).

- (2) In the UK, Wason (and later both Wason and Johnson-Laird), focused on the selection task and related problems (Wason, 1960; Wason and Johnson-Laird, 1970; Evans, Newstead and Byrne, 1993). The normative theory guiding this work was Popper's falsificationism. The empirical results showed that ordinary subjects were quite bad at using the strategy of seeking evidence which would falsify the hypothesis at hand.
- (3) Starting a bit later (in the late 1960s and early 1970s), Kahneman and Tversky began looking at various sorts of probabilistic reasoning (Kahneman, Slovic and Tversky, 1982). The normative theory they relied on was Bayesian. They made a great splash by finding a significant number of problems (base-rate problems, 'conjunction fallacy' problems, overconfidence problems and a host of others) on which subjects did very poorly indeed.

5.2 *Reactions*

The reactions and criticisms to these three lines of research on human reasoning were very different.

- (1) A number of people argued that the GPS strategy was much too simple for explaining how science actually works (e.g. Giere, 1989). But since the normative theory was open-ended, Simon and his collaborators often took the criticisms on board and sought to construct more complex problem-solving programs which could do a better job of simulating discoveries in the history of science. The approach has remained steadfastly individualistic, however, and thus it is often criticized by historians and sociologists of science for neglecting the social aspects of scientific discovery.
- (2) In response to the Wason and Johnson-Laird tradition, reactions were of two sorts. First, there was a great deal of experimental work looking at variations on the theme of the original selection-task experiments. Perhaps the best-known findings indicated that there are quite massive 'content effects' – there are *some* selection-task problems on which people do quite well. Many different explanations have been offered in an attempt to explain these content effects. The best known of these are the 'pragmatic reasoning schemas' proposed by Cheng and Holyoak (1985) and the social contract account first put forward by Cosmides (1989). (In addition, Sperber and colleagues have proposed a pragmatic explanation, based on relevance theory – see Sperber, Cara and Girotto, 1995.) Cosmides' work has been widely discussed in recent years because it is embedded within the theoretical framework of evolutionary psychology. The versions of the selection-task on which people do well, Cosmides argues, are just those that trigger special-purpose reasoning mechanisms which were designed

by natural selection to handle problems that would have been important for our hunter-gatherer forebears.

The second sort of reaction to the Wason/Johnson-Laird tradition was to challenge the Popper-inspired normative theory which had been assumed in analysing the original studies. Anderson's 'rational analysis' account was an influential first move along these lines (Anderson, 1990), and more recently, Chater, Oaksford and others have followed his lead (Oaksford and Chater, 1994). Moreover, Koslowski (1996) has argued in detail that once people's performance in causal and scientific reasoning tasks is interpreted in the light of a scientific realist perspective (particularly with the latter's commitment to underlying mechanisms), then much of that performance can be seen as normatively sensible and appropriate.

- (3) Perhaps because some of the major figures in the Kahneman and Tversky 'heuristics and biases' tradition made a point of stressing that their findings had 'bleak implications' for human reasoning, this work attracted the most attention and provoked a number of different lines of criticism. Some critics focused on alleged shortcomings in the experiments themselves – noting, for example, that the problems might be interpreted in ways that the experimenters did not intend (Adler, 1984). Cohen argued, in contrast, that the experiments could not possibly show that humans had an irrational reasoning 'competence', and thus that the results were at best the result of performance errors (Cohen, 1981).

Gigerenzer and others have mounted a sustained and very interesting attack on the normative assumptions that the heuristics and biases experimenters make (Gigerenzer, 2000). In case after case, he has argued that (for various reasons) it is far from clear *what* the 'right' or rational answer is to the questions posed to subjects. One theme in this critique has been that many statisticians favour a frequentist interpretation of probability; and on that interpretation, many of the heuristics and biases problems have *no* correct answer because they ask about the probabilities of single events not relativized to a reference class. Gigerenzer and his collaborators have carried out elegant experiments to demonstrate that reformulating some of the heuristics and biases problems in terms of frequencies rather than single event probabilities can dramatically improve performance.

In recent years, Gigerenzer has put an evolutionary spin on these results, claiming that our minds evolved to deal with probabilistic information presented in the form of 'natural frequencies' (Gigerenzer, Todd and the ABC Research Group, 1999; Gigerenzer, 2000). He has joined forces with Cosmides and Tooby and other evolutionary psychologists to argue that our performance on many of the heuristics and biases problems can be explained by the fact that they are not the sorts of problems that our minds evolved to deal with.

5.3 *An emerging synthesis?*

In the last few years, a number of people, notably Evans and colleagues (Wason and Evans, 1975; Evans and Over, 1996), and more recently Stanovich (1999), have proposed ‘dual-processing theories’ which may accommodate the findings (and many of the theoretical arguments) from both those within the heuristics and biases tradition, and their critics. On this account, reasoning is subserved by two quite different sorts of system. One system is fast, holistic, automatic, largely unconscious and requires relatively little cognitive capacity. The other is relatively slow, rule-based, more readily controlled and requires significantly more cognitive capacity. Stanovich speculates that the former system is largely innate and that, as evolutionary psychologists suggest, it has been shaped by natural selection to do a good job on problems similar to those which would have been important to our hominid forebears. The latter system, by contrast, is more heavily influenced by culture and formal education, and is more adept at dealing with the problems posed by a modern, technologically advanced and highly bureaucratized society. This new, slow system is largely responsible for scientific reasoning. Stanovich also argues that much of the individual variation seen in heuristics and biases tasks can be explained by differences in cognitive capacity (more of which is required for the second system), and by differences in cognitive style which lead to different levels of inclination to *use* the second system.

Important questions for our understanding of science to emerge out of this new synthesis include the extent to which scientific reasoning is influenced by the implicit system, and how scientific practices can control for or modify such influences. (See Evans, chapter 10 in this volume, for a discussion of these issues.)

6 **Themes and connections: a guide through the volume**

In this section we will say just a few words about each of the chapters in this volume, picking up and identifying a number of recurring themes and issues as we go along.

6.1 *Part one*

The first three chapters in the main body of the book all relate in one way or another to the question of the innate basis of scientific reasoning.

Mithen (chapter 2) distinguishes some of the key elements of scientific thinking, and then traces the evidence of their emergence from the early hominines of some 4 million years ago to the first human farming communities of 11,500 years ago. Mithen emphasizes that the cognitive components of scientific

reasoning are multiple (cf. Dunbar, chapter 8), and he sketches how they probably appeared at different stages and for different reasons in the course of human evolution.

Atran (chapter 3) explores the universal cognitive bases of biological taxonomy and taxonomic inference, drawing on cross-cultural work with urbanized Americans and forest-dwelling Maya Indians. His claim is that there is a universal, essentialist and innately channelled appreciation of species as the causal foundation for the taxonomic arrangement of life forms, and for inferences about the distribution of causally-related properties of living beings.

Carruthers (chapter 4) examines the extent to which there are continuities between the cognitive processes and epistemic practices engaged in by human hunter-gatherers, on the one hand, and those which are distinctive of science, on the other. He argues that the innately channelled architecture of human cognition provides all the materials necessary for basic forms of scientific reasoning in older children and adults, needing only the appropriate sorts of external support, social context and background beliefs and skills in order for science to begin its advance. This contradicts the claims of those who maintain that 'massive reprogramming' of the human mind was necessary for science to become possible (Dennett, 1991).

6.2 *Part two*

The seven chapters which make up part two of the book examine aspects of contemporary scientific cognition in both children and adults.

Varley (chapter 5) reports the results of a series of experiments testing the folk-scientific abilities of two severely a-grammatic aphasic men. Since the capacities of one of the two men were near-normal, Varley argues that core components of scientific reasoning, at least – specifically, hypothesis generation and testing, and reasoning about unseen causes and mechanisms – must be independent of language. This is an important result, since it is often claimed that scientific thinking is dependent upon language (Dennett, 1991; Bickerton, 1995).

Gopnik and Glymour (chapter 6) take their start from 'theory-theory' accounts of child development, according to which children propose and revise theories pretty much in the manner of adult scientists. They argue that the theories in question are best understood as 'causal maps', and that the recent development of computational 'Bayes nets' may provide the resources for us to understand their formation and change, in both science and child development.

Nersessian (chapter 7) makes a strong plea for enquiries into historical episodes in science to be seen as one important source of understanding of the cognitive basis of science. Her focus is on the cognitive basis of the model-based

reasoning practices employed in creative thinking, and on the way in which these can lead to representational change (or 'conceptual innovation') across the sciences.

Dunbar (chapter 8) reports and discusses the results of a series of *In Vivo* studies of science, in which the reasoning processes of scientists were observed and recorded 'on line'. He argues that these results support a pluralist conception of scientific activity, in which types of cognitive process which are used elsewhere in human activity are deployed in distinctive patterns and sequences in the service of particular goals, and with different such patterns occurring in different sciences and in different aspects of scientific activity. Dunbar also points to evidence of cross-cultural variation in scientific reasoning practices (cf. Faucher *et al.*, chapter 18).

Koslowski and Thompson (chapter 9) emphasize the important role of collateral information ('background knowledge') in scientific reasoning, both in proposing and in testing scientific theories. This role has generally been ignored (and often explicitly 'factored out') in psychological studies apparently demonstrating that naïve subjects are poor scientists. With the role of collateral information properly understood, Koslowski and Thompson present evidence that school-age children are able to reason quite appropriately in experimental contexts.

Evans, too (chapter 10), is concerned with the effects of background belief on reasoning, but from a different perspective. He reviews the experimental evidence of various forms of 'belief bias' in people's reasoning. He discusses the implications of this data for scientific practice, drawing conclusions which enable him (like Koslowski and Thompson) to be generally sanguine about the prospects for a positive assessment.

Finally in part two of the book, Hilton (chapter 11) is also concerned with the psychological evidence of irrationality in people's reasoning, this time in their reasoning about causality. While somewhat less optimistic in his approach than the previous two chapters, he does think that a significant proportion of this data can be explained in terms of (perfectly sensible and appropriate) pragmatic factors, and that those irrationalities which remain can be mitigated by appropriate social arrangements (a point which is also made by Evans, and which is further emphasized by Kitcher, chapter 14).

6.3 Part three

The three chapters making up part three of the book all share a concern with the place of motivation and emotion within science. Traditionally, science (like human reasoning generally) has been seen as a passionless enterprise, in which emotion can only interfere. Scientists are supposed to be *dispassionate*

enquirers. The separation of reason from emotion has come under vigorous attack (Damasio, 1994). And all three of the contributions in part three of the book continue that critique, focusing on scientific reasoning especially.

Thagard (chapter 12) argues against prevailing models in the philosophy of science by claiming that emotional reactions are an integral and ineliminable part of scientific practice, in all three of the main domains of *enquiry*, *discovery* and *justification*. The emotions which mostly concern him are broadly truth-directed ones such as *interest*, *surprise* and an aesthetic–intellectual response to *beauty* (‘elegance’, ‘simplicity’, etc.).

Hookway (chapter 13) explores some of the ways in which epistemic emotions such as doubt and dogmatism may help as well as hinder the pursuit of knowledge, especially in the way that they help guide enquiry by making certain questions salient for us while leading us to ignore others.

Kitcher (chapter 14) is concerned with the influence of rather more mundane and materialistic motivations – such as a desire for fame or for prestigious prizes – on scientific practice. He sketches a research programme for investigating the impact of different social arrangements and inducements on the conduct of science, and a framework for thinking about how such arrangements should be assessed within democratic societies such as our own.

6.4 Part four

The four chapters in part four of the book are, in various ways, about the social dimension of scientific cognition. (This is an important theme in a number of other chapters as well, including those by: Carruthers, chapter 4; Dunbar, chapter 8; Evans, chapter 10; Hilton, chapter 11; and Kitcher, chapter 14.)

Giere (chapter 15) argues that science can be better understood if we notice the extent to which scientific cognition is *distributed*, incorporating many factors outside the minds of individual scientists. These would include scientific instruments, libraries, calculations conducted with the aid of computers and a variety of forms of social arrangement and social structure.

Siegal (chapter 16) critically examines the idea proposed by Carey (1985), that childhood development may involve stages of ‘strong conceptual change’ analogous to revolutionary periods in science. He concentrates on the domain of biology in particular (and there are a number of connections here with the ideas of Atran, chapter 3). Siegal suggests that for key aspects of biology, the evidence for conceptual change in the strong sense is inconclusive, and that children’s understanding is both domain-specific and highly sensitive to the information they receive from their surrounding culture.

Harris (chapter 17) makes a powerful case for recognizing the importance of *testimony* in the development of childhood beliefs, contrasting this position with a view of the child as a ‘stubborn autocrat’ (a view quite common among

developmental psychologists), which sees the child as a lone little scientist, gathering data and forming and testing hypotheses for itself.

Faucher *et al.*, too (chapter 18), mount an attack on the sort of ‘child-as-scientist’ models proposed by Gopnik and others, arguing that even if they are right about childhood they are wrong about science. And, like Harris, Faucher *et al.*, too, emphasize what children acquire from the surrounding culture during development. But Faucher *et al.*’s emphasis is also on the acquisition of *norms* from surrounding culture, especially norms concerning the gathering of evidence and the assessment of theories. They also discuss recent data suggesting that culture-specific norms can have a deep and lasting impact on the operations of our cognition.

7 Conclusion

These are exciting times, not only for science and the scientific understanding of science, but also for our understanding of the human mind. The philosophy and psychology of science and scientific practice – like science itself – continues to make progress, and to raise challenging new questions. It is our view that the chapters in this volume should contribute substantially to that continued advance.

