1

An introduction to the debate

HELEN TAGER-FLUSBERG, SIMON BARON-COHEN, AND DONALD COHEN

This book focuses on a psychological theory of autism that has generated considerable interest in the past decade. It is known as the *theory of mind hypothesis of autism*. By 'theory of mind' is meant the ability of normal children to attribute mental states (such as beliefs, desires, intentions, etc.) to themselves and other people, as a way of making sense of and predicting behaviour. The theory of mind hypothesis of autism holds that in children with autism, this ability fails to develop in the normal way, resulting in the observed social and communication abnormalities in behaviour. We felt that this hypothesis deserved to be subjected to close and critical scrutiny, by leading authorities in the fields of psychology, psychiatry, and related disciplines, for several reasons.

First, if autism is indeed caused by a failure to develop a theory of mind, then studying autism might hold the clues to how this important ability is normally acquired so effortlessly. Secondly, studying autism from this perspective might show us what happens to a child when this ability is not available in the normal way. Thirdly, this hypothesis has been surrounded by fascinating debate about the role of affect and the nature of the cognitive mechanisms involved in supporting a theory of mind. These debates are important because resolving them will teach us about the relevant *processes* in development and pathology. Finally, we felt that subjecting the theory to scrutiny might help to reveal both its strengths and shortcomings, so as to guide future research in the field of autism.

For all of these reasons, we put together a book to debate the theory of mind hypothesis of autism. Most of our contributors presented their chapters in draft form at a two-day workshop in Seattle, in April 1991, with a set of key questions to guide our discussions: How do children acquire a theory of mind? What are the developmental origins of this ability, and what is its evolutionary history? Is autism a syndrome that is best understood in terms of a primary impairment in this capacity?

WHAT IS AUTISM? A NOTE ON DIAGNOSIS

Kanner's (1943) description of the syndrome of autism is a classic example of the contribution of clinical observation to psychiatric taxonomy. In his clear prose, he described a group of children who had impoverished or absent social relations from the very first years of life, and with language (when it was present) which was distinctively deviant. Although there have been modifications, Kanner's diagnostic criteria have proved remarkably robust and have been echoed in all subsequent psychiatric classification systems. Thus, thirty years after Kanner, Rutter (1978) reviewed the major studies and highlighted four essential features of autism: impaired social development; delayed and deviant language; insistence on sameness; and onset before 30 months. A similar set of features formed the basis of the diagnostic criteria in the American Psychiatric Association taxonomy in the 1980 edition of the Diagnostic and Statistical Manual (DSM-III), and in the World Health Organization taxonomy in the (1987) International Classification of Diseases, 9th Edition (ICD-9 1987).

To provide a more developmental approach, and in order to encompass the broad range of individuals with autism, the definition of autism was elaborated in the next edition of the Diagnostic and Statistical Manual (DSM-IIIR 1987). Thus, although the same three features are retained, DSM-IIIR provides a range of diagnostic items for each of these. For example, for the social impairment, it ranges from marked lack of awareness of others (for those with the most severe social impairments) to gross abnormalities in peer relations (for those who are least impaired). The next planned revisions of diagnostic criteria will appear in DSM-IV and ICD-10, both due to be published in the mid-1990s.

Changes in diagnostic criteria are not merely of academic interest: they have implications for which individuals receive the diagnosis of autism. For example, DSM-III and ICD-9 criteria encompass most of the children whom most clinicians would categorize as having autism. In contrast, DSM-IIIR criteria are broader and include some individuals whom some clinicians might feel fall outside the usual domain of the concept.* The shifting boundaries of the diagnostic criteria for autism reflect a fundamental problem for virtually all psychiatric disorders: the absence of an independent and fully accepted diagnostic 'gold standard' (Volkmar and Cohen 1988*a*). Given this limitation, the diagnosis of autism is remarkable for the general agreement among clinicians, over decades and across nations. There *are* paradigmatic cases of autism about which all experienced clinicians would agree. This is reassuring in relation to the use of categorical diagnosis as

* Using clinical diagnosis as the standard of comparison, DSM-IIIR appears to have increased sensitivity and decreased specificity (Volkmar *et al.* 1987).

an anchor in research studies on particular mechanisms, such as those described in the present volume.

There is, however, one major diagnostic issue which deserves special note in relation to autism research. If one collected all the individuals who are diagnosed as having autism in one room, probably the most striking fact would not be their similarity, but how vastly different they are among themselves (Volkmar and Cohen 1988*b*). Included among individuals with autism are three-year-old children and senior citizens, people with profound mental handicap and university graduates, adults who barely have a word of expressive vocabulary (and almost undetectable receptive language) and adults who read encyclopaedias for recreation and speak with pedantic exactitude. Some individuals with autism are self-destructive, while others are overconscientious about their physical well-being. There are individuals with autism who memorize road maps and train schedules, and others who couldn't make sense of either.

To counter the potential research problems associated with such variability, the majority of studies discussed in this book focus on high-functioning individuals with autism. By 'high-functioning' we mean individuals with only moderate or mild intellectual impairment. In this respect, they constitute the upper 25 per cent of the population with autism (Rutter 1978). The selection of this subject group reflects a research strategy which enables us to identify autism-specific impairments independently of the effects of mental handicap in general. Ultimately, it will be important that research focusing on this group of individuals with autism should be extended to the full range of people with the condition.

A BRIEF HISTORY OF THEORY OF MIND

The literature on the development of a theory of mind has grown exponentially over the last ten years. Studies on children's developing understanding of the mental world arguably began with Piaget (1929). He claimed that children younger than seven years of age were unable to make the ontological distinction between the mental and physical realms. The discussion about the ontogeny of an understanding of minds was reopened in 1978 with the publication of Premack and Woodruff's seminal paper: 'Does the chimpanzee have a theory of mind?' Premack and Woodruff described a series of experiments that suggested to them that their famous chimp Sarah, who had knowledge of a symbol system, was able to predict and interpret a human's actions in terms of mental states such as intentions. They argued that Sarah's success indicated that she had a theory of mind. Commentaries on this paper, especially by Dennett, Pylyshyn, and Bennett, pointed out that it is not until one demonstrates an understanding of false belief (where

the mental state conflicts with reality) that one can unequivocally attribute a theory of mind to an individual, human or otherwise.

Within a few years, developmental psychologists began devising ingenious experiments to tap children's understanding of false belief, using the ideas suggested by Dennett and others. Wimmer and Perner (1983) published their important study of three- and four-year-olds' understanding of false belief, involving the now famous Maxi and the chocolate scenario. In this task, an object (a bar of chocolate) is unexpectedly moved whilst the main protagonist, Maxi, is out of the room. The child is then asked to predict where Maxi thinks the chocolate is, or where he will look for it. The main findings, which have been replicated many times, are that only older three-year-olds and over can pass this task. This study set in motion a flurry of research investigating young children's knowledge of false belief, other mental states, and related cognitive and linguistic achievements (see Astington *et al.* 1988; Butterworth *et al.* 1991; Frye and Moore 1991, and Whiten 1991, for recent collections of papers).

The first extension of this line of work to the study of autism which utilized Wimmer and Perner's (1983) false-belief test was carried out by Baron-Cohen *et al.* (1985). They used this test in order to ask the question 'Does the autistic child have a theory of mind?' This study, and subsequent replications, provided strong evidence that children with autism have a specific impairment in their understanding of false belief.* Given that it had been argued in the philosophy of mind and language (Dennett 1978; Grice 1975) that a theory of mind was *necessary* for social understanding and communication, it seemed plausible that a deficit in this area might account for at least two of the core symptoms in autism.

Earlier, Hobson (1981) had proposed the theory that children with autism have a primary problem in the development of a concept of other persons, specifically in coming to understand that people have minds. His approach was to investigate the understanding of expressions of emotion by children with autism, which grew out of his view that affective impairments (especially a relative lack of empathic responsiveness to others) could lead to impairments in conceptual development. The 1985 paper of Baron-Cohen *et al.* placed the emphasis on a primary *cognitive* deficit, and the debate about the primacy of affect or cognition in this domain continues to fuel new research ideas.

* Inevitably, many contributors in this book make reference to this early study. Whilst this may create some redundancy, each chapter uses this simply as a starting point for their own empirical and theoretical directions.

THE DEBATES

In this volume there are several fascinating theoretical debates that resurface with vigour between contributors, each time from a different angle. Here we mention the key issues, indicating in which chapters they are taken up:

The first set of debates focuses on underlying processes and developmental origins of a theory of mind: Does a theory of mind require a capacity for metarepresentation? If so, what is meant by metarepresentation (Leslie and Roth; Perner)? Does a theory of mind, or metarepresentation, arise *de novo* in the second to third year of life, or are there infancy precursors to either or both of these (Wellman)? If there are infancy precursors, what are these? Imitation (Meltzoff and Gopnik)? Joint-attention (Baron-Cohen; Mundy, Sigman, and Kasari)? Narrativity (Bruner and Feldman)? Affective sensitivity (Hobson)?

The second set of debates focuses on what consequences one would expect if a theory of mind was impaired: What would the effects be on language and communication (Tager-Flusberg; Loveland and Tunali) and on social development (Lord)? Is an inability to deceive a cardinal example of theory of mind failure (Sodian and Frith)? And what are the clinical implications of such deficits (Baron-Cohen and Howlin)?

The third set of debates focuses on alternative theories of the data from autism: Is task performance better understood in terms of executive control systems? If so, do deficits in the latter make better sense of the lack of imaginative play in autism, than the metarepresentation theory of autism (Harris)? Which symptoms of autism are successfully explained by the theory of mind hypothesis, and which cannot be (Klin and Volkmar)? Is the notion of a theory of mind a mistaken notion? If so, is it more appropriate to emphasize the development of self? (Samet); or social desire? (Mayes, Cohen, and Klin).

A final set of debates centre on what can be learnt about autism and the development of a theory of mind from the study of non-human primates (Gómez, Sarriá, and Tamarit; Whiten), or from philosophy of mind (Samet), or from a psychoanalytic perspective (Mayes *et al.*).

THE THEORY OF MIND HYPOTHESIS OF AUTISM: A PARADIGM CASE IN THE APPLICATION OF DEVELOPMENTAL PSYCHOPATHOLOGY

This volume exemplifies the significance of the field of developmental psychopathology, a field that uses theories and research on normal populations to advance our understanding of atypical children, and in turn acknowledges the influence that the study of atypical populations can have

on our understanding of normal development (Cicchetti 1984; Sroufe and Rutter 1984).

As our brief historical review shows, it is straightforward to recognize the influence of developmental psychology, and more broadly cognitive science, on the evolving theoretical and empirical work in the field of autism during the past ten years, and these are widely represented and acknowledged in numerous chapters in this volume. The other side of the equation—the contribution of the study of psychopathology to our understanding of normal development—can, with hindsight, also be discerned. Indeed, in having influenced theories of normal development it stands out as exceptional. Without the study of autism it is debatable whether the field would have been focusing on the significance of joint-attention in the development of a theory of mind, for example, or would have considered the modularity of a theory of mind in neuropsychological terms. Indeed, even the link between theory of mind and pretend play owes much to the associated deficits uncovered in autism. Finally, the burning questions of the primacy of affect and cognition in the development of a theory of mind can be seen as strongly influenced by the parallel debates about the primary impairment in autism.

ORGANIZATION OF THIS BOOK

In Part I, we begin with introductory chapters that review the development of a normal theory of mind, and the core social abnormalities characteristic of autism, since it is these that the theory of mind hypothesis set out to explain. Parts II and III then take up the central debates. Part II includes chapters that advance the theory of mind hypothesis of autism, all taking a cognitive approach, though there is by no means unanimous agreement among the contributors in this section on the nature of the deficit in autism. In Part III, a range of critical perspectives on the cognitive approach to the theory of mind hypothesis are presented. In Part IV, the debates are broadened still further to include philosophical, evolutionary, psychoanalytic, and developmental theories of autism and the theory of mind. In the final chapters, the implications of the theory of mind hypothesis of autism for clinical issues and for future research are considered.

We hope that this volume will provide an impetus for future work that will bring us closer to a more complete psychological understanding both of autism and of the normal development of a theory of mind.

REFERENCES

Astington, J., Harris, P., and Olson, D. (1988). *Developing theories of mind.* Cambridge University Press, New York.

Baron-Cohen, S., Leslie, A.M., and Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, **21**, 37–46.

Butterworth, G., Harris, P., Leslie, A., and Wellman, H. (1991). *Perspectives on the child's theory of mind.* Oxford University Press/British Psychological Society.

Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Development*, **55**, 1–7.

Dennett, D. (1978). *Brainstorms: philosophical essays on mind and psychology.* Harvester Press.

DSM-III-R (1987). *Diagnostic and statistical manual of mental disorders.* (3rd rev. edn). American Psychiatric Association, Washington DC.

Frye, D., and Moore, C. (1991). *Children's theories of mind.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Grice, H.P. (1975). Logic and conversation. In *Syntax and semantics: speech acts* (ed. R. Cole and J. Morgan). Academic Press, New York. (Original work published in 1967).

Hobson, R.P. (1981). The autistic child's concept of persons. In *Proceedings of the National Society for Children and Adults with Autism, International Annual Conference* (ed. D. Park). Washington.

ICD-9 (1987). *International Classification of Diseases* (9th edn). World Health Organization, Geneva.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child*, **2**, 217–50. Reprinted in Kanner, L. (1973). *Childhood psychosis: initial studies and new insights.* Wiley, New York.

Piaget, J. (1929). *The child's conception of the world.* Harcourt Brace, New York.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a 'theory of mind'? *Behavioral and Brain Sciences*, **4**, 515–26.

Rutter, M. (1978). Diagnosis and definition. In *Autism: a reappraisal of concepts and treatment* (ed. M. Rutter and E. Schopler). Plenum, New York.

Sroufe, A., and Rutter, M. (1984). The domain of developmental psychopathology. *Child Development*, **55**, 17–29.

Volkmar, F., and Cohen, D. (1988*a*). Diagnosis of pervasive developmental disorders. In *Advances in clinical child psychology* (ed. B. Lahey and A. Kazdin). Plenum, New York.

Volkmar, F., and Cohen, D. (1988*b*). Classification and diagnosis of childhood autism. In *Diagnosis and assessment in autism* (ed. E. Schopler and G. Mesibov). Plenum, New York.

Volkmar, F., Sparrow, S., Goudreau, D., Cicchetti, D., Paul, R., and Cohen, D. (1987). Social deficits in autism: an operational approach using the Vineland Adaptive Behavior Scales. *Journal of the American Academy of Child and Adolescent Psychiatry*, **26**, 156–61.

Whiten, A. (ed.) (1991). *Natural theories of mind.* Basil Blackwell, Oxford.

Wimmer, H., and Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, **13**, 103–28.