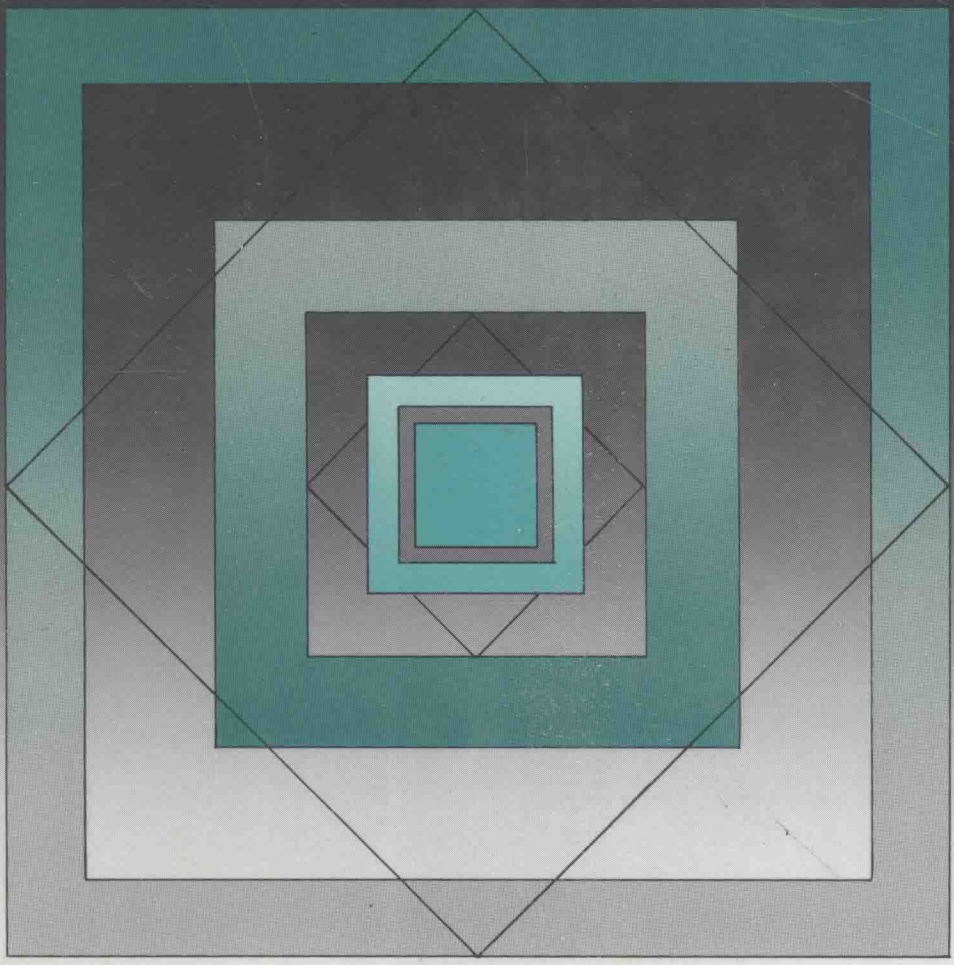
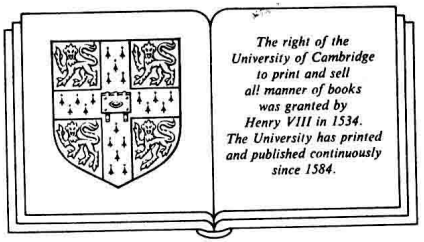
DEVELOPING THEORIES OF MIND



Edited by
JANET W. ASTINGTON, PAUL L. HARRIS, & DAVID R. OLSON

DEVELOPING THEORIES OF MIND

Edited by
Janet W. Astington
Ontario Institute for Studies in Education
Paul L. Harris
University of Oxford
David R. Olson
Ontario Institute for Studies in Education



CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
NEW YORK NEW ROCHELLE MELBOURNE SYDNEY

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1988

Printed in Canada

Library of Congress Cataloging-in-Publication Data

Developing theories of mind.

Papers from the International Conference on Developing Theories of Mind, University of Toronto, May, 1986, and the Workshop on Children's Early Concept of Mind, St. John's College, Oxford, June, 1986.

1. Cognition in children – Congresses. 2. Children – Attitudes – Congresses. I. Astington, Janet W. II. Harris, Paul L. III. Olson, David R. IV. International Conference on Developing Theories of Mind (1986 : University of Toronto) V. Workshop on Children's Early Concept of Mind (1986 : St. John's College University of Oxford)

BF723.C5D468 1988 155.4'13 87–34141

British Library Cataloguing in Publication Data

Developing theories of mind.

1. Children. Mental development
I. Astington, Janet W. II. Harris, Paul L.
III. Olson, David R.
155.4'13

ISBN 0 521 35411 0

Preface

In this book psychologists are developing theories to account for children's developing theories of mind. Its neatly ambiguous title, for which we thank Alison Gopnik, comes from the International Conference on Developing Theories of Mind, organized by Janet Astington, Lynd Forguson, Alison Gopnik, and David Olson through the McLuhan Program in Culture and Technology at the University of Toronto in May 1986. Many of the contributors to that conference subsequently participated in the Workshop on Children's Early Concept of Mind, organized by Paul Harris at St. John's College, Oxford, in June 1986. Both conferences generated an excellent set of papers, a very high level of constructive discussion, and an enormous amount of excitement and enthusiasm. The similarity of theme, participants, and goal was so great that the papers were combined into this single volume, jointly edited from Toronto and Oxford. In the year following the conferences, drafts and revised drafts of the papers were exchanged among the authors, as can be seen in the numerous arguments and agreements to which they refer in the published chapters. The result is much more than a set of conference proceedings: It is a coherent, sustained attack on a set of fundamental issues in developmental psychology and cognitive science. The volume, therefore, should be of interest to academics, researchers, graduate students, and advanced undergraduates who are interested in the nature and development of children's understanding of mental life.

We would like to acknowledge the support provided to the Toronto conference by the Connaught Foundation through a grant to the McLuhan Program. We are also grateful for support provided by the Faculty of Arts and Science, the Division of Life Sciences of Scarborough College, the Department of Psychology, and University College, all of the University of Toronto, and by Field Services and Research of the Ontario Institute for Studies in Education. In Oxford, support and hospitality were provided by St. John's College. We would like to thank all those who participated in the conferences, both those who contributed papers to the volume and all the other participants, including Inge Bretherton, Jerome Bruner, Susan Carey, Colin McGinn, and Andrew Woodfield. We also take this oppor-

tunity to thank Sylvia Wookey for ensuring that the Toronto conference ran smoothly, and Denese Coulbeck for secretarial assistance during preparation of the manuscript for publication. Finally and especially, we would like to thank Helen Wheeler, editor at Cambridge University Press, for her advice and guidance, and for expediting publication of the volume.

J.W.A.
P.L.H.
D.R.O.

Contributors

JANET W. ASTINGTON, Ontario Institute for Studies in Education
CAROLE R. BEAL, Dartmouth College
GARY BONITATIBUS, Saint Anselm College
MICHAEL CHANDLER, University of British Columbia
CAROL FLEISHER FELDMAN, New York University
JOHN H. FLAVELL, Stanford University
LYND FORGUSON, University of Toronto
ALISON GOPNIK, University of Toronto
DANA GROSS, University of Minnesota
PAUL L. HARRIS, University of Oxford
JÜRGEN HOGREFE, University of Salzburg
CARL NILS JOHNSON, University of Pittsburgh
ALAN M. LESLIE, MRC Cognitive Development Unit, University of London
DAVID R. OLSON, Ontario Institute for Studies in Education
JOSEF PERNER, University of Sussex
DIANE POULIN-DUBOIS, Concordia University
JAMES RUSSELL, University of Cambridge (previously University of Liverpool)
MARILYN SHATZ, University of Michigan
THOMAS R. SHULTZ, McGill University
BEATE SODIAN, University of Munich
MARJORIE TAYLOR, University of Oregon
HENRY M. WELLMAN, University of Michigan
HEINZ WIMMER, University of Salzburg
ILAN YANIV, University of Michigan

Contents

1

Introduction

DAVID R. OLSON, JANET W. ASTINGTON, and
PAUL L. HARRIS

Sometime after they learn to talk but before they begin formal schooling, children come to display a new understanding of perception, action, and talk that is symptomatic of a new sensitivity to the life of the mind. Children begin to recognize themselves and others as "things which think," as things which *believe*, *doubt*, *wonder*, *imagine*, and *pretend*. This, in itself, is a remarkable achievement as it marks their coming to make a systematic distinction between the world and *mental representations* of the world. But perhaps even more remarkable, the achievement of this new understanding of mind appears to spill over into a number of apparently unrelated domains, including children's understanding of the distinction between appearance and reality, and their understanding of the relation between an utterance and its interpretation. Indeed, the repercussions are sufficiently broad that it is not misleading to think of children's new understanding of mind as constituting a new stage or level of intellectual development.

To characterize just what is achieved and how it is achieved is the major focus of this volume. Although accounts vary somewhat, there is agreement that this development does not consist simply of the addition of a new piece of knowledge. Rather, it involves a fundamental alteration or shift in children's orientation to their knowledge. Their mental representations cease to be transparent and invisible. Instead, they become *opaque*. Children begin to recognize mental states in themselves and others; they come to recognize beliefs as beliefs, desires as desires, and intentions as intentions. That recognition reflects children's acquisition of a *theory of mind*, a set of explicit and interconnected concepts for representing those representational states; that is, it involves the development of a set of *meta-representations*. The chapters of this volume examine young children's acquisition of this theory of mind, the impact of its acquisition on children's understanding of their own and others' behavior, its role in their consciousness of their own mental states, and its role in their ascription of mental states to others. As such, these chapters represent a state-of-the-art report

This chapter was written collaboratively; order of authorship indicates successive contributions.

on the conceptual and empirical advances being made in an important new area of research on children's cognitive development.

As the evidence for the impact of children's acquisition of a theory of mind on their talk, action, and interaction mounts, a number of related questions comes to the fore. Primary among them is how best to characterize this achievement. Is the development a matter of acquiring a general capacity for recursion, for embedding relations within higher-order relations? Is it a matter of acquiring a set of quasi-theoretical concepts for interpreting behavior? Is it a matter of becoming self-conscious, of learning to think about representations instead of thinking about the world? Or is it a matter of coming to understand the various causal connections between states of the world and intentional states of the mind such as thinking, knowing, and believing?

A second set of questions revolves around how we are to explain this remarkable achievement, the acquisition of a theory of mind. Is the explanation to be found in some fundamental and endogenous reorganization of the child's cognitive processes? Or is it to be found in the socializing practices of adults who impart a particular way of talking and thinking about language and action?

A third set of questions concerns the epistemological status of the child's theory of mind. Do children, irrespective of culture, discover a set of intrinsic and universal properties of the mind? Alternatively, have they acquired a "folk theory," a set of concepts about mental life that is culturally transmitted, just as any other set of beliefs is, whether true or not? The appeal to mental acts and mental states, that is, to *intentionality*, in cognitive psychology is one of the vexed problems in the cognitive sciences. On the one hand theorists such as Fodor (1981), Pylyshyn (1984), and Searle (1983) defend the view that beliefs, desires, and intentions are essential to psychological explanation. On the other, theorists including Quine (1960), Dennett (1978b), Churchland (1984), and Stich (1983) maintain that psychological explanation cannot be based on the ascription of intentional states and that meta-representations are at best part of a "language game," a social psychology, or a "folk psychology" that will sooner or later fall to the advance of a realist, empirical cognitive theory. This volume will not, of course, resolve that issue. But a careful examination of children's acquisition of a set of concepts for representing their own mental states and those of others, and the dramatic effect that such an achievement has on their behavior, may permit us to see the problem from a fresh perspective.

Acquisition of a theory of mind

Why can this accumulation of achievements be deemed the acquisition of a *theory* of mind? A theory, as Wellman (Chapter 4) reminds us, consists

of a referential domain, that is, the set of events to be explained and predicted, an interrelated set of concepts that make up the theory, and a set of rules for using the theoretical concepts to explain and predict events in the referential domain. Consider, briefly, what is involved in having a theory of planetary motion. The events referred to are the orbital motions of a set of heavenly bodies, the theoretical terms are *planets*, *gravity*, and *centripetal force*, and the theoretical concepts can be used to explain and predict events, such as orbital planetary motion. If the theory successfully explains the events, the entities or forces identified by the theoretical terms – planets, gravity, and centripetal force – are treated as real entities.

So what is a theory of mind? The events to be explained and predicted are talk and action (some would say behavior). The theoretical concepts are those of *belief*, *desire*, *intention*, and *feeling*. And, third, these concepts may be used to explain and predict the events in the referential domain, namely, talk and action. Finally, if the theory provides the best explanation and prediction of the events in the referential domain, the entities specified by the theoretical terms may be treated as real entities.

Children sometime between their second and sixth years, at least children in western cultures, acquire such a theory. They begin to acquire the relevant set of mental predicates or concepts, beginning with *pretending* in their second year, as Leslie (Chapter 2) shows, and they quickly acquire the lexical terms for a set of mental concepts including *know*, *think*, *remember*, *forget*, *dream*, *pretend*, and the relations among those theoretical terms, as Wellman (Chapter 4) shows. Then they begin to understand, predict, and explain their own and others' talk and action by means of the concepts expressed by those mental terms, as Perner (Chapter 8), Wimmer, Hogrefe, and Sodian (Chapter 9), Flavell (Chapter 13), and others show. These meta-representational concepts are also theoretical in that they represent states that are not directly observable but inferred from overt action and speech and used to understand some properties of talk and action. In acquiring these meta-representational concepts, children are acquiring a "folk psychology," part of what Forguson and Gopnik (Chapter 12) refer to as the commonsense view of the world. It is these concepts that permit children to reflect on their own and others' mental states. It is important to stress that having a theory of mind goes beyond merely possessing such states. It involves the possession of concepts of those states. Because these concepts represent such states as beliefs, desires, intentions, and feelings, they constitute representations of representations and in this sense constitute a recursive or meta-representational system.

Premack and Woodruff (1978) were among the first, at least in recent times, to ask whether a theory of mind is the unique possession of human beings. They claimed that the behavior of chimpanzees could be most perspicuously explained by allowing that they imputed wants, beliefs, and intentions to others, that is, that they possessed a theory of mind. How-

ever, critics such as Dennett (1978a) noted that the training required to test the hypothesis that chimpanzees have beliefs about beliefs in others, could engender beliefs and habits that in themselves could explain the behavior of the animals without appealing to beliefs about beliefs.

Children, of course, are immature members of a different species, as Chandler (Chapter 20) points out, and they may have a theory of mind as part of their native endowment. Moreover, unlike chimpanzees, children are linguistic creatures, and, exercising some ingenuity, one can ask them about their own and others' mental states and so determine if they possess the relevant concepts and make the relevant distinctions. Further, by experimentation one can determine if and when they use a theory of mind in predicting and explaining the talk and action of themselves and others. Indeed, this volume consists of reports of ingenious devices that researchers have used to examine children's explicit concepts about their own and others' minds, and their recourse to such concepts in explaining and predicting what they or other people might do or say.

In arguing that it is helpful and appropriate to think of children's acquiring a theory of mind, we should underline certain caveats. First, in claiming that children have a theory of mind, we do not intend to suggest that they have a theory in which the referential domain is the mind itself. We mean only that in their explanation and prediction of talk and action, children have recourse to mental constructs. Thus, their theory, strictly speaking, is of action and talk; the mind enters as a theoretical construct, not as a domain requiring explanation. Similarly, it can be misleading, strictly speaking, to describe Newton as advancing a theory of gravity, if by that phrase one implies that the referential domain was gravity. Rather, Newton advanced a theory of planetary motion and of falling bodies, in which gravity served as a theoretical postulate. He explicitly avoided advancing an explanation of gravity itself.

A second caveat concerns the importance of explicit talk about the mind as opposed to predictions that appear to presuppose the mind. When children appear to take a given entity into account in making a prediction, we are not thereby entitled to infer that children do actually deploy that entity as a theoretical construct. Consider, for example, Piaget's observation of infants in the first year of life (Piaget, 1937/1954). In attempting to catch sight again of a rapidly moving object, they frequently anticipate where the object will have fallen and scan the floor. In some sense, they appear to take the consequences of gravity into account. Yet we clearly would not wish to assert that infants employ gravity as a theoretical construct on the basis of such observations. Similarly, from the mere fact that children appear to take beliefs into account in predicting behavior, we are not thereby entitled to infer that children employ beliefs as a theoretical construct in their predictions. However, children offer us much more evi-

dence for their theory of mind than simple prediction. They make explicit reference to the theoretical entities that motivate their predictions. Thus, they refer explicitly to what they or others *know*, *think*, and *pretend*. As various authors show (Wellman, Chapter 4; Harris & Gross, Chapter 15), children are often capable of backing up the judgments and predictions that they make by reference to the relevant mental states. We cannot rely just on children's use of mental terms, as Perner (Chapter 14) emphasizes; it is the converging evidence from prediction and explanation that strengthens our conviction that children do indeed adopt a theory of mind.

The empirical findings reported in this volume and the interpretations of these findings are far too rich to summarize briefly. Yet there is enough agreement between both the data presented and the authors' interpretations to state the primary achievements involved in developing a theory of mind. First, children achieve some means for disconnecting – "decoupling," as Leslie (Chapter 2) describes it – representations from the things they are representations of. Sometime in their second year these detached representations become embedded in representational functions such as pretending. Second, children acquire a set of concepts for representing mental activities such as *thinking*, *dreaming*, *imagining*, and *pretending*, and the terms for referring to these concepts, sometime between their second and fourth years. Third, they become skilled in using these concepts for predicting and explaining actions premised on false beliefs, present beliefs discrepant from prior beliefs, appearances discrepant from reality, intentions discrepant from actions, utterances discrepant from beliefs and intentions, and facial expressions discrepant from actual feelings. This third achievement, the use of mental concepts to understand and predict what is said and done, begins, at least for children in our cultures, when they are about 4 years of age. Some would argue that only when these concepts function in the theoretical way just mentioned, can the child be credited with a theory of mind. Even then it will be some time before children are sufficiently skilled in using that theory of mind to understand such issues as perceptual and linguistic ambiguity, and the relativity of knowledge.

How are we to account for these relatively systematic changes? A number of factors are implicated. The swift and regular emergence of children's understanding of their own and others' minds between 2 and 4 years of age suggests that the development is, at least in part, maturational, that children at that age become capable of recursive operations enabling them to represent the contents of an earlier or nonveridical representation. Second, the fact that children have recently become language-using creatures when they begin to form meta-representations, such as those implicated in pretend play, suggests that language or a more general symbolic capacity may be an important element in the development of this understanding. Third, the fact that children are acquiring a mentalistic vocab-

ulary about this time suggests that the particular concepts represented by that vocabulary play an important role; they appear to serve as a distinctive set of predicates for the formation of embedding (and embedded) assertions, such as "John *pretends* that the banana is a telephone" or "John *thinks* that the chocolate is in the cupboard." The extent to which language in general, and such distinctive predicates in particular, are relevant to the achievement of a theory of mind remains an important but unresolved issue.

A brief overview of the volume

The acquisition of a theory of mind is not a simple matter but involves a number of achievements. The early steps, taken in the second and third year of life, are discussed in Part I. Thereafter, the child elaborates the basic theory in a variety of domains. Two domains in particular have been fruitful for empirical research: the child's understanding of the relationships among perception, knowledge, and reality (considered in Part II), and the child's understanding of the relationship between minds engaged in various types of dyadic interaction (considered in Part III). Further implications of children's acquisition of a theory of mind are discussed in Part IV. Below we give a brief overview of the child's initial theory, and its subsequent elaboration.

I. Developmental origins of children's knowledge about the mind

The first and perhaps the most important step in the development of a theory of mind is the ability to form meta-representations. This ability requires that children isolate, or "decouple," as Leslie calls it (Chapter 2), primary representations of the world from their normal input–output functions. In pretend play, for example, a banana that in its primary representation is something to eat, comes to be represented in a secondary way as, say, a telephone. How can the secondary representation be formed without confusing the child about the intrinsic properties of bananas? To resolve that predicament, Leslie proposes that representations are "decoupled" from the things they are representations of and so become eligible for a secondary representation that can violate the normal reference, truth, and existence properties of the primary representation. Secondary representations with these properties, Leslie calls "meta-representations," and they are expressed by such predicates as *pretend* and *think*. Leslie suggests that this meta-representational ability provides the basis for the child's acquisition of these mentalist terms.

Not only do children begin to show evidence of a new ability to operate on their representations, as suggested by their pretend play, they also

begin to report on their own mental states and on the general properties of these states, as Johnson (Chapter 3) and Wellman (Chapter 4) point out. Johnson argues that a primary source for such mental concepts is children's conscious experience of their own thoughts, dreams, and feelings, and that the theory of mind provides an expression for these conscious experiences. As Johnson points out, such experiences are rich and comprehensible. His claim that they are an adequate source for a theory of mind provides a possible counterargument to the claims of Sellars (1963), Nisbett and Wilson (1977), Quine (1960), and others, to the effect that people have little or no direct introspective access to higher mental processes and that what passes for introspection is often a retrospective interpretation along commonsensical or "folk theory" lines.

Wellman shows that 3-year-olds have an explicit understanding of the distinction between real objects and events and decoupled representations such as dreams, thoughts, and images. They have also mapped out some of the interdependencies among these concepts. It seems clear that children by this age possess the concepts needed for a theory of mind, but the fact that they cannot use these concepts to understand cases in which action is premised on a false belief, for example, suggests important limitations to that theory.

Unlike decoupled representations, the creation of primary representations depends upon appropriate perceptual access. Children's ability to assess another person's knowledge must depend, in part, on their awareness of whether or not that person had the appropriate perceptual access. Yaniv and Shatz (Chapter 5) show that even 2½- to 3-year-olds are aware of the conditions affecting perceptibility: seeing, hearing, smelling, and touching. From these observations they infer that young children have theories about perceptual access and its relation to mental processes. Although the latter part of this claim is still open to question because Wimmer, Hogrefe, and Sodian (Chapter 9) show that 3-year-olds are uncertain of what others know on the basis of what the other sees and hears.

Children must also understand the concept of intention if they are to predict and explain behavior, since a person's action is dependent not merely on what he knows or believes but also on what he desires or intends. Poulin-Dubois and Shultz (Chapter 6) review research on children's attribution of intentions to others, showing the beginnings of this ability in children as young as 2½ to 3 years of age. Even younger children show some knowledge of others' agency, if not of their intentionality, and Poulin-Dubois and Shultz discuss how children's understanding of the concept of intention may arise out of their concept of agency.

These cognitive achievements are reflected in children's linguistic ones; if children can embed mental representations in other representations, they can also do so in language, as Feldman shows (Chapter 7). Indeed,

Feldman suggests that the linguistic device for turning predicates into subjects, a fundamental recursive procedure that she observes in young children's language, may be the device for turning representations into meta-representations. In this way a mental attitude, treated as an object of discourse, acquires the status of a mental state.

II. Coordinating representational states with the world: Understanding the relationships among perception, knowledge, and reality

Somewhat later, roughly in their fifth year, children begin to ascribe false beliefs to other people, in the attempt to predict and explain their behavior. In an important paper, Wimmer and Perner (1983) reported that between 4 and 6 years of age children come to anticipate that others will act in accordance with the beliefs they hold, even if they are false, rather than in accordance with the way things are in the world. In the current volume, they take their initial findings in somewhat different directions.

For Perner (Chapter 8), the crucial new competence at this stage is the ability to represent the process of representation, and Perner would not apply the term "*meta*-representational" to children's abilities until this stage is reached. Two-year-olds can form representations, including hypothetical and counterfactual representations, and can compare these models to the world (that is, to their model of the world) so that they can enter a pretend scenario and can tell you that dreams are not real. However, 2-year-olds cannot represent the process of modeling whereas 4-year-olds can represent that process. Only then can they understand that someone entertaining a false belief takes it to be a true representation of the real world, and is the world in which he will act, even though it *mis*represents the actual situation.

An alternative proposal for what changes between the beginnings of meta-representations in the 2-year-old and the management of false belief by the 4-year-old is that the child acquires a new understanding of the sources of information about the world. Wimmer, Hogrefe, and Sodian (Chapter 9) show that the older, but not the younger, children recognize the role of perception and communication in the acquisition of beliefs, and suggest that children's success with false belief and appearance–reality tasks is a manifestation of their new understanding of informational conditions. Younger children, by contrast, failing to grasp this relation, discount the fact that the protagonist in a false belief story has not had perceptual access to the true state of affairs. Wimmer et al. also show that by 6 years of age children come to understand inference as a source of information. They emphasize that perception, communication, and inference function as informational sources for even the youngest children tested; what is gained between 4 and 6 years is an understanding of these processes.