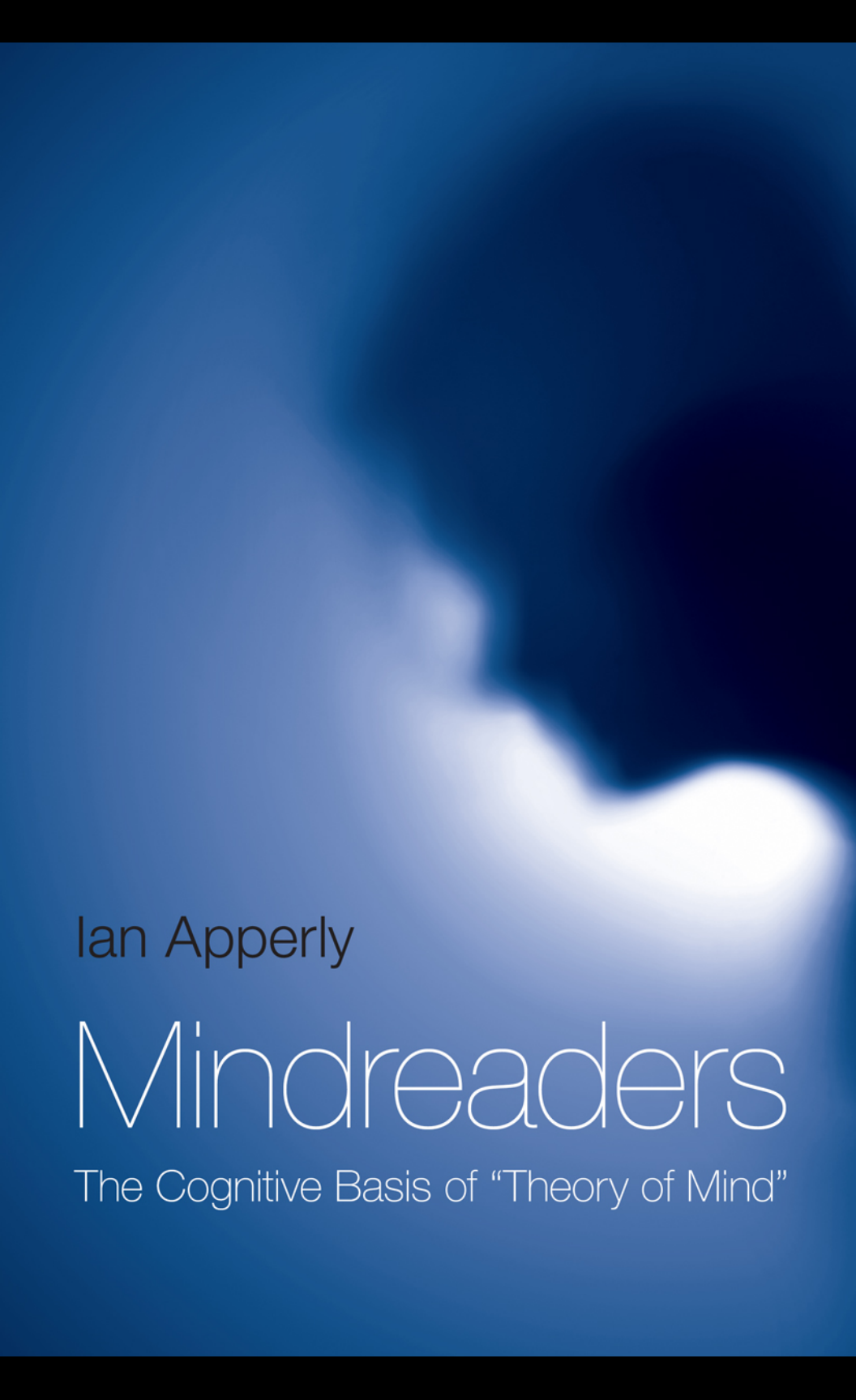
Ian Apperly
Mindreaders
The Cognitive Basis of "Theory of Mind"

Mindreaders

Theory of mind, or "mindreading" as it is termed in this book, is the ability to think about beliefs, desires, knowledge and intentions. It has been studied extensively by developmental and comparative psychologists and more recently by neuroscientists and cognitive psychologists. This book is the first to draw together these diverse findings in an account of the cognitive basis of "theory of mind", and establishes the systematic study of these abilities in adults as a new field of enquiry.

Apperly focuses on perceptions, knowledge and beliefs as paradigm cases of mindreading, and uses this as a basis from which more general lessons can be drawn. The book argues that an account of the cognitive basis of mindreading is necessary for making sense of findings from neuroscience and developmental and comparative psychology, as well as for understanding how mindreading fits more broadly into the cognitive system. It questions standard philosophical accounts of mindreading, and suggests a move away from the notion that it consists simply of having a "theory of mind".

This unique study into the cognitive basis of mindreading will be ideal reading for academics and advanced students from the diverse disciplines that have studied theory of mind in particular, and social cognition more generally.

Ian Apperly is currently a Reader in psychology at the University of Birmingham. He studied natural sciences at Cambridge University, and completed his PhD at the University of Birmingham.

Mindreaders

The Cognitive Basis of "Theory of Mind"

Ian Apperly

Psychology Press
Taylor & Francis Group
HOVE AND NEW YORK

Published in 2011
by Psychology Press
27 Church Road, Hove, East Sussex BN3 2FA

Simultaneously published in the USA and Canada
by Psychology Press
270 Madison Avenue, New York, NY 10016

This edition published in the Taylor & Francis e-Library, 2011.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.

Psychology Press is an imprint of the Taylor & Francis Group, an Informa business

Cover design by Andy Ward
Cover photo © eva serrabassa/iStockPhoto

This publication has been produced with paper manufactured to strict environmental standards and with pulp derived from sustainable forests.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Apperly, Ian.
Mindreaders : the cognitive basis of "theory of mind" / Ian Apperly.
p. cm.
ISBN 978-1-84169-697-3 (hb)
1. Cognitive psychology. 2. Philosophy of mind. I. Title.
BF201.A66 2010
153–dc22

2010019000

ISBN 0-203-83392-9 Master e-book ISBN

ISBN 978-1-84169-697-3 (hbk)

Contents

Here, as he looked at the table, at the malachite cover of her blotter, and an unfinished letter lying on it, his thoughts suddenly underwent a change. He began to think of her, of what she was thinking and feeling. For the first time he really pictured to himself her personal life, her ideas, her desires; and the notion that she could and should have a separate life of her own appeared to him so dreadful that he hastened to drive it away. This was the abyss into which he was afraid to look. To put himself in thought and feeling into another being was a mental exercise foreign to Karenin. He considered such mental exercise harmful and dangerous romancing.

(Leo Tolstoy, 1954, pp. 159–160)

List of figures

Preface and acknowledgements

Before I left to begin my PhD in 1995 my undergraduate supervisor gave me a good piece of parting advice: ''Don't work on theory of mind – the bubble is about to burst.'' What he correctly diagnosed was a tailing off in the 15 years of frenetic research activity that had followed seminal papers on the ability of chimpanzees, typically developing children and children with autism, to reason about the knowledge, beliefs, desires and intentions of others. It was not so much that all of the interesting problems had been solved, but that key theoretical positions had become firmly entrenched, and there were just so many experiments that nobody could keep all of the details in their head at once in order to make sense of them. This, at least, was my excuse. I half listened to my undergraduate supervisor, and spent five very interesting years studying theory of mind in ''older'' children, which at the time meant 5- to 7-year-olds rather than 3- to 4-year-olds. But although my participant group was unusual, and although this somewhat avoided the ''neurotic task fixation'' of the literature on younger children, this work was squarely in the existing tradition of research on conceptual development. And as predicted, there was a sense that many researchers were growing just a bit bored with endless talks and endless papers on theory of mind.

What was difficult for anyone to foresee was that research on theory of mind, or ''mindreading'' as I shall call it in this book, would be rejuvenated in just a few years. Impetus for this change has come from at least three distinct disciplines. Firstly, just as many researchers were converging on the view that chimpanzees could not mindread there came a series of studies with much more positive evidence, not only in chimpanzees, but in other species, including distant relations of humans, such as birds. Secondly, evidence that infants might pass suitably adapted mindreading tasks has shaken the long-standing conclusion that the preschool years are the main period of change in mindreading abilities. Thirdly, the trickle of pioneering papers on the neural basis of mindreading in the mid nineties has grown rapidly into a tide by joining with the new research programme of ''social neuroscience''. Meanwhile, research on young children has continued to make advances, just a little more quietly than before.

Although these developments are very exciting, there is a curious absence in the literature on mindreading. Despite the many hundreds of papers on the abilities of human brains, children, infants and clinical populations, and non-human animals, there remains very little research on mindreading in human adults. This is curious since, in comparable topics, such as spatial cognition or number cognition, cognitive psychological research on how adults think about space or number provides the core model for understanding what children must develop, what brains must process and what might be different about the cognition of clinical populations or non-human animals. It is almost as if researchers investigating mindreading are building the walls of a tower but omitting to put in the floors.

Before I put off too many potential readers, I am not suggesting that existing research is wrong-headed when viewed on its own terms. As will become clear, I shall draw heavily on this work throughout the book both for inspiration and for a wide range of empirical findings. What I shall argue, however, is that besides being a worthy project in its own right, investigating the cognitive basis of mindreading, including studies of adults, provides critical connections between research in each of the sub-disciplines in which mindreading has typically been studied. Making such connections leads to a range of novel empirical questions and empirical predictions for many different areas of research on mindreading. Moreover, findings that appear very puzzling when viewed from the perspective of one sub-discipline, such as infants' success on tasks that seem very difficult for 3-year-olds, might be less puzzling if viewed in the broader perspective of the mature system.

My interest in the cognitive basis of mindreading grew out of my PhD work on children's development. With Liz Robinson I examined children's ability to think about the particular ways in which people represented information, and how this constrained reference and description when talking about what they thought, or interpreting what they said. This forced us to think about mindreading as an on-line process, rather than just a set of concepts or a body of knowledge that children needed to acquire, and it stretched the limits of what it seemed possible to achieve with the tasks typically used in research with children. It seemed natural to look for inspiration in research on adults, where there were many methods available for studying on-line cognitive processes. What I found surprising was that nobody had used such methods to study mindreading in adults, and more surprising still was the realization that despite all of the work on development, this meant that we actually knew rather little about what mindreading abilities children eventually developed. I have been very lucky to have a number of collaborators – especially Dana Samson – who were willing to share their invaluable expertise to address this question, and to have a head of department, Glyn Humphreys, who was willing to back our rather speculative ventures.

It is inevitable that a book of this kind will not do justice to the full range and depth of careful experimentation and clever theorizing that exists in the

literature on mindreading. As a developmental psychologist by training I am particularly aware that this is true for the variety of accounts that have set out to explain how children's mindreading emerges from earlier, simpler processes. This book does not present an alternative to these accounts of development. It does aim to present a different, broader perspective on mindreading. Among other things this should have implications for how children develop these abilities, and I hope that this gain justifies my abridgement of many interesting and important ideas.

I owe many people thanks for making this book possible. Liz Robinson planted the idea in my head, where, for some time, it looked set to stay. I am grateful to the publisher for being patient, my friends for not reminding me too often, and the School of Psychology for granting me the sabbatical that eventually gave me time to write. Being much more used to collaborative work, writing as a single author is strange, and brings with it an uncomfortable sense of laying claim to ideas that I know are not wholly my own. This goes in particular for Dana Samson and Steve Butterfill who have been co-conspirators in the development of so many of the ideas presented here, and whom I cannot thank enough.

Sarah Beck was an invaluable sounding board throughout the process, and with an office adjacent to mine, bore the brunt of many problems over endless coffee. Even after this she generously gave her time to provide feedback and encouragement on draft chapters, as did Steve Butterfill, Charlotte Easter, Kevin Riggs, Dana Samson and Laura Shapiro.

Finally, I thank my partner, Laura, who made me read *Middlemarch* before I started.

1 Introduction

What is "mindreading"?

I know what you're thinking. Just in case you've come here by mistake let me be clear that there is nothing mystical about the kind of mindreading I shall discuss. But even without mysticism there remains a good deal that is interesting, surprising and, at times, mystifying, about our everyday ability to "get inside the heads" of other people and think about what they know, want, intend and believe.

Reasons for interest include claims that mindreading is uniquely human (e.g., Penn & Povinelli, 2007; Saxe, 2006), that it is at the heart of key cognitive processes for social interaction and communication (e.g., Grice, 1989; Sperber & Wilson, 1995), and that impaired mindreading may be part of the explanation for a variety of psychiatric and developmental disorders such as schizophrenia and autism (e.g., Baron-Cohen, 1995; Frith, 2004). Reasons for surprise include the discovery of apparently sophisticated mindreading abilities in human infants and some non-human species (e.g., Call & Tomasello, 2008; Emery & Clayton, 2009) alongside apparently poor mindreading abilities in human children (e.g., Carpendale & Lewis, 2006; Doherty, 2009), and sometimes in human adults (Apperly, Samson & Humphreys, 2009). And mindreading is mysterious because there are genuine conceptual puzzles about how it is even possible to know the minds of others. Most obviously, we do not have direct access to what other people know, want, intend or believe, but must infer these mental states on the basis of what they do and say. Although this is clearly something that we actually do on a regular basis, how we do it raises some significant theoretical and empirical questions that have only been partially addressed in the literature. My overall aim is to justify the reasons for interest in mindreading, but to suggest that some of the surprises and some of the mysteries are the result of the literature conceptualizing the problem of mindreading in a rather narrow way.

By way of illustration, researchers have often asked me "Why are you studying adults? Don't they already have a theory of mind?". This question seems very telling about the issues that have dominated research on

mindreading. Firstly, the vast majority of work to date has focused exclusively on *when* children are able to reason about mental states such as beliefs, desires and intentions and what factors affect this development. Secondly, many researchers hold that the development of mindreading consists of acquiring abstract mental state concepts, and that these concepts constitute a "theory" about how the mind works. This has led to "theory of mind" becoming the predominant term for mindreading in the academic literature (though "folk psychology", "mentalizing" and "social cognition" are also common). More importantly, this has led to the view that once such concepts are present – and this presence is evidenced by young children's success on simple tasks that require reasoning about beliefs, desires, etc. – psychologists can pack up their toolkits and move on to another problem; there is really nothing further to explain.

I find this baffling. Most cognitive psychologists only study adults, but appear to hold down their jobs despite the fact that young children already have a wide range of cognitive abilities, including aptitude in "conceptual domains" such as number, space, time or physics. The general pattern is that the literatures on adults' cognition not only have lives of their own, but also provide the model of the system that children are developing. Surely there is no reason for mindreading to be an exception?

The emphasis on concepts in the literature on mindreading also gives a deceptive impression of simplicity. It feels like we know what we mean when we credit a child who passes false belief tasks with the concept of belief, but do we really? It should make us nervous that there is little consensus in the theoretical literature on what it means to have a concept. And do we really suppose that unless a child passes a test suggesting that she has a given mental state concept then she is quite insensitive to the way in which behaviour is governed by mental states? Or that once a child does pass such a test there is nothing further for her to learn? Probably most researchers of mindreading would baulk at such black and white characterizations of their views, but common habits of thought in the literature nonetheless proceed along these lines. Witness the fact that the very great majority of work on mindreading concerns the abilities of 3- to 5-year-old children with little attention to younger children and still less on older children and adults. I shall argue that research on mindreading – including research on children – would benefit from a broader perspective on the kinds of data and participant groups that may be of interest and on the kinds of cognitive process that may be at work. We can keep many of the insights behind conceptual change accounts of the development of mindreading, but also accept that having key mental state concepts might be neither the beginning nor the end of mindreading abilities.

So I take as my starting point the fact that we live in a world populated by agents whose behaviour can be understood by supposing that they have mental states (beliefs, desires, etc.) that interact according to psychological laws (e.g., people act to satisfy their desires on the basis of their beliefs).[1] If

we take this as the problem domain of mindreading then our question is, what cognitive processes enable us to operate successfully in this domain of agents? Part of the answer to this question is surely that at least some organisms – adult humans at the very least – reason in a sophisticated way about the causes and consequences of mental states *as such*.

But this is unlikely to be the full answer, even for adult humans. To see why, briefly consider another domain, that of normal objects (i.e., non-agents) whose behaviour can be understood by supposing that they have physical characteristics (mass, velocity) that interact according to physical laws (e.g., an object maintains its velocity unless acted on by a force). Some of the cognitive processes that enable human adults to operate in this domain clearly involve reasoning about physical characteristics and physical laws, and such reasoning can be very sophisticated. But what happens when we catch a ball? Of course, a suitably educated adult with sufficient time could make reasonable estimates of the relevant physical parameters, derive a prediction about the ball's trajectory and their own capacity to move, and work out the point of intersection where a catch would be possible. But this is demonstrably not how people catch balls[2] (e.g., McLeod, Reed & Dienes, 2003), or even how they arrive at everyday intuitions about the behaviour of balls[3] (e.g., Hood, 1995; Hood, Carey & Prasada, 2000; McCloskey, 1983; McCloskey et al., 1983). In fact, when operating in the domain of physical objects, people appear to have a wide variety of tricks and strategies, rules of thumb and specialized cognitive processes. These enable useful work in the domain but do not necessarily operate over conceptual representations of forces and physical characteristics, or at least not the same ones as would have been learned in physics classes. It seems plausible that the same will be true for mindreading.

Why "mindreading"?

Readers already familiar with the literature will know that most researchers talk about "theory of mind" rather than "mindreading". I have done the same in publications, because this is what everyone expects, but it always chafes, for two reasons. The important one is that the term is enormously tendentious because it implies that mindreading consists in having a theory about how the mind works. In fact this is, at best, just one theoretical possibility among many. And for my purposes it unhelpfully suggests that "theory of mind" is something that one *has* rather than something that one *does*. The more trivial reason is that "theory of mind" is linguistically awkward. It is bad enough as a sort of compound noun for describing a putative psychological faculty, and much worse as an adjective to describe "theory of mind processes" or "theory of mind brain regions", and as for inflection, if I have a theory of mind do I really want to be a "theory of minder"?

Since I shall have to use the term a lot, I propose to talk about "mindreading" instead. I prefer the term because it is theoretically neutral, yet

captures something of the character of the problem. When we read text we not only process the words in front of us but also make elaborated inferences about the meaning behind them, which go far beyond what is on the page. When we mindread we often have to process perceptually accessible social stimuli, and we always have to make significant further inferences to arrive at the underlying thoughts, desires, knowledge or intentions that we cannot directly perceive. Others have opted for other terms, such as "social cognition" or "social understanding" (e.g., Carpendale & Lewis, 2006), but these are much too broad for my purposes. I shall keep returning to the point that there is a very great deal more to social cognition than mindreading alone, and this is discussed in many other sources that I shall refer to along the way. But mindreading is the topic of this book.

Epistemic mental states as a case study

We have very rich mental lives. I have already mentioned beliefs, knowledge, desires and intentions, and we could keep going for some time charting further kinds of epistemic, emotional and motivational states. A full account of mindreading will ultimately require an explanation of how we navigate a social world of agents who have all of these mental states, not only for completeness but because different mental states combine together as the causes and consequences of behaviour. However, we are not only very far from having the evidence required for a full account, but there remains a good deal of work in conceiving the form such an account should take. Since my aim is to examine the cognitive basis of mindreading in greater depth than is typical in the literature, this comes, for the time being, at the expense of examining a broad range of mental states. My simplifying strategy is to limit most of my discussion to epistemic mental states such as perceptions, knowledge and belief. Rightly or wrongly, these occupy the centre of the literature on mindreading, giving me the largest evidence base to draw upon. My hope is that the account developed for epistemic mental states can serve as a model for the broader family of mental states, and ultimately for a fully integrated account of mindreading.

Normative models and philosophical theories

From its inception the modern study of mindreading has involved close collaboration between psychologists and philosophers. When Premack and Woodruff (1978) asked whether the chimpanzee has a "theory of mind" it was the commentaries of three philosophers (Bennett, 1978; Dennett, 1978; Pylyshyn, 1978) that led to paradigms involving perspective differences – such as the false belief task (Wimmer & Perner, 1983) – becoming the workhorses of empirical studies for the next thirty years. Some of the most prominent psychologists in the field have had strong interests in philosophy

of mind, and full-time philosophers of mind have continued to play an important role in the development of empirical research (e.g., Carruthers & Smith, 1995; Davies & Stone, 1995a,b). It is philosophy of mind that supplies the normative model for mindreading. Agents (you and I) hold attitudes (e.g., seeing, knowing and believing) to propositional mental states (e.g., It is raining, again). The mindreader's task is therefore conceived as reasoning about such relationships. As you see me leave my house you may predict my next action – that I will return to fetch an umbrella – if you can judge that I will see that it is raining, that I know that this will make me wet, and that I believe there is an umbrella back inside the house. As we shall see in later chapters, the literature is dominated by questions about how children become able to reason this way, and whether non-human animals ever do so.

Philosophy of mind also supplies the two dominant theoretical accounts of how mindreading is possible. According to so-called "theory-theories" mindreading depends on us having mental state concepts and principles that describe their interactions (Davies & Stone, 1995a,b). This contrasts with "simulation theories", according to which mindreading may not require a fully specified theory of the mind if only we could use our own minds to model (i.e. simulate) the minds of others.

However, although philosophy of mind has made many positive contributions to the study of mindreading, I think it is also partly responsible for the topic becoming rather hidebound. Although the debate between theory-theory and simulation-theory may have given us a useful sense of the different flavours that mindreading might have, it has been remarkably poor at generating empirically testable predictions for propositional mental states such as seeing, knowing and believing.[4] Although there are particular cases that seem to favour one theory or the other, nobody has yet come up with a generalizable test, or set of criteria, that could be used to discern whether a given mindreading problem was solved by simulation or by theorizing (e.g., Apperly, 2008; Stich & Nichols, 1997). I am not the first, and probably not the last, to wonder whether experimental psychologists might be better off without these theories. At the very least that we might do well to set them aside while we pursue empirically tractable hypotheses, and perhaps return to them later to see if they offer us any further insight into what we have found. This is the strategy that I shall pursue here.

I do not propose to be as gung-ho about the normative model, which characterizes mindreading as the ascription of attitudes with propositional content to agents. It will be a working hypothesis that we do in fact mindread in something like this way at least some of the time, and that it is important to know how we do so. However, I do not want to be tied exclusively to this characterization, firstly because I suspect that our very best mindreading abilities might only approximate to the assumptions of the normative model, and secondly because I suspect that much of the work of everyday mindreading – and possibly all of the mindreading of infants and

non-human animals – is achieved in ways that deviate very substantially from the normative model. This theme will be picked up in more detail in the second part of the book.

Mindreading perceptions, knowledge and beliefs: The structure of the book, and how to read it

My interest, then, is in what cognitive processes make it possible to operate successfully in a world of agents who have perceptions, knowledge and beliefs. Although surprisingly few studies have tackled this question directly, the existing literature on adults, children, infants, atypical development, cognitive neuroscience and non-human animals provides much useful information to help construct an account. My aim in the first part of the book, from chapters 2 to 5, is to identify a set of questions about the cognitive basis of mindreading, and use these to conduct a very selective review of these literatures. Much more complete reviews and assimilations of different branches of the literature have been conducted elsewhere, and I shall draw attention to these as I go along. The unique features of my review will be the breadth of evidence considered and the focus on what can be learned about cognitive processes. In the second part of the book, in chapters 6 and 7, I shall draw together these different strands of evidence to develop a schematic account of the cognitive basis of mindreading. I shall then explore implications of this account for the study of mindreading in children, adults and non-human animals.

This structure means that readers who are most interested in the theory could skip ahead to chapters 6 and 7. Chapter 6 begins with a distillation of evidence from earlier chapters, and this may be sufficient, if you are willing to take my word for it. However, much of the evidence comes from very recent studies conducted in different corners of the literature. There remains a great deal of controversy about how they should be interpreted, and I discuss these debates together with the detailed findings throughout chapters 2 to 5. I hope that many readers will find it worthwhile to engage with these details because, ultimately, it is this evidence, and the patterns that emerge across diverse tasks and participant groups, that really justifies the effort of taking a new look at mindreading in the later chapters.

A cognitive psychologist's wish-list

Cognitive psychology does not have a universal taxonomic system for characterizing cognition. But to give some structure to my discussion of the literature I shall distinguish between questions about representations, processes and architectures for mindreading. These dimensions are clearly inter-related, but different cognitive processes nonetheless vary with substantial independence on each.

What are the representational characteristics of mindreading? Is mindreading "special"? In what sense does mindreading depend on a specific set of concepts? Does mindreading pose unusual representational problems? Is the representational basis of mindreading domain-specific (that is to say, specialized for the purpose), or is it domain-general (that is to say, the same representational resources serve other cognitive functions)? What role does language play in mindreading?

What are the processing characteristics of mindreading? Is mindreading effortful? Does it make significant demands on scarce cognitive resources for working memory, inhibition or other aspects of "executive function". Is it "fast" or "slow"? To what degree is mindreading automatic? Will mindreading always occur in the presence of certain stimuli or will its occurrence depend on context?

What are the architectural characteristics of mindreading? Is mindreading a unitary faculty? Are there distinct sub-components serving different functions (such as processing perception versus knowledge)? Is there a hierarchy with different systems for simple mindreading and more complex mindreading, using different representations and processes? If so, what information passes between them? How is mindreading integrated with other cognitive processes such as communication or action planning? Might cognitive processes for communication or action planning have their own "on board" system for mindreading, so that a given mindreading function (such as processing perception) is performed more than once by separate systems that do not talk to each other?

It is not that these questions have never been asked before in the context of mindreading. All of them have. And, indeed, some people think they have been answered already. It is not uncommon for researchers to simply state that mindreading depends on a specialized cognitive module that is innate, fast, automatic and domain-specific (e.g., Leslie, 2005; Sperber & Wilson, 2002; Stone, Baron-Cohen & Knight, 1998). Curiously, others are similarly trenchant in the view of the development of mindreading as a protracted process that depends on a series of conceptual insights gained through domain-general learning and reasoning (e.g., Gopnik & Meltzoff, 1997; Perner, 1991; Wellman, 1990). Clearly, these contradictory interpretations derive, in part, from contradictory theoretical commitments about the nature of cognition in general and development in particular. But I think they also result from questions about the cognitive basis of mindreading cropping up piecemeal as the by-product of questions about the nature of mindreading in children or non-human animals, the neural basis of mindreading, or the functional basis of some other process such as communication. I want to ask these questions in a more systematic way, using them to interrogate different fields of the literature for relevant evidence. My aim is to construct something closer to a coherent picture of the cognitive basis of mindreading (chapter 6), which I shall then use to cast light back on the existing literatures that I have drawn upon (chapter 7).

Unspoken problems at the heart of mindreading

So far, so good. My aim is to understand the cognitive processes that make it possible to mindread perceptions, knowledge and beliefs, and we should be able to make progress in this direction by applying a cognitive psychologist's questions to the existing literature on mindreading. I want to keep this structure for the first half of the book. But it may be useful to have some further questions in the back of the mind because of their importance for the second half. These questions concern two related problems that are not much discussed in the literature on mindreading, but which seem important for developing a cognitive account.

The first problem is that the proposed functions of mindreading appear to make contradictory cognitive demands. Consider the following examples, both of which appear to require mindreading. As a member of the jury in a court of law, your job is to judge the guilt or innocence of the person in the dock. Thinking about the defendant's mental states will clearly be critical to this judgement. Did they see the victim spoon the white powder into their tea? Did they know that the white powder was poison not sugar (e.g., Young, Cushman, Hauser & Saxe, 2007)? Did they leave the white powder next to the kettle with the intention that the victim should put it in their tea or were they just careless? Did they know that the poison would kill the victim rather than just make them sick? Of course, this is just one fanciful example. But if we scale up from the complexity already inherent in this simple example, and consider the range of cases that are tried in courts of law it is clear that a juror must be extremely flexible in their mindreading abilities. A case could concern any conceivable person in any conceivable set of circumstances where almost any conceivable information might turn out to bear on the judgement of guilt or innocence. Such flexibility is not expected to come easily. Indeed, we expect jurors to take their time to assimilate the relevant information, to think about it carefully, and to discuss it with fellow jurors before reaching an opinion. These characteristics seem common across a wide range of other situations. Viewed this way, mindreading would seem to be as flexible as any general reasoning process, and as potentially demanding of limited general processing resources for memory and executive control.

On the other hand, consider the case of playing a competitive sport like football. The success of passing the ball depends on whether or not opponents nearby the path of the ball will be able to see it in order to intercept it, and of course on whether the receiving player on your team is in a position to take account of what you are doing. Faking a pass in order to trick an opposing player to commit themselves in the wrong direction seems to depend upon deliberately giving them a false belief about your intentions. Realizing that an opponent cannot see your team mate behind them who is waving for the ball clearly confers a competitive advantage. So

playing football surely involves some sort of mindreading (and indeed, mind manipulation), but to be of any use at all, these judgements must be made extremely quickly. And in order not to interfere with other on-going processes involved in playing the game, it would seem very useful, if not essential, that these judgements were made very efficiently, making few demands on limited general processing resources.

So it seems that mindreading must be as flexible as any reasoning process, but at the same time, fast and efficient enough to guide judgements made on the fly. The problem is that these demands are largely contradictory. That is to say, in cognitive systems flexibility and efficiency tend to be negatively correlated. The difficulty in having one system that is both flexible and efficient is apparent from the high prevalence of "two-systems" accounts in cognition, whereby in a given domain, be it social categorization, number cognition or general reasoning, the contradiction is resolved by having two types of cognitive system that operate in the domain, which make complementary trade-offs between flexibility and efficiency. What this means for current purposes is that, although it is common to talk about mindreading (or theory of mind, mentalizing or folk psychology) as if it were one kind of cognitive process, we should not be too surprised if this turned out to be incorrect. Indeed, in chapter 6, I shall argue for a "two-systems" account of mindreading.

A related point that has also not been confronted very clearly in the literature is that mindreading comes up very hard against some deep conceptual problems in cognitive science. Csibra and Gergely (2007) discussed the "inverse problem" in relation to the task of making inferences about an agent's goals and actions. Put simply, the mindreader who wants to infer an agent's goal (or intention, or belief, etc.) from observation of the agent's action (or to predict the agent's action given their goal) is faced with the problem of many-to-one mappings between goals and actions. A given action could be evidence for any number of different goals, and a given goal could lead to any number of actions. This presents a significant problem for inferences about the relations between observed behaviour and mental states. A related concern is the "frame problem" or the "problem of relevance" (e.g., Pylyshyn, 1987). In the general case, the frame problem is to know what are the causes or consequences of a change in a system whose components are all potentially (though not necessarily) linked. Applied to mindreading, the problem is that an agent may have any number of beliefs (and other mental states), any of which might be relevant when trying to judge what the agent will think or do in a given situation. It follows that even the most conscientious juror with all the time in the world should not be able to judge the guilt of the person in the dock: the problem is computationally intractable.

Clearly jurors do make decisions, and footballers do make passes, and so an important objective of the second half of the book is to discuss how this might be possible.

Notes

1 There is, of course, a longstanding debate in philosophy of mind about whether agents really do have such mental states, or whether assuming that they do is just a parsimonious strategy for explaining and predicting their behaviour (e.g., Dennett, 1987; Fodor, 1975). However, for current purposes this makes no difference because both sides of the debate agree that mindreading involves thinking about mental states, whether these are real entities or useful fictions.

2 A series of studies by McLeod and colleagues suggest that fielders running to catch a ball do not explicitly decide their speed or direction of running (as they would if they computed the solution to simultaneous equations governing their own trajectory and that of the ball). Instead, their running is governed by two parameters recovered from visual tracking of the ball's trajectory: ". . . they [fielders] run so that their angle of gaze elevation to the ball increases at a decreasing rate, while their horizontal gaze angle to the ball increases at a constant rate . . ." (McLeod et al., 2003, p. 244).

3 People's everyday reasoning about falling balls, as well as many other physical interactions, is subject to systematic errors and biases reflecting the use of rules of thumb or "naïve theories". For example, McCloskey (1983) found that when adults are asked to predict what happens to a ball dropped by a moving person, many people judge that it will fall straight downwards from the point at which it is dropped. In fact it will follow a parabolic trajectory resulting from the combination of the forward velocity with which it starts and the downward acceleration due to gravity.

4 A much more convincing case can be made that some mindreading of non-propositional mental states, such as action schemas and emotions, is achieved via simulation (e.g., Goldman, 2006). But it clearly does not follow from this success that the debate between simulation-theory and theory-theory will carve at the joints of propositional attitude ascription, which is the topic of this book.

2 Evidence from children

Although the modern literature on mindreading originated with a question about chimpanzees (Premack & Woodruff, 1978), much of the running since then has been made in studies of human development. Not only is this by far the largest literature on mindreading, but many of the central theoretical themes and most prominent empirical paradigms have arisen out of research on children. The current chapter, therefore, is not just designed to provide evidence bearing on questions about the cognitive basis of mindreading, but also to give important background information for the rest of the book. In this chapter I shall describe evidence from typically developing children aged 2-years upwards, followed by evidence from atypical development. These studies set the long-standing benchmarks in terms of theories, paradigms and, until recently, age of acquisition for children's mindreading.

This literature has been cast in a new light by studies emerging in the last few years that suggest that infants have much more impressive mindreading abilities than previously thought. I shall describe these studies in the next chapter. Although this ordering might seem odd, it is much easier to appreciate why the recent studies of infants are so interesting, surprising and controversial if they can be seen against the background of the literature on older children.

Perspective differences as a key empirical device

In their commentaries on Premack and Woodruff (1978), Bennett (1978), Dennett (1978) and Pylyshyn (1978) cautioned that we could never be sure that an animal (or anyone else for that matter) was taking account of another agent's mental states if the animal's own mental states coincided with those of the other agent. Let's imagine we were nearing the end of a long meeting, and your mind started wandering to the subject of beer. You might think it a minor miracle of everyday mindreading if I suggested that we go to the pub. But you might be less sure of my skills if you then found out that this was precisely what I wanted to do myself. It would be a much better test if I had a blistering hangover, so that going to the pub was the