In B. Repacholi and V. Slaughter (eds), *Individual Differences in Theory of Mind: Implications for Typical and Atypical Development*. Macquarie Monographs in Cognitive Science. Hove, E. Sussex: Psychology Press, 2003.

Psychological Understanding and Social Skills

MARTIN DAVIES AND TONY STONE

Suppose we use the term 'theory of mind' in a neutral and inclusive way, so that having a theory of mind is simply having an ability to engage in our everyday folk psychological practices of attribution, interpretation, and prediction. This use of the term is non-committal as to how the ability is to be explained. Perhaps the explanation is to be given in terms of possession of a substantive theory about how the psychological world works; perhaps in terms of a capacity to identify with others in imagination and to simulate their mental lives; or perhaps in terms of a mixture of these things. The neutral and inclusive use of 'theory of mind' is also non-committal as to whether the basis of our folk psychological abilities resides in a special-purpose module of the mind.

Once we abstract away from questions about theory versus simulation and from questions about modularity, it can seem truistic that having a theory of mind – being able to engage in our everyday folk psychological practices or having *psychological understanding* – is fundamental to social functioning. Thus, for example, at the beginning of *The Child's Theory of Mind*, Henry Wellman says (1990, p. 1): 'Arriving at some sort of understanding of mind is an important accomplishment of childhood . . . An understanding of the mind is . . . fundamental to an understanding of the social world.' Indeed, it seems natural to suppose that, just as a naïve or everyday understanding of physics enables us to negotiate the physical world, so also a naïve or everyday understanding of psychology enables us to negotiate the mental, and in particular the social, world.[1] In a similar spirit, Simon Baron-Cohen says (1995, p. 30): 'Mindreading [the ability to imagine or represent states of mind that we or others might hold] is good for a number of important things, including social understanding, behavioral prediction, social interaction, and communication.'

The truistic-seeming idea that having a theory of mind is fundamental to social functioning might suggest that individual differences in theory of mind will go neatly in step with differences in social functioning. However, as the chapters in this book make clear, the relationship between psychological understanding and social skills is very much more complicated. One indication of the complexity of the relationship is that there are striking mismatches between psychological understanding as evidenced in experimental tasks – including, centrally, false belief tasks – and social skills as evidenced in daily life. In some individuals, social skills outrun psychological understanding. Other individuals have a high level of psychological understanding even while their social skills leave much to be desired. We begin with a mismatch of the first kind.

1. Mismatches between understanding and skills: Dunn's paradox

In line with the neutral and inclusive use of the term 'theory of mind', Janet Astington[2] says that having a theory of mind is a matter of attributing mental states to people so as to interpret their behaviour. The core of this interpretive practice is 'that we assume that people's actions are motivated by their desires in light of their beliefs' (ms. p. 3). So, having a theory of mind involves understanding the relations amongst the three core

[1] See Heider, 1958, p. 5 (quoted by Wellman, 1990, p. 2): 'In the same way one talks about a naïve physics . . . one can talk about a "naïve psychology" which gives us the principles we use to build up our picture of the social environment and which guides our reactions to it.'

[2] 'Sometimes necessary, never sufficient: False-belief understanding and social competence', in this volume.

notions – belief, desire and action – and the relations between these and intention, attention, perception, emotion, and the rest.

While it seems obvious that having this kind of understanding of the mental domain is a fundamental requirement for functioning well socially, there is a paradox here, pointed out by Judy Dunn.[3] Even very young children are social creatures who show a measure of understanding of the actions and feelings of other people. But children up to the age of five years are apt to fail at experimental tasks that are designed to assess their folk psychological understanding. Why, in these children, does social engagement and interaction outrun experimental performance?

In general, mismatches between psychological understanding and social skills invite a variety of explanatory strategies. Where real-life social skills outrun experimentally tested psychological understanding, at least three possibilities suggest themselves. First, it might be that the aspects of psychological understanding tested by the experimental tasks are indeed absent but that they are not centrally relevant to everyday social life. After all, many people who fluently negotiate their physical environment have a flawed understanding of how the physical world works. Second, it might be that the experimental tests reveal the absence of aspects of psychological understanding that normally play a central role in social life but that subjects are able to make up for this deficit by employing rules and work-arounds. Third, it might be that the experimental tasks test central aspects of psychological understanding that are actually present but that these aspects of understanding are manifested more readily in real-world performance than in laboratory performance. That is, it might be that the real-world situation imposes fewer demands, or offers more resources in support, than does the laboratory situation. So, in the case where skills manifested in the real world outrun understanding manifested in the laboratory (S>U), we can summarise the three explanatory strategies as follows: (S>U)1: irrelevance; (S>U)2: compensation by rules; (S>U)3: difference in demands or resources.

Now, the experimental tasks by which young children's theory of mind is assessed mainly test understanding of the fact that people's thoughts, actions, and feelings flow from beliefs that may not represent reality accurately. But it will not do to say, as the first strategy, (S>U)1: irrelevance, would suggest, that understanding of false beliefs simply has nothing to do with the social life of a three-year-old. Even if an understanding of desires and emotions is more centrally relevant to the pre-schooler, there is still the fact that, by three years of age, children engage in deceptive behaviour that seems to manifest an appreciation of the possibility of false belief. So why does this aspect of social interaction outrun performance on false belief tasks? The second strategy, (S>U)2: compensation by rules, is not adequate to explain this mismatch since these children do not seem to be just relying on learned deceptive routines (ms pp. 5–6).

Astington mentions at least two factors that may help to explain the mismatch in a way that is more in line with the third strategy, (S>U)3: difference in demands or resources. A child's developing understanding of false belief may be more readily manifested in real-world social behaviour than in an experimental task, first, because the child's social behaviour is very often 'supported by more competent others' (ms p. 28) and, second, because in a real-world social situation the child is likely to be emotionally involved and to 'have the motivation to act in an appropriate manner' (ibid.). In short, the

[3] Astington says, ms p. 3: 'At the [1988] Yale meeting mentioned earlier, Dunn pointed to a paradox that is at the heart of the issue. It is this: babies and toddlers are fundamentally social creatures, who are tuned in to other people, seemingly well aware of other people's behaviour and emotional reactions. Yet, until the end of the preschool years, children fail at experimental tasks that are designed to assess their understanding of another person's point of view.'

real-world situation offers supporting resources that are missing from the laboratory situation. (For the purposes of our simple taxonomy of explanatory strategies, we treat factors that provide motivation as 'resources'.)

In the title of her chapter, Astington says that false-belief understanding is 'never sufficient' for social competence. Factors other than false-belief understanding, including familiar settings, emotional significance, interaction with parents and others, interest, motivation, and emotional understanding, make essential contributions to social competence. We have just seen that the 'never sufficient' claim helps to solve Dunn's paradox because some of the factors that are required if understanding is actually to be manifested may be present in real-world situations but absent from the experimental situation.

But what the 'never sufficient' claim primarily suggests is a strategy for explaining a low level of social skills in subjects with a high degree of false-belief understanding (S<U). For it may be that social competence in a real-world situation imposes additional demands by comparison with an experimental test of false-belief understanding. For example, acting appropriately in a social situation may require understanding of both false belief and emotion. This is (S<U)3: difference in demands or resources. And we can rapidly see that there are at least two other strategies for explaining cases in which understanding manifested in the laboratory outruns skills manifested in the real world.

First, if experimentally tested psychological understanding is not centrally relevant to real-life social skills then the two may vary fairly independently, and so skills may be low while understanding is high. This is (S<U)1: irrelevance. Second, the understanding manifested in the laboratory might not be systematic or deep. Even in the absence of genuine psychological understanding, a subject might be able to work out the correct responses to experimental tasks such as false belief tasks, using rules and work-arounds. But laboriously reasoning one's way to solutions for problems posed in the laboratory does not make for fluent negotiation of the social domain. This is (S<U)2: compensation by rules.

2. Normal development, autism, and Asperger syndrome

The proposed solution to Dunn's paradox is that real-world social behaviour is likely to be supported and facilitated by resources that are absent from the laboratory. This still leaves it open that understanding of false beliefs may be only occasionally relevant to pre-school social life. But the complex relations between false-belief understanding and social behaviour in normally developing children are very much clarified by the research of Astington and her colleagues in Toronto. False-belief understanding is 'never sufficient' but it is 'sometimes necessary' for social competence.

2.1 Normal development

Astington reports that, while false-belief understanding is not related to mere frequency of pretend play, it is related to specific aspects of sophisticated pretend play: role assignment and joint planning. It is also related to certain pragmatic aspects of language use having to do with informativeness and to aspects of social competence that involve recognising and taking account of other people's mental states. (All this independently of age and language ability.) Furthermore, in the case of role assignment and joint planning in pretend play, there is evidence that the direction of causation is from development of false-belief understanding to these features of pretend play. So the aspects of psychological competence tested by false belief tasks are causally relevant to pre-school social life in the real world – and not only in respect of deceptive behaviour.

At the beginning of his chapter, Thomas Keenan[4] commends a framework for considering individual differences in the development of psychological understanding that was originally proposed by Karen Bartsch and David Estes (1996). According to their framework, there are three categories of research into individual differences in theory of mind development focusing on three kinds of differences:[5]

> Differences in the *antecedents* of theory of mind development that might facilitate or delay the acquisition of psychological understanding (e.g. differences in the number of siblings[6]).
>
> Differences in the developmental *consequences* of the specific timing – early or late – of acquisition of psychological understanding (e.g. differences in peer status).
>
> *Qualitative* differences in the child's theory of mind (e.g. a child might develop a 'theory of nasty minds'[7] if raised in a hostile environment).

In the present volume, many of the chapters report investigations of the consequences of individual differences in psychological understanding for the development of social skills – a category of research described by Bartsch and Estes, (1996, p. 287) as being still in its infancy.[8] Some chapters are concerned with the antecedents of psychological understanding, especially the role of language (see below, section 3). And qualitative differences in children's theories of mind, especially attributional biases such as aggressive children's bias towards attributing hostile intentions to an agent, are important for the chapters that discuss the relationship between psychological understanding and anti-social behaviour (see below, sections 4 and 5).

Keenan himself reports an investigation into the relationships between psychological understanding (assessed by a false belief task), peer acceptance or likeability, and social skills (as rated by a teacher) in four- and five-year-old children. The relationship between false-belief understanding and peer acceptance turned out to be complex, with a significant correlation (independent of language ability) when likeability was rated by girls but not when it was rated by the entire group or by boys. There were also significant correlations (again, independent of language ability) between false-belief understanding and social skills and between likeability (when rated by the entire group or by girls) and social skills.

Keenan also considers the prospects for a measure of psychological understanding that could detect continuing development beyond the age of five when a basic understanding of false beliefs is in place. He suggests that we borrow from research on social cognition the idea of 'empathic accuracy' or everyday mind reading as a measure of the accuracy of a subject's attributions of thoughts and feelings to other people.[9] While passing a test of false-belief understanding certainly requires accurate attribution of a thought to another person, tests of empathic accuracy typically draw on a rich set of resources going far beyond basic conceptual understanding of mental states. So, even

---

[4] 'Theory of mind and social development in the pre-school years and beyond', in this volume.

[5] The three kinds of differences mentioned by Bartsch and Estes are all inter-individual differences. Keenan adds a fourth category, namely, intra-individual differences. People may deploy their psychological understanding in different ways – for example, with different degrees of sophistication – in different contexts – for example, in relationships with different affective quality. See O'Connor and Hirsch, 1999.

[6] Perner, Ruffman and Leekam, 1994; Jenkins and Astington, 1996.

[7] Happé and Frith, 1996.

[8] Bartsch and Estes refer to Astington and Jenkins, 1995; Jenkins and Astington, 1996; Lalonde and Chandler, 1995.

[9] Ms pp. 22–3; see Ickes, 1993, 1997.

when basic understanding is securely in place, empathic accuracy continues to develop. Keenan suggests that, if existing empathic accuracy tasks could be adapted for use with younger children, then this measure would have a developmental span ranging from the pre-school years right through into adulthood.

Because we shall later be giving a rather specific sense to the word 'empathy', we shall use the more neutral term 'attributional accuracy' instead of 'empathic accuracy'. But the idea remains the same. Attributional accuracy requires complex inferences about the mental states of other people and these inferences may draw on observation, memory, and knowledge including, for example, knowledge about the specific individual in question and knowledge about particular kinds of relationships.[10] The hope is that, by using attributional accuracy tasks alongside more familiar tasks including false belief tasks, we could tease apart two rather different aspects of psychological understanding. On the one hand, there is the basic conceptual understanding of mental states of various types; not just beliefs but also desires, intentions, emotions, and so on. On the other hand, there is the further knowledge about the psychological world that is needed for accurate attributions of mental states to various people in various situations.

Tests of basic conceptual understanding of mental states, such as false belief tasks, often require subjects to attribute mental states to the protagonist in a story. But, even in these cases, there is a difference between having a basic understanding of mental states and actually attributing mental states to people. It is a difference between competence and performance, on one way of developing that distinction from Chomsky (1965). Basic understanding of mental states – of belief, for example – is a kind of conceptual competence. Actually attributing a mental state – attributing to Sally a belief about the location of her marble, for example – is performance. Competence is never sufficient for performance, and there are many factors that could lead to an incorrect attribution even when basic conceptual understanding is in place. For example, actually giving the correct answer in a false belief task may require inhibiting a pre-potent response to indicate the position where the marble really is.

What is more important for Keenan's proposal, however, is that accurate attribution of mental states often requires, not only basic conceptual understanding of beliefs, desires, intentions, emotions, and the rest, but also substantive empirical knowledge about mental states.[11] So, even when basic conceptual understanding is in place, and even when factors like pre-potent responses are not at issue, still a flawed view about how the psychological world works may lead to incorrect attributions of mental states. The flawed view might not be articulated. It might just be an attributional bias; and Keenan makes the helpful suggestion that such biases can be conceived as implicit theories about people and their psychological properties (ms p. 26). Someone who has a bias towards attributing hostile intentions to other people does not have an inability to represent

[10] Because of the rich variety of factors that are implicated in performance of these inferences – including, not only knowledge about individuals and kinds of relationships, but also motivation, for example – we would expect attributional accuracy to be subject to intra-individual differences.

[11] This is analogous to the situation in the study of language. Competence is linguistic knowledge; performance is the use of that knowledge. But linguistic performance, such as utterance interpretation, also requires knowledge that goes beyond knowledge of language. Both linguistic knowledge (competence) and real-world knowledge are needed for utterance interpretation.

As Kim Sterelny points out to us, it is not straightforward to apply the competence-performance distinction in the context of a simulation-theory account of our folk psychological practices. Indeed, it is a substantive question whether simulation theory can furnish a substantive account of our basic conceptual understanding of mental states. The authors in the present volume largely ignore simulation theory; but see section 9.3 below.

intentions or an impaired conceptual understanding of what intentions are. But such a person is likely to have an impaired ability to attribute intentions accurately.[12]

We can connect the issues in the last couple of paragraphs with the multiple uses of the term 'theory of mind'. Some theory of mind tasks assess psychological understanding in the sense that they assess basic conceptual understanding of one or another type of mental state. Other theory of mind tasks assess psychological understanding by probing the subject's knowledge about how the psychological world works. Either having a basic understanding of mental states or knowing how the psychological world works could be described as having a theory of mind. Either lacking a basic understanding of some or all mental states or having false empirical views about how the psychological world works could be described as having an impaired theory of mind. But it is important to recognise that, in these cases, the term 'theory of mind' would be used in two different ways.

Furthermore, the neutral and inclusive use of the term 'theory of mind' with which we began is different from both these uses. Having an ability to engage in our everyday folk psychological practices of attribution, interpretation, and prediction – especially if this is conceived as an ability to offer accurate attributions, interpretations, and predictions – requires both having a basic conceptual understanding of mental states and having some knowledge about how the psychological world works.[13]

2.2 Autism

From Astington's research, we have seen that, in normally developing children, the aspects of psychological understanding tested by false belief tasks are related to social competence (at least where this involves appreciation of other people's mental states), to pragmatic aspects of language use (having to do with informativeness), and to aspects of sophisticated pretend play (role assignment and joint planning). It is particularly interesting to compare all this with the case of children with autism since they typically have a triad of impairments: in social skills, in communication, and in pretend play.

As Helen Tager-Flusberg explains,[14] a typical pattern in research on the 'theory of mind hypothesis of autism' is that theory of mind as assessed by false belief tasks is found to be related to some measure of social skills, but the relationship disappears once age and language ability are factored into the analysis. So, despite the evident explanatory power of the theory of mind hypothesis, there is still a genuine question whether a theory of mind deficit explains discourse deficits and impairments in everyday social functioning in people with autism.

One distinctive aspect of the research being carried out by Tager-Flusberg and her colleagues in Boston is that they assess psychological understanding with a test battery that is broader than just false belief tasks. It includes, at the lower end of the scale, tests of pretence and of predicting action based on desire and, at the upper end, tests of judging intentions given information about personality traits and of moral judgements. The range of this test battery corresponds, in normally developing children, to an age range from eighteen months to early adolescence. In contrast, first-order false belief tasks, such as

[12] Strictly speaking, being biased is one thing and being inaccurate is another. Someone who has a bias, in the sense of being far more likely than most subjects are to attribute hostile intentions to other people, may have a superior, rather than an impaired, ability to attribute intentions accurately. For a biased subject might be surrounded by people with hostile intentions.

[13] Or perhaps a way of arriving at such knowledge, such as mental simulation.

[14] 'Exploring the relationships between theory of mind and social-communicative functioning in children with autism', in this volume.

the Sally-Anne task or the Smarties box task, correspond to an age range from three to five years.[15]

Psychological understanding assessed in this way is related, in children with autism, to a measure of social competence, to a pragmatic aspect of language use (namely, staying on topic), and to severity of autism diagnosis.[16] (All this independently of language ability.) In respect of both social and communicative functioning, these results provide a striking parallel with Astington's results for normally developing children, even though the specific measures of social competence and of language use differ between the two research programmes.[17]

Tager-Flusberg's research improves on earlier work, and provides important support for the theory of mind hypothesis of autism. This is because the demonstrated relationships between psychological understanding, on the one hand, and social and communicative functioning and severity of diagnosis, on the other, remain even when factors such as age, IQ, and – most important – language ability, are factored into the analysis. She attributes this improvement to the wider developmental span of her theory of mind test battery. But perhaps it is worth observing that the test battery is broader than just first-order false belief tasks in two ways. On the one hand, it covers a greater developmental span – as it would, though to a lesser extent, if it included the false belief explanation task, at the lower end of the scale, and second-order and other more complex false belief tasks, at the upper end of the scale. On the other hand, it assesses understanding of a wider range of psychological phenomena, including pretence, the role of desire in action, character traits, and moral commitment.[18]

2.3 Asperger syndrome

The research reported by Cheryl Dissanayake and Kathleen Macintosh[19] sheds some further light on the social skills of children with autism and also of children diagnosed with Asperger's disorder. On a test battery of four first-order false belief tasks, the children with autism performed significantly less well than typically developing children and also significantly less well than the children with Asperger's disorder.[20] But, despite this difference in performance on false belief tasks, on various measures of social skills the two clinical groups did not differ from each other.[21] Nor, for the most part, were individual differences in false belief performance within the groups related to social skills.

Dissanayake and Macintosh consider various possible explanations for these negative findings. Some of these concern the relatively narrow range of tasks used to test

---

[15] Wellman, Cross and Watson, 2001.

[16] But, Tager-Flusberg suggests, psychological understanding is not related to all aspects of autism; in particular, not to 'restricted repetitive and stereotyped behaviors, interests or activities' (ms pp. 18–19).

[17] Tager-Flusberg's work does not assess the relationship between psychological understanding and pretend play, presumably because a pretending task is already included in the theory of mind test battery.

[18] Keenan's proposed measure of attributional accuracy would also have a wide developmental span and it would range over a variety of psychological phenomena. But it would be different from Tager-Flusberg's test battery in a significant respect. For it would test knowledge of how the psychological world works and not just basic conceptual understanding of mental states.

[19] 'A comparative study of mind reading and social functioning in children with Autistic Disorder and Asperger's Disorder', in this volume.

[20] The children with Asperger's Disorder did not perform differently from the typically developing children, but this may be the result of a ceiling effect in the latter group.

[21] '[T]he children with high functioning autism and Asperger's disorder were largely indistinguishable on parent report of their adaptive behaviour, and on both parent and teacher ratings of social skills', ms p. 31.

psychological understanding and the relatively broad range of measures used to assess social functioning. In the case of the mismatch found in the children diagnosed with Asperger's disorder, between experimentally tested psychological understanding and real life social skills, these explanations belong in the category (S<U)3: difference in demands or resources. The basic idea here is that many aspects of good social functioning impose demands that go beyond anything required for understanding of false beliefs.

But Dissanayake and Macintosh also consider a different kind of explanation, the so-called 'hacking' hypothesis (Bowler, 1992), which belongs in the category: (S<U)2: compensation by rules. In the case of the children with Asperger's disorder, the idea would be that their success on false belief tasks is not the result of real psychological understanding but rather of laborious compensatory strategies that do not make for fluent negotiation of the social domain. This kind of explanation seems to be broadly in line with the following description of individuals with Asperger syndrome offered by Uta Frith and Francesca Happé (1999, p. 7):[22]

> There is reason to believe that the understanding of mental states developed by these individuals is rather different from the effortless automatic ToM of the normal preschooler. First, they require much higher verbal ability to pass ToM tasks than do normal children, and do so at later stages (typically in adolescent, not preschool, years). Second, even as adults they are prone to making tell-tale slips in mental state attribution. . . . Third, their approach to social tasks has been said to resemble slow, conscious calculation. They appear to do better with written than spoken communication, where the fast to and fro of mental state appraisal is avoided.

Frith and Happé regard people with Asperger syndrome as belonging within the autistic spectrum, differing from other people with autism in being able to 'pass tests of mental state attribution' and in having 'higher social and communication abilities than those without ToM ability' (ibid., p. 6). In apparent contrast, Dissanayake and Macintosh's two clinical groups are contrasted as children with high functioning autistic disorder, on the one hand, and children with Asperger's disorder on the other. But Dissanayake and Macintosh note that the 'hacking' hypothesis could also be used to explain how some high functioning children with autism are able to pass false belief tests and they regard their results as supporting the claim that 'Asperger's disorder is on a continuum with Autistic disorder' (ms pp. 35, 39).

In fact, it is not easy to obtain a clear picture of Asperger syndrome. One diagnostic criterion is normal language development, but impairments in pragmatic aspects of language use are often present. People with Asperger syndrome are reckoned to have higher social abilities than do most people with autism, but still show impaired social understanding and abnormal, inappropriate, and gauche social interactions.[23] The combination of relatively high language abilities and abnormal social interactions is seen in the two measures of social functioning on which children with Asperger's disorder differed from children with autism in Dissanayake and Macintosh's study. In the school playground, they spent more time conversing with their peers and they initiated interactions with peers more frequently even though this resulted in no more time spent in ongoing interactions.

[22] See also Raffman, 1999.

[23] For a brief review, see Ellis and Gunter, 1999.

3. The role of language

In order to explore the contribution that psychological understanding makes to social skills we need to take into account that the development of both the understanding and the skills will be correlated with factors such as age. So experimenters need either to choose subjects matched for age but differing in psychological understanding or else to use statistical methods to determine whether psychological understanding makes a contribution to the prediction of social skills that is independent of age. What goes for age goes also for language ability, and Astington and Tager-Flusberg both stress the importance of controlling for language ability when investigating the relationship between psychological understanding and social skills. However, we are also interested in the role that language ability plays in the development of psychological understanding and of social skills.[24] To what extent is language ability a causal factor in the development of psychological understanding or social skills and to what extent is it a causal consequence?

If we view language ability from the perspective of Paul Grice (1989), then it may seem that language ability must rest on prior psychological understanding. For, according to Grice, understanding an utterance is a matter of discerning the intentions and beliefs of the speaker. So linguistic understanding would be impossible for someone utterly unable to attribute mental states to others. But, while it is correct that some degree of psychological understanding is required even to understand literal utterances as speech acts, it is not obviously correct that the kind of psychological understanding that is assessed by false belief tasks, for example, is required for the basic communicative use of language.[25] Thus, it remains open that language ability may be an important causal factor in the development of psychological understanding and Astington provides evidence that, in three-year-old children, language ability does indeed play a causal role in the development of false-belief understanding, rather than the other way around (ms p. 20).

In an illuminating discussion of the time lag between understanding of desires and understanding of beliefs, Paul Harris (1996) proposes a specific role for conversation in the development of belief understanding. The basic idea is that children are involved in planned actions and in conversation. Planning and acting together with others, and coming to understand other people as agents, puts goals or desires centre stage. But exchanging information through conversation involves understanding other people as epistemic subjects; so it puts beliefs and knowledge centre stage (ibid., p. 208):

> My central claim is that children's understanding of other people as epistemic subjects develops in the context of their increasing proficiency at conversation involving the deliberate exchange of such information. Hence, my explanation for the lag is that a critical precondition for understanding beliefs but not desires – participation in the exchange of information through conversation – is not attained by most children until the third year.

According to Harris's hypothesis, it is conversational abilities and not verbal abilities as such, pragmatic aspects of language use rather than knowledge of syntax or vocabulary size, that are the crucial causal factors in the development of an understanding of beliefs. If this is right, then it will be important to investigate which pragmatic aspects are causal factors in the development of belief understanding and which are causal consequences.[26]

---

[24] See the concluding section of Astington's paper in this volume.

[25] See Langdon, Davies and Coltheart, 2002; Sperber, 2000.

[26] See Peterson and Siegal, 2000, pp. 139–40: '[T]o the extent that a theory of mind is a necessary component of a skilled conversationalist's pragmatic understanding of an interlocutor's mind and

Harris's paper originated in a conference held in 1994, before the work of Candi Peterson and Michael Siegal with deaf children was widely known. He notes that, if his hypothesis is correct, then we should expect 'that children with limited or delayed exposure to conversation (e.g. deaf children) should show difficulties on tests of belief understanding'.[27] As we now know, this is just how it turned out.

In a review of eleven separate studies, Peterson and Siegal (2000) found that false-belief understanding in deaf children from hearing families is delayed in comparison with hearing children. It is also delayed in comparison with 'native signers' (that is, deaf children of signing deaf parents or those who have a native speaker of sign language in their immediate household). Indeed, deaf native signers seem to develop false belief understanding at the same age as children of normal hearing, while the performance of deaf children from hearing families is markedly similar to that of autistic children of similar mental age.

Peterson and Siegal propose that the explanation of this delay in developing false belief understanding is that (ibid., p. 132):

> until they enter a signing (or Total Communication) primary school, many profoundly deaf children have no readily available means of conversing with any of their hearing family members, especially about topics like mental states which have no obvious visual referent.

This explanation seems to be roughly in line with Harris's hypothesis. But the situation is really quite complex.

Harris's hypothesis is that it is participation in conversation *as such* – as the exchange of information – that is crucial for the development of an understanding of beliefs, and not for the development of an understanding of desires, goals, and plans, for example. He distinguishes this from the hypothesis that what is crucial is participation in conversations *about mental states* (Harris, 1996, p. 211). Conversation about mental states – about beliefs or thoughts, desires or plans, itches or tickles – typically requires the use of specific lexical items. And where these lexical items are propositional attitude verbs, such as 'believes', 'desires', 'hopes', and the like, their use typically involves a characteristic syntactic construction: 'x Vs that p' ('Sally believes that the marble is in the box'; 'Sally desires that Anne should come and play'; 'Sally hopes that she will find the marble soon').[28] So, the 'conversation about mental states' hypothesis is connected with hypotheses about specific lexical items and syntactic constructions, and it is concerned, not only with beliefs, but also with other propositional attitudes. In contrast, the 'conversation as such' hypothesis is not connected with hypotheses about lexical items or syntactic constructions, and it is concerned specifically with beliefs, and not with other propositional attitudes such as desires or hopes.

It seems clear that the explanation proposed by Peterson and Siegal is intended to be in line, not so much with Harris's 'conversation as such' hypothesis, as with the competing 'conversation about mental states' hypothesis. So important questions arise about the development, in deaf children, of an understanding of desires and other mental

---

intentions, impairments in pragmatic skill and mental state understanding are likely to be reciprocal and inextricably interconnected.'

[27] Harris, 1996, p. 220, n. 7; the observation is credited to Peter Carruthers.

[28] It is worth noting that desires and hopes, for example, are often expressed without using the complement construction 'Vs that p'. Thus: 'Sally wants Anne to come and play'; 'Sally hopes to find the marble soon'. On competence with the 'x Vs that p' construction as a predictor of false-belief understanding, see de Villiers, 2000. Note too the relevance to these issues of investigations of psychological understanding in children with specific language impairment.

states. It is to such questions that Peterson turns in her chapter.[29] The experiments that she reports confirm, once again, that deaf children from hearing families lag behind hearing children in their development of false belief understanding. The results also show that the deaf children with an average age between nine and ten years are less successful than hearing children with an average age between four and five years in tests of desire and emotion understanding.

These results seem to be broadly consistent with Peterson's proposal that 'the present group of deaf children may have lacked the opportunities often available to hearing preschoolers in hearing-speaking families for querying, commenting on, or justifying seemingly false or deviant beliefs, feelings and desires' (ms pp. 24–5). And the results do not sit so comfortably with Harris's hypothesis that participation in conversation makes a contribution specifically to the development of belief understanding. However, the situation remains somewhat unclear since the discrepancy between the deaf children and the hearing children is greater in the case of belief understanding than in the case of desire and emotion understanding. One possibility is that the developmental sequence is the same in deaf children as in hearing children – belief understanding lags behind desire understanding – and that the overall delay is to be explained in terms of the limited opportunity for participation in conversation about mental states. Another possibility is that the limited opportunity for conversation as such results in a specific delay in the development of belief understanding by comparison with desire understanding, and that the overall delay – including the delay in understanding desires different from one's own – is attributable to some other factor. Perhaps both possibilities will figure in a full explanation of these phenomena.

4. Understanding, empathy, and anti-social behaviour

We began from the truistic-seeming idea that having an ability to engage in our everyday folk psychological practices of attribution, interpretation, and prediction is fundamental to social functioning. But Dunn's paradox highlights the fact that social skills manifested in the real world may outrun experimentally tested psychological understanding; and Astington's claim, 'Sometimes necessary, never sufficient', leads us to expect mismatches in the opposite direction as well. Various strategies are available for explaining these mismatches; one is an appeal to differences in demands or resources. Thus, where real-world skills outrun understanding manifested in the laboratory, this may be because the real-world situation imposes fewer demands, or offers more resources in support, than does the laboratory situation. Equally, where experimentally tested understanding outruns skills manifested in the real world, this may be because social competence in a real-world situation imposes additional demands by comparison with an experimental test of psychological understanding. This is especially clear if the experimental test battery includes only false belief tasks. For appropriate behaviour in a social situation may surely require understanding, not only of beliefs and desires, but also of emotions.

One of the findings of Astington's research is that, while false-belief understanding is related to aspects of social competence, language use, and pretend play, it is not related to empathy, to popularity, or to aggression. So we should keep separate the ideas of belief understanding, on the one hand, and emotion understanding and empathy, on the other. But it is also important to distinguish between emotion understanding and at least two notions of empathy. Astington's measure of empathy involved asking children to

[29] 'The social face of theory-of-mind: The development of concepts of emotion, desire, visual perspective and false belief in deaf and hearing children', in this volume.

nominate classmates who are kind when others are sad.[30] But empathy in this sense, which is perhaps better called 'sympathy', is different from understanding of emotions. Someone who understands sadness and understands kindness might be kind, unkind, or indifferent towards another person who is sad. Furthermore, both understanding of emotions and kindness to sad people (sympathy) should be distinguished from the more common notion of empathy as feeling with, or identifying in imagination with, another person. First, while the imaginative process of empathy may enhance emotional understanding, understanding of emotions, particularly third-person understanding, does not require empathy. Second, while empathy may lead to sympathy, it is neither necessary nor sufficient for sympathy.

Sympathy, manifested by kindness towards people who are sad, is thus several steps removed from false-belief understanding. Belief understanding does not guarantee emotion understanding; emotion understanding does not guarantee empathy; and empathy does not guarantee sympathy.[31] This last point is well explained by Peter Goldie (2000, p. 215):

> [Imaginative processes such as empathy] are consistent with at least three kinds of response which do not involve the sort of ethical motivation that is involved in sympathy. First, they are consistent with indifference: you can imagine the other's suffering, yet simply disregard it . . . Secondly, they are consistent with a response which is the *opposite* of sympathetic, involving *rejoicing* in the other's suffering, or even, like the subtle and imaginative inquisitor, exploiting your sensitivity to the other's feelings to help you exacerbate his suffering. And thirdly, they are consistent with motivations and actions aimed at alleviation of one's own suffering, rather than the other's. For example, one might turn away at the sight of blood . . .

So, third-person understanding of emotions and even the kind of first-person (or 'from the inside') understanding that is connected with empathy can be deployed for good (e.g. sympathy) or ill (e.g. cruelty).

Several of the chapters in this volume look at the relationship between psychological understanding and anti-social behaviour, defined by James Blair[32] as 'any action that impinges on the rights and welfare of others' (ms p. 5). It includes the behaviour of children and adults who bully others, children and adults who are Machiavellian, and individuals who are classified as psychopathic. Bartsch and Estes's (1996) framework for considering individual differences in the development of psychological understanding is also helpful for thinking about the relationship between theory of mind development and anti-social behaviour. In particular, one question that has been prominent in the literature is whether anti-social behaviour is a *consequence* of delayed or impaired psychological understanding.

---

[30] Astington, this volume, ms p. 16.

[31] There is considerable variation in the use of the terms 'empathy' and 'sympathy'. Alvin Goldman says (1993/1995, p. 197): 'To empathize with someone, it its most frequent sense, is to sympathize or commiserate, which involves shared attitudes, sentiments, or emotions.' But, strictly speaking, having an emotional experience of the same sort as someone else is one thing, and commiserating with someone is another thing. As we use the terms here, empathy is an imaginative process that leads to the first thing; sympathy is the second thing. But it is the distinction, rather than the terminology, that is important.

[32] 'Did Cain fail to represent the thoughts of Abel before he killed him? The relationship between theory of mind and aggression', in this volume.

The thought here might be this. Psychological understanding is required for socially competent behaviour; bullying, Machiavellianism and psychopathy are incompetent social behaviours; so bullies, manipulators and psychopaths must be deficient in psychological understanding. But, as the distinctions between psychological understanding, empathy, and sympathy reveal, and as Jon Sutton[33] points out in his chapter, this is not really a promising line of thought. It confuses social competence or social skills, on the one hand, with social conformity or moral correctness, on the other; and it pathologises rather than explains psychological differences (ms pp. 26, 7). Sutton recommends – as do Betty Repacholi, Virginia Slaughter, and their colleagues[34] – that psychological understanding should be regarded as a collection of neutral social tools to be assessed by tests that are 'value free' (ms p. 8). With psychological understanding seen in this way, there can then be an investigation of the motivations and other factors that influence how the neutral tools of psychological understanding are put to use for good or ill in social interaction.

4.1 Bullying

This leaves the theory of mind abilities of bullies, Machiavellians and psychopaths as an open empirical question, and the consensus view of the authors in this volume is that those abilities are not impaired. In the case of bullies, for example, Sutton remarks that 'there is little empirical evidence to support the popular stereotype of a bully as physically powerful yet intellectually simple or backward' (ms p. 7).

If the psychological understanding of those engaging in anti-social behaviour is not deficient, might it actually be better than average? Perhaps a successful bully, manipulator or psychopath requires theory of mind abilities that are superior to those that result from the typical course of development. On this question, the evidence does not allow an unequivocal answer. Sutton reports a study in which bullies demonstrate better false belief understanding than other children – better, in particular, than children classified as victims – even when the tasks require some understanding of the role of displays of emotion in the production of beliefs. But he agrees that, overall, the literature does not reveal such a straightforward picture of the relationship between psychological understanding and bullying. Similarly, the studies of Machiavellianism reported by Repacholi, Slaughter, and their colleagues and the studies of psychopaths reported by Blair reveal normal, but not superior, theory of mind abilities.

If the theory of mind abilities of these people are normal, then where should we look to explain their anti-social behaviour? If psychological understanding is a collection of neutral social tools, then why does one person use the tools anti-socially and another not? Recalling again the Bartsch and Estes framework mentioned in Keenan's chapter, we might ask: Are there *antecedents* in a child's development of psychological understanding that might result in its differential use? Sutton asks, in particular, whether certain features of children's family environments might have a tendency to result in their using theory of mind abilities to bully. He notes that the families of children who are bullies are 'often characterised by a lack of cohesion and an imbalance of power between the parents' and that 'ringleader bullying appears to be associated with insecure attachment'. But the question that remains is whether these features of family relationships actually impair the development of psychological understanding or rather encourage a 'cold, manipulative' way of deploying that psychological understanding (ms pp. 15–16).

[33] 'ToM goes to school: Social cognition and social values in bullying', in this volume.
[34] 'Theory of mind, Machiavellianism, and social functioning in childhood', in this volume.

There are several possibilities to be teased apart here. If bullies show normal psychological understanding on test batteries that are dominated by false belief tasks, it remains possible that they have an impaired understanding of mental states other than beliefs. Perhaps they do not understand emotions. On the other hand, it might be that bullies have a normal third-person understanding of emotions but do not attribute the correct emotions to their victims because they are not good at recognising emotions from the facial and bodily expressions of other people.[35] But Sutton does not regard either of these possibilities as being generally the case: 'It is as if some children who bully understand the emotions their actions cause and go ahead not only despite that, but because of that' (ms p. 17).

Other options that need to be distinguished include the possibility that bullies lack empathy in the sense that they do not identify with, or share, the joy or sadness of other people and the possibility that they do not understand what Sutton calls the 'moral emotions', such as guilt, love, remorse, sympathy, and shame. Finally, there is the possibility that, whether or not they empathise with other people's joy or sadness, bullies do not experience, and do not show, sympathy or remorse when it would be appropriate to do so. In this last case, it may be that exercises in excusing oneself are important for sustaining the lack of guilt, remorse, or shame; and it is plausible that these exercises would draw on psychological understanding. We might, then, expect that lack of remorse, for example, would be related to some measure of psychological understanding, and Sutton briefly mentions some evidence that this is indeed so (Sutton, Reeves and Keogh, 2000).

4.2 Machiavellianism

The overall picture that is suggested by Sutton's chapter is that bullies deploy their psychological understanding in a way that is skilful but cynical, manipulative and morally unattractive. In short, bullies seem to be rather Machiavellian and bullies do indeed score higher than control children on a questionnaire designed to assess Machiavellianism in children as young as nine years of age (the Kiddie-Mach scale).[36]

As Doris McIlwain[37] explains in her chapter, Machiavellianism has several components: 'There is a cynical view of others, a willingness to manipulate and exploit others to the point of harm, and a distinctive "cool" affective style' (ms p. 2). The cynicism component amounts to a belief that other people are untrustworthy; the Machiavellian expects the worst of others and gets in first. In his own eyes, his actions are no worse than what other people would do to him, given the chance. The cool affective style has two aspects. First, the Machiavellian lacks empathy. Indeed, McIlwain suggests that the combination of normal or superior psychological understanding with impaired empathy lies at the heart of the Machiavellian personality style. The second aspect of the affective coolness is that the Machiavellian betrays little in the way of emotion or affect. Because of this second aspect, he is well placed to escape detection as he engages in manipulation and exploitation.

---

[35] Sutton, Smith and Swettenham, 1999, show that bullies can answer questions about emotional expressions on faces. But, as McIlwain, this volume, notes, Machiavellians are impaired in recognising emotional expressions.

[36] Sutton and Keogh, 2000. However, as Repacholi, Slaughter, and their colleagues point out (ms pp. 13–4), there are important differences between bullies and Machiavellians, both in the type of anti-social acts committed and in the underlying intentions and motivations.

[37] 'Bypassing empathy: Mapping a Machiavellian theory of mind and sneaky power', in this volume.

Successful manipulation and exploitation require an accurate view of the mind of the other person and here the Machiavellian's lack of empathy may present a potential problem. So the second aspect of the cool style is important in part because tactics such as deception and flattery can be used in an information-gathering project that compensates for the first aspect, that is, for the Machiavellian's inability to use empathy or imaginative identification to understand the mind of another person.

Machiavellianism is, then, characterised by cynicism, manipulativeness, and cool affect. But McIlwain also mentions another way of seeing Machiavellianism as having a tripartite nature: it involves 'beliefs, tactics and morality' (ms p. 19). The central belief is that other people are untrustworthy. This is a cynical belief, but it also attributes cynicism to other people. The Machiavellian's tactics include deception and flattery. And the Machiavellian's morality is that it is better to exploit than to be exploited. It is the morality of getting one's retaliation in early. But there is a little more that can be said about the moral aspect of Machiavellianism, for the personality style includes lack of 'empathic concern' (ms pp. 10–11).

We take this to mean, first, that Machiavellians do not show the moral emotions of sympathy, remorse, and so on and, second, that this is plausibly explained in terms of the lack of empathy or 'feeling with' the person who is being manipulated. McIlwain makes use of a distinction between hot and cold empathy here. Cold empathy is a matter of being able to work out the mental state, particularly the emotional state, of another person. It is a kind of third-person emotional understanding and Machiavellians are not impaired in this respect. Hot empathy is assessed by measures of 'personal distress to another's negative experience' and 'an affective response for the distressed other' (ms p. 10). In other words, hot empathy encompasses what we have called empathy and what we have called sympathy. It is also implicit in the use of the term 'hot empathy' that empathic personal distress is normally involved in bringing forth the affective response, that is, the moral emotion.[38] On McIlwain's account, it is in these connected respects of empathy and sympathy that Machiavellians are impaired.

Since Machiavellianism is a multi-faceted personality style, it is natural to ask why the facets go together. To a considerable extent, the combination of normal or superior psychological understanding with impaired empathy, stressed by McIlwain, does help to make sense of the co-occurrence of the beliefs, the tactics, and the morality – the cynicism, the manipulativeness, and the coolness of affective response. But, partly because empathy is not strictly speaking necessary for sympathy, it seems possible to imagine people with normal or superior psychological understanding and impaired empathy who are, nevertheless, not inclined to anti-social behaviour. What else might explain the use of psychological understanding for anti-social purposes?

Repacholi, Slaughter, and their colleagues investigated whether young Machiavellians have social-cognitive biases that would support the use of psychological understanding for anti-social purposes. They found that nine- to twelve-year-old children who scored highly on the Kiddie-Mach scale were more likely to attribute negative intent to story characters in ambiguous social situations, and more likely to predict that the situations would lead to negative outcomes, than children with low Mach scores. These biases might lead to the cynical Machiavellian belief that other people are untrustworthy and, indeed, cynical themselves and to the Machiavellian morality of getting one's

[38] Or, at least, empathic personal distress may, in one way or another, 'diminish the likelihood that we will harm or exploit that other'; see ms p. 11.

retaliation in early. There would remain the question of what antecedent factors might figure in the aetiology of these social-cognitive biases.[39]

Repacholi, Slaughter, and their colleagues also investigated whether the High Mach children showed any impairment in empathy.[40] They did find a negative correlation between Mach scores and empathy scores, but this was explained by the fact that female children scored higher on empathy and lower on Mach than male children. Once the effects of gender were taken into account, the negative correlation was no longer significant. Given the image of a Machiavellian as someone with normal or superior psychological understanding but impaired empathy, this is a somewhat surprising result. But Repacholi, Slaughter, and their colleagues suggest an interesting and important possible explanation. In respect of empathy proper, they suggest that Machiavellians might not be impaired, but might be able to regulate their emotional feelings, especially when these could interfere with their personal goals. In addition, empathic personal distress might be offset by the positive emotional feelings associated with a personal goal. In respect of the moral emotions, such as sympathy, guilt, or remorse, they suggest that these might be absent, not because of an impairment – not because of impaired sympathy resulting from impaired empathy – but because the cynical Machiavellian belief and its associated morality provide justifications for manipulative and exploitative actions, even actions that lead to harm.

If this explanation is correct then it remains the case that '[t]he Machiavellian presumably knows how their manipulative behaviour will impact another person's feelings, but this knowledge is not accompanied by any feelings of concern, sympathy or compassion'. It may also be true that '[w]ithout this emotional arousal [the moral emotion], antisocial behaviour is less likely to be inhibited' (ms p. 27). But the primary explanation of the absence of the moral emotion would not go via a presumed connection between empathy and sympathy. Instead, the absence of the moral emotion would be the result of an exercise in self-justification. The Machiavellian's psychological understanding would thus enter the picture twice over, in the project of manipulation and in the project of excusing. This might seem to suggest that Machiavellians need to have, not just typical, but superior, psychological understanding. But, in this study, there was no significant difference in psychological understanding between High Mach and Low Mach children.[41]

It is an open question whether Machiavellianism properly so-called can be present in children below the age for which the Kiddie-Mach scale is appropriate. In their chapter, Repacholi, Slaughter, and their colleagues report on pioneering research with children between the ages of four and six years, using a new Mach rating scale. Here, as with the older children, there was no relation between Mach scores and psychological understanding as assessed by false belief tasks. And, as with older children tested in other laboratories, Mach scores were found to be positively correlated with aggression scores and negatively correlated with pro-social behaviour scores. Finally, Mach scores were not

[39] As in the case of bullying, we might naturally look to features of children's family environments to help explain these biases. Betty Repacholi reports (personal communication) that the High Mach children in their study classified their parents as 'permissive-neglectful'. This parenting style is described as 'disengaged parents who are motivated to do whatever is necessary to minimize the costs in time and effort of interaction with the child' (Maccoby and Martin, 1983, p. 000).

[40] As measured by the Bryant Empathy Index; see Bryant, 1982.

[41] As Repacholi, Slaughter, and colleagues note, High Mach children were simply those in the original sample with Mach scores in the upper quartile. In the absence of normative data, it remains open whether these children were High Mach by any more absolute criterion. So it remains a possibility that genuinely High Mach children would show superior psychological understanding.

related to social preference scores (a measure of the extent to which a child is amongst other children's most liked, rather than least liked, classmates). But they were related to social impact scores (a measure of the extent to which a child is amongst other children's most liked *or* least liked classmates). Machiavellians are noticed, favourably or unfavourably, by their peers. Overall, these first results with the new Mach scale are consistent with the idea that young Machiavellians, or proto-Machiavellians, deploy average psychological understanding in a distinctive anti-social way.

5. A double dissociation: Autism and psychopathy

Several of the chapters in this volume report the use of correlational analyses to investigate the relationship between psychological understanding and social skills. But, as James Blair[42] notes at the beginning of his chapter, the existence of correlations may not tell us very much about functional architecture (ms p. 6).

It is a familiar point that correlation is one thing and causation is another. But, even if a correlation between measures of two cognitive abilities X and Y is the result of a causal relationship, there remain questions about the nature of this relationship. The direction of causation may be from X to Y or from Y to X; or it may be that both X and Y depend causally on some third factor Z. And, even when the direction of causation is settled, that still leaves us some distance from a conclusion about functional architecture. For example, if there is causation in the direction from X to Y, that still does not show that the cognitive system that underpins X is a component of the system that underpins Y. It might be, for example, that possession of ability X is crucial for the acquisition of ability Y but is not directly implicated in the exercise of ability Y. That is, X might be a distal, but not a proximal cause of Y. Equally, if X and Y both depend on some third factor, this factor might be crucial for the acquisition of both X and Y but directly implicated in the exercise of neither.[43]

5.1 Developmental cognitive neuropsychology

Because of the questions that can be raised about correlational analyses,[44] Blair adopts the approach of developmental cognitive neuropsychology instead. But, in cases where cognitive neuropsychology delivers only findings of associations between deficits, important questions remain. If abilities X (for example, psychological understanding) and Y (for example, some social skill) are impaired together then this may only reflect facts about neuroanatomy. But suppose that there is a functional explanation for the co-occurrence of deficits. Suppose, for example, that both deficits result from some third impairment to ability Z (for example, a linguistic ability or an executive ability). Still, just

---

[42] 'Did Cain fail to represent the thoughts of Abel before he killed him? The relationship between theory of mind and aggression', in this volume.

[43] See below, section 7, for the distinction between distal and proximal causes. Blair also mentions the possibility that a cognitive ability X might be implicated in performance of tasks that are used to assess cognitive ability Y even though the cognitive system that underpins X is not a component of the cognitive system that underpins Y. For example, performance on false belief tasks depends on executive ('inhibitory') systems (ms p. 6).

[44] Blair explicitly mentions the problems faced by the correlational approach when it is adopted with normally developing individuals. But presumably the same point applies in the case of people with autism, for example. Thus, as we have noted earlier, both Astington – working with normally developing children – and Tager-Flusberg – working with people with autism – stress that it is important to control for language ability, as well as age and IQ, when investigating the relationship between psychological understanding and social skills. Of course, even with language ability controlled for, a correlation still does not settle questions about causation, let alone functional architecture.

as in the case of a correlation between measures of X and Y, this might not tell us much about the structure of the cognitive systems that underpin X and Y. In particular, the cognitive system that underpins ability Z might not be part of the cognitive system for ability X or the system for ability Y.

It is a very familiar point that neuropsychological findings of dissociations between impairments are apt to impose more constraints on functional architecture than do findings of associations. Blair's strategy is thus to use dissociations to investigate the relationship between psychological understanding and social skills. For example, if there are subjects in whom psychological understanding is absent or severely impaired while a particular social skill Y is intact then this counts against the hypothesis that psychological understanding is normally necessary for the possession of Y.[45]

However, when we discussed Dunn's paradox we noted three possible ways of explaining cases in which real life social skills outrun experimentally tested psychological understanding. It may be that the aspects of psychological understanding tested by the experimental tasks are absent but that they are not centrally relevant to Y. Alternatively, it may be that psychological understanding normally plays a central role in Y but that the subjects with impaired psychological understanding employ rules or work-arounds in order to manifest social skill Y. Or again, it may be that the real-world situations in which Y is manifested impose fewer demands, or offer more resources in support, than does the laboratory situation in which psychological understanding is tested. So, even given cases in which psychological understanding is impaired while social skill Y is intact, we need to rule out two kinds of explanation – (S>U)2: compensation by rules and (S>U)3: difference in demands or resources – before we conclude that psychological understanding is not normally implicated in the social skill.

In a neuropsychological context, the explanation in terms of demands or resources can be ruled out, or rendered highly implausible, by presenting cases of the reverse dissociation, that is, psychological understanding intact but social skill Y impaired.[46] So the evidence that Blair reviews takes the form of a double dissociation. On the one hand, people with autism are usually thought to have a severe impairment of psychological understanding while certain social functions are intact. On the other hand, psychopathic individuals are impaired in those social functions but have intact psychological understanding. Examining these two groups, people with autism and people with psychopathy, is an important way of investigating the relationship between theory of mind abilities and anti-social behaviour.

Two of the social functions on which Blair focuses are empathic responding and moral development.[47] People with autism orient towards adults who display signs of distress and they show autonomic responses to distress cues (sad faces) just as normally

[45] A dissociation of the form 'X impaired *at present* but Y intact at present' does not, in general, count against the hypothesis that *earlier* presence of X is normally necessary for the acquisition of Y. If X is a distal, rather than a proximal, cause of Y then it may be impaired after it has made its contribution to the acquisition of Y. For more on these issues, see below, section 7. But, in the case of autism, psychological understanding is not, and never has been, present. So, if a particular social skill Y is intact in people with autism then this counts against the hypothesis that psychological understanding is normally necessary either as a distal or as a proximal cause of Y.

[46] Also, in the case of the particular social functions under discussion, particularly autonomic responses to distress cues, the explanation in terms of compensation by rules is not plausible.

[47] Another is inhibition of anti-social behaviour. The apparent lack of instrumental anti-social behaviour in people with autism is consistent with the claim that there is a dissociation here: theory of mind impaired but inhibition of anti-social behaviour intact. However, the social interactions of people with autism are dysfunctional in any case (ms p. 7).

developing children do. Furthermore, the balance of evidence is that people with autism are able to recognise emotional expressions on faces. This is not yet to say that people with autism show normal empathy, or normal sympathy, as we have been using those terms. But the evidence certainly does not suggest a total lack of empathy, despite the severe impairment to psychological understanding.

In addition, people with autism understand the difference between moral and merely conventional transgressions. Across cultures, normally developing children make this distinction from around thirty nine months of age; and what is crucial for the distinction is the presence or absence of a victim. For example, children's classification of an unknown transgression (signified by a nonsense word) as moral or conventional depends on the consequences of the transgression. If it has a victim ('X has done dool and made Y cry'), then it is classified as moral; if it has no victim ('X has done dool and the teacher told him off') it is classified as conventional. In Blair's experiments, children with autism – even children who failed all false belief tasks – made this distinction. So, while this is not yet to say that people with autism show the moral emotion of sympathy to those who are experiencing distress, it does suggest that they 'generate appropriate aversion to acts that typically result in harm to others' (ms p. 12).

The reverse pattern of dissociation is found in individuals diagnosed with psychopathy: 'a disorder characterised in part by callousness, a diminished capacity for remorse, impulsivity, and poor behavioural control' (ms p. 12). These people have intact psychological understanding. But they are impaired in their autonomic responses specifically to distress cues, in recognising sad and fearful expressions on faces, and in making the distinction between moral and conventional transgressions. The cases of autism and psychopathy thus provide a double dissociation between psychological understanding and appropriate responses to distress; and it is impairments in the latter, rather than the former, that are associated with anti-social behaviour.

5.2 The violence inhibition mechanism

The connection between the impairments shown by psychopathic individuals and their anti-social behaviour is made by Blair's (1995) theory of a violence inhibition mechanism (VIM). In healthy subjects, this mechanism initiates a withdrawal in response to signs of distress such as a sad facial expression. So an observer – and, in particular, an aggressor – who sees a victim's distress cues will be predisposed to withdraw from the situation. In the course of normal development, a child 'will be negatively reinforced by the distress cues every time he engages in any aggressive activity' (ibid., p. 5). As Blair summarises the VIM theory (ms p. 14):

> [S]ad and fearful facial and vocal expressions act as punishing stimuli that, when experienced, reduce the probability that a healthy individual will engage in any action associated with the display of these expressions. In other words, the healthy individual is punished for engaging in antisocial activity by the distress of the victims.

Indeed, Blair suggests, even the thought of aggression may come to trigger the VIM, so that the child will become less likely to engage in violent actions. The VIM is also implicated in the development of the moral emotions and of the distinction between moral and conventional transgressions.

According to the VIM theory, absence of the violence inhibition mechanism has several consequences in addition to the absence of an aversive response to distress cues. Violent action is not inhibited, moral emotions such as sympathy, guilt and remorse are

lacking, and the distinction between moral and conventional transgressions is not drawn. Furthermore, if empathy is conceived simply as 'an emotional reaction to a representation of the distressed internal state of another', then the VIM normally plays a role in the development of empathy. Where the VIM is absent, empathy is missing as well. Thus, the constellation of impairments found in psychopathic individuals – impairments in autonomic responses to distress cues, in recognising sad and fearful expressions on faces, and in making the distinction between moral and conventional transgressions – is just what would be expected if, for physiological or social reasons, the VIM is absent. And, of course, if the VIM is absent then anti-social behaviour is liable to be forthcoming.

However, it is important not to oversimplify the relationship between the violence inhibition mechanism and actual behaviour. The presence of the VIM in healthy subjects does not guarantee withdrawal from a situation in which distress cues are present. Despite the predisposition to withdraw, an aggressor may be motivated to overrule the aversion generated by the VIM; or a bystander may approach and help the distressed victim. Equally, the absence of the VIM does not guarantee anti-social behaviour. The absence does not, of itself, provide any motivation for aggression and, in any case, other systems may serve to inhibit aggressive behaviour.[48] Nevertheless, it is clear in outline how the absence of the VIM could, given other factors, lead to the anti-social behaviour of psychopathic individuals.

5.3 Psychological understanding, the VIM, and anti-social behaviour

Having in mind Blair's account of the anti-social behaviour of psychopathic individuals, we can return briefly to bullying and Machiavellianism. From Sutton's chapter, recall two points. First, bullies do not have impaired psychological understanding. Second: 'It is as if some children who bully understand the emotions their actions cause and go ahead not only despite that, but because of that' (Sutton, this volume, ms p. 17). This does not suggest any impairment in recognising sad or fearful expressions on faces. On the other hand, Sutton notes that 'young bullies frequently ignore the submissive behaviour of their victims and carry on inflicting pain' (ibid.). This sounds like the absence of a predisposition to withdraw and to that extent it suggests an impaired or absent VIM. But, as Blair says, an aggressor with an intact VIM may overrule the aversion that it generates. So it would be interesting to know whether bullies are impaired in their autonomic responses to the distress of others and also whether they have any difficulties with the distinction between moral and conventional transgressions. For the time being, we should leave open the possibility that bullies engage in anti-social behaviour despite having an intact VIM.

In the case of Machiavellianism, McIlwain's discussion of an impairment to 'hot empathy' is broadly in line with the idea that Machiavellians have an impaired or absent VIM. In contrast, Repacholi, Slaughter, and their colleagues suggest an account that is more in line with the idea that Machiavellians overrule the deliverances of an intact VIM because of their own goals and because of the cynical Machiavellian belief and its associated morality. Once again, further information would be useful.

The Machiavellian belief that other people are untrustworthy and the associated bias towards attributing negative intent and expecting a negative outcome might be described

[48] Blair, 1995, p. 11: 'a lack of VIM need not result in the individual becoming a psychopath. . . . A lack of VIM does not of itself motivate an individual to commit aggressive acts. A lack of VIM just means that one source of the interruption of violent action is lost. . . . It is perhaps possible that the development of the psychopath may require deficits within executive functioning as well as within VIM; that both sources of behavioural inhibition must be impaired for the child to develop as a psychopath.'

as impaired psychological understanding. But it is not an impairment of the kind that is associated with autism; it is not an inability to represent and understand the nature of psychological states. Machiavellians are in general able to represent and understand other people's intentions, but there is a bias in their assessments of what these intentions are likely to be.

The overall picture seems to be that psychological understanding, particularly belief understanding, is necessary for the anti-social behaviour seen in bullying, Machiavellianism, and psychopathy. Also, anti-social behaviour does not seem to be the product of impaired emotion understanding, especially if this is conceived as third-person understanding of emotions. In addition, it is not clear that anti-social behaviour in bullies and Machiavellians is the product of impaired empathy or sympathy. It appears that a good understanding of beliefs, desires, intentions, emotions, and other mental states may be deployed in a way that leads to anti-social behaviour, especially in cases where an attributional bias has the result that the mental states attributed to others tend to have negative connotations. This may happen as the result of a child's experience of a developmental environment in which the attribution of negative mental states to others is often adaptive.[49]

A different route to anti-social behaviour does involve impaired empathy or sympathy. Blair's theory of the violence inhibition mechanism makes very good sense of the anti-social behaviour of psychopathic individuals. More work is required to assess its applicability to the cases of bullying and Machiavellianism. Also, more work of an inter-disciplinary kind is needed in order to relate the theory of the violence inhibition *mechanism* to more philosophical accounts of empathy, sympathy, and the moral emotions more generally.[50]

In this, we might begin with the role that Hume assigns to 'sympathy' in the production of moral approval and disapproval. For Hume, sympathy is a disposition to feel what others are feeling; so it is close to what we have been calling empathy. Its operation is the business of imagination rather than reason:[51]

---

[49] See Keenan, this volume, ms p. 10: 'For a child growing up in an abusive or hostile environment, such a modification to one's theory of mind might be quite adaptive in the short term, helping to protect the child from aggressive acts. For example, a hypersensitivity to aggressive intents may prove to be a useful adaptation to their social environment, helping an abused child to avoid acts of aggression and abuse.'

However, there are further complexities here because a bias towards attribution of hostile intentions is more closely associated with reactive aggression than with proactive aggression. For some discussion, see Sutton, this volume, ms pp. 8–9; Blair, this volume, ms pp. 18–19.

[50] See also Nichols, 2001, for an account that postulates a Concern Mechanism (CM) to do some of the work that is done, on Blair's theory, by the VIM. The normal operation of the VIM is that distress cues activate predispositions to withdraw. On the input side, *representations of distress* come to activate the VIM only by way of *classical conditioning* as a result of pairings of distress cues with representations of distress. On the output side, the operation of the VIM gives rise to *moral emotions*, such as sympathy, and thence to caring behaviour only by way of a process in which the arousal that is induced by the activation of the VIM is *interpreted* as a moral emotion (Blair, 1995, pp. 4–5). In contrast, the basic operation of the CM is that representations of distress give rise either to empathic distress or else directly to the moral emotion of sympathy. Thus, the basic operation of the CM, unlike the basic operation of the VIM, requires representation of a mental state. But, according to Nichols, 'basic altruistic motivation [provided either by a distinctive emotion of sympathy or else by empathic distress] requires only a minimal capacity for mindreading, the capacity to attribute negative affective or hedonic mental states like distress' (Nichols, 2001, pp. 450, 445).

[51] The following three quotations are from Hume, *Treatise*, Book 3, Part 3, Section 1. See also Stroud, 1977, chapter 9.

> When I see the *effects* of passion in the voice and gesture of any person, my mind immediately passes from these effects to their causes, and forms such a lively idea of the passion as is presently converted into the passion itself.

That is, when I perceive the vocal and bodily signs of an emotion, I form a representation of the emotion, and thence experience the same emotion myself.

> In like manner, when I perceive the *causes* of any emotion, my mind is conveyed to the effects, and is actuated with a like emotion.

That is, when I perceive a situation that would bring forth an emotion in another person, I experience the same emotion myself. Hume provides some vivid examples:

> Were I present at any of the more terrible operations of surgery, it is certain that, even before it begun, the preparation of the instruments, the laying of the bandages in order, the heating of the irons, with all the signs of anxiety and concern in the patient and assistants, would have a great effect upon my mind, and excite the strongest sentiments of pity and terror.

6. The theory of mind hypothesis of autism and schizophrenia

A number of authors in this volume discuss the theory of mind hypothesis of autism that is associated particularly with Simon Baron-Cohen and his colleagues. This hypothesis concerns the cognitive explanation of many of the symptoms of autism, especially of the characteristic impairments in imaginative play, social functioning and communication. According to the theory of mind hypothesis, this explanation can be provided in terms of a deficit in psychological understanding that results from damage to a domain-specific module or neurocognitive mechanism – or, perhaps better, damage to one or two components of a network of mechanisms. On Baron-Cohen's account, our normal ability to 'mindread' – that is, 'to imagine or represent states of mind that we or others might hold' (1995, p. 2) – is subserved by a network of four mechanisms, the intentionality detector, the eye-direction detector, the shared-attention mechanism, and the theory of mind module proper. He suggests that, in autism, there is an impairment to the shared-attention mechanism and a consequent dysfunction in the theory of mind module.

As we have already seen, Tager-Flusberg provides support for the theory of mind hypothesis of autism by showing that, in children with autism, psychological understanding is related to severity of the autism diagnosis, to a measure of social competence and to a pragmatic aspect of language use. On the other hand, Peterson shows that, in development, impaired psychological understanding is not specific to children with autism. In the case of deaf children, the impairment seems to be the result of their limited opportunities for participation in conversation. So this raises the possibility of an explanation of autism in which impaired psychological understanding is not the manifestation of a deficit in a cognitive module but rather a consequence of more basic linguistic and social difficulties.[52]

[52] See again Peterson, this volume; Peterson and Siegal, 2000; Garfield, Peterson and Perry, 2001. Any putative explanation of autism along these lines must take into account that, although deaf children's impaired psychological understanding persists even into late adolescence, it is usually described in terms of delay rather than deficit. Some high functioning children with autism are able to pass tests of false-belief understanding, as are people with Asperger syndrome; but there is some plausibility to the idea that this success reflects laborious compensatory strategies rather than systematic or deep psychological understanding (see again Dissanayake and Macintosh, this volume). So it would be interesting to know the results of a detailed and relatively demanding investigation of the psychological understanding of deaf adults who grew up in hearing families.

The theory of mind hypothesis of autism offers a unified cognitive explanation of many of the symptoms of autism and it grounds this cognitive explanation in neurobiology.[53] Christopher Frith has proposed a similar theory of mind hypothesis of schizophrenia (1992, p. 121):

> My proposal is that people with schizophrenia resemble people with autism in that they too have impairments in the mechanism that enables them to mentalise. . . . The autistic person has never known that other people have minds. The schizophrenic knows well that other people have minds, but has lost the ability to infer the contents of these minds.

So we can consider a unified theory of mind hypothesis of autism and schizophrenia according to which both disorders result from damage to a single neurocognitive mechanism, the theory of mind module. In the early-onset developmental disorder of autism this mechanism never develops properly; in the late-onset disorder of schizophrenia the mechanism malfunctions after achieving its mature state.

This bold hypothesis aims at what Philip Gerrans and Victoria McGeer[54] call 'horizontal integration' and 'vertical integration' (ms p. 445). Horizontal integration is the within-level unification provided by a single cognitive functional explanation of various symptoms. Vertical integration is the cross-level unification provided by an account of the neural basis of that explanatory cognitive function. In their chapter, Gerrans and McGeer critically assess and ultimately reject the hypothesis for the fundamental reason that 'there is no single [theory of mind] module responsible for successful social reasoning and behaviour in normal subjects' (ms p. 448). That is, there is no module or neurocognitive mechanism that, on the one hand, explains our ability to do the things that mindreading is supposed to be good for, such as 'social understanding, behavioral prediction, social interaction, and communication' (Baron-Cohen, 1995, p. 30) and, on the other hand, could be damaged, early or late, so as to give rise to most or many of the symptoms of autism or schizophrenia.

Their argument for rejecting the theory of mind hypothesis of both autism and schizophrenia comes in two stages. First, they argue that the case of autism does not support the idea of a genuine cognitive domain of psychological understanding or mindreading that is subserved by an innate module specific to that domain. This first stage of the argument leaves it open that the cognitive capacity for mindreading may be underpinned by a dedicated cognitive mechanism. The point of the first stage is to make it plausible that, even if the capacity for mindreading is subserved by a dedicated module, still it is developmentally constructed capacity rather than one that is 'achieved by a module whose cognitive architecture is genetically pre-specified' (ms p. 450). Second, they argue that the case of schizophrenia makes it plausible that the capacity to do the things that mindreading is supposed to be good for is not only developmentally constructed but also architecturally virtual rather than architecturally real. That is, 'it gives the appearance of being achieved by a dedicated mechanism [albeit a

---

Perhaps it is worth noting that, even if impaired psychological understanding in autism is a consequence of more basic linguistic and social difficulties, this does not undermine Blair's double dissociation argument discussed in the previous section.

[53] Baron-Cohen, 1995, chapter 6, suggests that the theory of mind module proper may be grounded in the orbito-frontal cortex (though see Blair, this volume, ms p. 16). But, on his account, psychological understanding depends on a network of four mechanisms and he suggests that the neural basis for this network involves the superior temporal sulcus and the amygdala, as well as the orbito-frontal cortex.

[54] 'Theory of mind in autism and schizophrenia: A case of over-optimistic reverse engineering', in this volume.

developmentally constructed one] but in fact the underlying architectural reality is of a number of interacting, possibly lower-order, quasi-independent, subsystems' (ms p. 453).

In the case of autism, a starting point for the argument is that there are symptoms of autism that do not have a satisfying explanation in terms of impaired psychological understanding. The symptoms that are problematic for the theory of mind hypothesis are not just the 'restricted repetitive and stereotyped behaviors, interests or activities' that Tager-Flusberg (this volume, ms pp. 18–19) found to be not significantly related to psychological understanding, but include also sensory-motor problems and abnormalities in perceptual processing. Gerrans and McGeer suggest that we make better sense of the co-occurrence of symptoms if we see the symptoms that do fit the theory of mind hypothesis as resulting from early sensory and perceptual disturbances. For these disturbances could have severe consequences for the social interactions and interpersonal engagement of a child with autism. The explanation of impaired psychological understanding in people with autism is thus structurally similar to the explanation offered by Peterson in the case of deaf children from hearing families.

The explanation is also structurally similar to the explanation of specific language impairment (SLI) that is offered by Paula Tallal (1988). According to this account, SLI is the result of a hearing deficit that makes it difficult for children to gather the information that would be needed for learning about rules of language such as the rules of inflectional morphology. This kind of account does not appeal to a genetic impairment in an innately specified language module; indeed, the postulated cause of SLI does not have any intrinsic connection with language. But the account still leaves it open that inflectional morphology is a real cognitive domain and that, in normal mature language users, the capacity to negotiate this domain is an architecturally real capacity grounded in the presence of a dedicated piece of cognitive machinery.

Thus, the state of play at the end of the first stage of the argument is as follows. There are symptoms of autism that are problematic for the theory of mind hypothesis. The co-occurrence of these symptoms with impaired psychological understanding can be explained if the theory of mind impairment is a causal consequence of the problematic symptoms. This explanation would still allow that psychological understanding is a genuine cognitive domain.

In the case of schizophrenia, Gerrans and McGeer point out that, alongside the theory of mind hypothesis of schizophrenia, Frith has more recently developed a rather different style of explanation for the experiences, and corresponding delusions, of alien control and thought insertion. Both styles of explanation appeal to a disorder of self-monitoring; in particular, an inability to monitor one's own intentions (Frith, 1992, pp. 114–5); but there is a crucial difference. According to the theory of mind hypothesis of schizophrenia, the breakdown in the representation of one's own intentions results from an impairment of the mechanism that supports the representation of mental states in general, the theory of mind module. But, in the alternative explanation, the key idea is that internal monitoring of self-initiated action is lost as a result of a breakdown in the component of the motor control system that compares feedback from a limb movement with a forward model or efference copy. This more recent explanation does not make any essential appeal to a general problem with representing mental states. The experience of alien control – which leads, in the delusion of alien control, to an incorrect attribution of intention – results from the breakdown of a comparator in the motor control system rather than from a malfunction in the theory of mind module.

If, as Gerrans and McGeer suggest, the second kind of explanation is to be preferred then in the case of schizophrenia, as in the case of autism, the theory of mind hypothesis

does not really meet the requirement of horizontal integration. It does not provide a single cognitive functional explanation of the symptoms.

Now, in the case of autism, Gerrans and McGeer suggested that the symptoms that do not receive a satisfying explanation in terms of impaired psychological understanding might play a role in the disturbed course of development that results in that very impairment. But, in the case of schizophrenia, the challenge of explaining the co-occurrence of symptoms must be met in a different way. For it is not plausible that impaired psychological understanding – which is indeed found in people with schizophrenia – is a causal product of factors like the breakdown of a comparator. Nor is it plausible that impaired psychological understanding accounts for all the symptoms that are left unexplained by the hypothesis about a failure of self-monitoring.[55]

Gerrans and McGeer propose that the most likely explanation of the co-occurrence of symptoms in schizophrenia will be neurobiological rather than cognitive. Strictly speaking, this would still leave it open that normal mindreading ability is an architecturally real capacity. But the crucial point is that, even if psychological understanding were underpinned by a cognitive module, we would not have a horizontally integrated cognitive explanation of schizophrenia.[56] So, once it is allowed that the unification of symptoms is to be achieved neurobiologically rather than cognitively, the reality or virtuality of the capacity for psychological understanding or mindreading becomes substantially irrelevant.

For the purposes of explaining the symptoms of schizophrenia, the apparent unity of the domain of psychological understanding might as well be regarded as virtual. We might as well suppose that mindreading is the manifestation of a collection of disparate cognitive systems. Thus, Gerrans and McGeer arrive at their main claim: '[Theory of mind] theorists go too far in proposing that high-level processing of social information constitutes a developmentally set or, indeed, [even an] architecturally real cognitive system whose malfunction is implicated in both autism and schizophrenia' (ms p. 455).

7. Distal causes and on-line processes

Correlation is one thing and causation is another. But, where there are correlations between aspects of psychological understanding, as evidenced in experimental tasks, and social skills, as evidenced in daily life, it is natural to ask whether there is a causal relationship and, if so, in what direction the causal explanation runs. Some researchers stress that both language and social experience play a role in the development of psychological understanding.[57] If this is right, then the development of some social skills is plausibly explanatorily prior to the development of some aspects of psychological

[55] The delusions of alien control and thought insertion are amongst the positive symptoms of schizophrenia. The analogue of the strategy followed in the case of autism would be to say that the impairment that explains the positive symptoms causes impaired psychological understanding and that this, in turn, explains the negative symptoms of schizophrenia, such as flattening of affect, poverty of speech, and social withdrawal. But Gerrans and McGeer say, 'the negative symptoms are prime candidates for a mechanical [rather than cognitive] explanation' (ms p. 480).

[56] That is, the explanation of the co-occurrence of symptoms would take the form of a vertically integrated account that unifies the symptoms neurobiologically. Speaking of the mechanical explanation of the disorder phenylketonuria, Gerrans and McGeer describe the strategy of vertical integration in the absence of horizontal integration (ms p. 10): 'we explain the neural malfunction and treat the resultant behaviour as the outcome of haphazard interference with the development of a number of *arbitrarily involved* cognitive functions'.

[57] See again Peterson, this volume; see also Garfield, Peterson and Perry, 2001. This kind of position is sometimes associated with Lev Vygotsky; see e.g. Astington, 1996.

understanding. On the other hand, some kinds of social interaction clearly require psychological understanding. It seems to us beyond dispute that negotiating the social world is, in part, negotiating a world in which people have beliefs, desires, and emotions that are different from our own and that successful negotiation of such a world may, on occasion, require an understanding of these mental phenomena.[58] If this is right, then the development of some aspects of psychological understanding is plausibly explanatorily prior to the development of some social skills.

If a cognitive ability Y depends causally on another cognitive ability X in the course of normal development, we can ask a further question. Does the development of X operate as a distal cause of the development of Y – perhaps contributing to the conditions in which Y can be effectively learned – or is X a proximal cause, directly implicated in the exercise of Y at a given time?[59] For an example where the answer to this question is intuitively clear, consider again Tallal's account of SLI. There is evidence that knowledge of inflectional morphology depends causally on specific aspects of hearing (having to do with the detection of acoustic changes over very short time intervals). It seems clear that these aspects of hearing operate as a distal cause of the development of knowledge of inflectional morphology by contributing to the conditions in which the rules of morphology can be learned. The crucial aspects of hearing are not directly implicated in the presence or use of morphological knowledge at a given time. Suppose, for example, that a subject gains the morphological knowledge and is able to use it and then suffers a hearing impairment of the crucial type. Intuitively, there is no reason to expect that the subject's performance on morphological tasks will immediately be impaired. Or suppose that a subject has impaired knowledge of morphology because of a hearing impairment. The morphological impairment is distally caused by, but not presently constituted by, the hearing impairment. If the subject's hearing were to be instantaneously improved there would, intuitively, be no reason to expect the subject immediately to perform at normal levels on morphological tasks.[60]

When we first considered the example of SLI, it was in the context of discussing the role of social interactions and interpersonal engagement in the aetiology of psychological understanding. But we can consider psychological understanding as a cause as well as an effect. Thus, consider some aspect of psychological understanding, such as false-belief understanding, and some social skill. We can ask whether the development of false-belief understanding is a distal cause of the development of the social skill, as crucial aspects of hearing are a distal cause of the development of morphological knowledge and as certain kinds of social interactions are a distal cause of the development of psychological understanding. Or is false-belief understanding a proximal cause, directly implicated in the existence of the social skill at a given time? Is false-belief understanding actually exercised on-line in the day-to-day manifestations of the social skill?

This kind of question also arises, for example, in the case of psychological understanding and pragmatic aspects of language use. There is considerable evidence of the co-occurrence, in people with autism, of impaired psychological understanding and impairments in pragmatic aspects of language use. We have already seen (Tager-

[58] See e.g. Currie and Sterelny, 2000, p. 145, for this point. It appears to be disputed by Garfield, Peterson and Perry, 2001, p. 525.

[59] For a general discussion of the importance of the distinction between distal and proximal causes for developmental cognitive neuropsychology, see Jackson and Coltheart, 2001, Chapter 2, 'Proximal and distal causes of individual differences in reading'.

[60] Improved hearing might, though, permit gradual linguistic improvement as a result of training. See Tallal et al., 1995.

Flusberg, this volume) that there is a relationship between psychological understanding and the ability to stay on topic. Also, people with autism show a poor understanding of metaphorical and ironical utterances (Happé, 1991, 1993, 1995) and a difficulty with distinguishing the different intentions of speakers who make jokes and speakers who lie (Leekam and Prior, 1994).[61] But evidence of an association, a correlation, or even a causal relation, between poor psychological understanding and poor pragmatics in autism may tell us rather little about the role that psychological understanding plays in the *on-line* processes that underpin pragmatic aspects of language use in normal adults. An argument for the claim that day-to-day exercises of communicative skills involve the on-line deployment of psychological understanding could be strengthened considerably by evidence about disorders of cognition that result from brain injury or a late-onset disorder such as schizophrenia, after the attainment of mature psychological understanding. Thus, for example, the claim is supported by findings that disorders that impair the ability to interpret other minds also impair pragmatic aspects of language use.[62]

In the case of people with schizophrenia, Robyn Langdon and her colleagues report that patients are impaired in false-belief understanding, and that this impairment cannot be accounted for in terms of more general problems that are evident in patients' performance on a picture-sequencing task. Patients also have difficulty identifying appropriate uses of irony and metaphor, though they do not differ significantly from control subjects in their ability to recognise appropriate literal uses of expressions (Langdon, Davies and Coltheart, 2002). But their problems with irony and metaphor do not reflect a single impairment in pragmatics. For Langdon's results also show that interpretation of irony and interpretation of metaphor involve distinct cognitive processes. Relatively sophisticated psychological understanding, of a kind that is disrupted in patients with schizophrenia, is implicated in the interpretation of irony. But only a more basic ability to attribute mental states, which is intact in patients with schizophrenia, is necessary for the interpretation of metaphor. The difficulty with metaphor interpretation that people with schizophrenia have is the result of something other than their impaired psychological understanding.

If we are to learn more about the contribution of psychological understanding to social skills then there is clearly much to be gained by the parallel investigation of psychological understanding – and, in particular, of individual differences in psychological understanding – in several populations. These include normally developing children, children with autism, deaf children and blind children, and also normal adults, people who have suffered brain injury, and people with schizophrenia.

8. Psychological understanding and delusional thinking in schizophrenia

As Langdon explains in her chapter,[63] people with autism and with schizophrenia both show impaired psychological understanding. But, although some of the negative symptoms of schizophrenia are similar to symptoms of autism, there are marked differences between the two disorders. The most striking of these is the presence of delusions in schizophrenia. Also, it may well be that impaired psychological understanding in autism is the result of a conceptual or representational deficit. That is, children with autism may be unable to represent representational mental states such as beliefs, intentions or pretendings. But, whether or not that is the correct account in the

[61] See also the test battery used by Tager-Flusberg, this volume.

[62] See, for example, see Happé, Brownell and Winner, 1999, for a review of research with patients who have sustained right hemisphere brain damage.

[63] 'Theory of mind and social dysfunction: Psychotic solipsism versus autistic asociality', in this volume.

case of autism, it certainly cannot be the correct account of poor psychological understanding in the case of schizophrenia. For schizophrenic delusions often involve the attribution of quite elaborate beliefs and intentions to other people – consider, in particular, persecutory delusions.

The occurrence of persecutory delusions in schizophrenia is enough to show that patients' poor psychological understanding is not the result of an inability to represent representational mental states. But it appears that flawed psychological understanding in people with schizophrenia, or in healthy adults with a high rating on schizotypal personality traits, may be related, not just to persecutory delusions, but to delusional thinking more generally. So the question arises whether there is some aspect of impaired psychological understanding in people with schizophrenia that could explain the tendency towards delusions.[64] Is there something that might explain the 'solipsism', or the private world, of a delusional patient with schizophrenia in contrast to the asocial 'aloofness' of a person with autism?

Langdon addresses this question by examining visual perspective-taking in high-schizotypal healthy adults and in patients with schizophrenia. Now, one of the tasks included in Peterson's test battery involved visual perspective-taking. The deaf children did not perform significantly differently from hearing children on this task even though they were impaired relative to hearing children on various measures of psychological understanding involving beliefs, desires, and emotions. This same pattern – understanding of visual perspective intact, but understanding of false belief impaired – is found in children with autism and in blind children.[65] But, in apparent contrast to these results for children with autism, and deaf or blind children, high-schizotypal healthy adults and patients with schizophrenia do show a problem with visual perspective-taking. And their performance on the visual perspective-taking task is related to their performance on a picture-sequencing task used to assess false-belief understanding.

The particular visual perspective-taking task that presents difficulties for Langdon's subjects involves judgements about how an array of coloured blocks would look when viewed from a different position.[66] Answering the question, 'Imagine moving to sit in the chair [90 degrees to the right]. Would the blocks look like this?', is equivalent to answering the question, 'Imagine turning the array [90 degrees to the left]. Would the blocks look like this?' But it is only the question that asks the subject to imagine adopting a different perspective that presents a problem. Given this striking result, it is natural to suggest that these subjects show poor psychological understanding because they find it difficult to adopt in imagination a different point of view and, in particular, the point of view of another person. That is, it is natural to suggest that the co-occurrence of the visual perspective-taking problem with poor psychological understanding counts in favour of a simulation-theory account of psychological understanding (Langdon and Coltheart, 2001).

But Langdon moves beyond this natural suggestion to a proposal that draws on ideas about egocentric and allocentric frames of reference.[67] When we represent a place in an

[64] We can ask this question about the role of impaired psychological understanding in the aetiology of delusions, such as the *delusion* of alien control, even if we accept that alien control *experiences* are caused by a breakdown of a component of the motor control system. See Davies et al., 2001.

[65] For discussion, see Peterson and Siegal, 2000, pp. 125–7; Garfield, Peterson and Perry, 2001, pp. 512–3.

[66] The question how the array would look is an *appearance* question, 'Would the blocks look like this [as presented on a computer monitor]?', rather than an *item* question, 'Would you see a yellow block in the front on your right?'.

[67] Eilan, McCarthy and Brewer (eds), 1993/1999; Evans, 1982, chapter 6, 'Demonstrative identification'.

egocentric frame of reference, this representation is closely linked with perception and action. I hear a sound as coming from a direction 45 degrees to the left of straight ahead and I turn in that direction. I see a drink as being about a metre away to the right and I move and reach towards it. But we also have a more detached way of representing places by using a cognitive map of a region. Quite independently of where I am or, indeed, of any facts about me, Macquarie University is very much further from the Sydney Opera House than the Harbour Bridge is, but in (roughly) the same direction.

It is a familiar thought in philosophy that this map-like style of representation is crucial for our objective conception of the world. One aspect of this objective conception is that I have a general, though highly fallible, ability to integrate egocentric and allocentric frames of reference so as to be able to make use, in my own actions, of the information provided by a cognitive map. Thus, the objective conception involves an appreciation that I am one *object* among others. I, and the objects in my egocentric space, have locations that can be represented on a cognitive map. Conversely, in virtue of my location, I have a subjective perspective or point of view on a region of objective space.[68]

Another aspect of the objective conception often stressed in philosophical discussions is that I appreciate that the course of my experience over time is determined jointly by what is happening at various places and by where I am located from time to time.[69] This means that facts about worldly occurrences cannot be wholly woven out of facts about my own experience and it allows for the possibility of existence unperceived. An event may occur unperceived by me because it happens at a place different from where I am located.

The conditions that are required for such an objective conception are famously explored by Peter Strawson in *Individuals*. For the most part, Strawson carries out his exploration by considering the conceptual scheme of a single subject. Under what conditions can a subject conceive of particular things as existing independently of himself and his experiences? Or, as Strawson puts it, under what conditions can the conditions of a 'non-solipsistic consciousness' be fulfilled? Here, a non-solipsistic consciousness is defined as: 'the consciousness of a being who has a use for the distinction between himself and his states on the one hand, and something not himself or a state of himself, of which he has experience, on the other' (Strawson, 1959, p. 69). But Strawson also notes that it is tempting to gloss the notion of something 'objective' as something 'public', so that an objective conception comes to involve 'the ideas of other enjoyers of experience and of shared surroundings' (ibid., p. 68). He briefly indicates a line of thought that might support the idea that objectivity requires publicity though he does not commit himself on the issue.

Here, we do not need to decide on the merits of the somewhat controversial philosophical theory that an objective conception absolutely requires a plurality of subjects. What is important is just to recognise that, in my ordinary thinking about the world, I appreciate that I am not only one object among others but also one *subject* among others. I am one of many subjects, each with a point of view onto a common world that exists independently of all of us. This is the idea that Langdon appeals to in order to explain the striking finding that psychotic and psychosis-prone adults have a selective impairment in visual perspective-taking when this is tested by questions of the form: 'Imagine moving to sit in the chair [90 degrees to the right]. Would the blocks look like this?' Her proposal is that the psychotic or psychosis-prone subject has a flawed

[68] See Evans, 1982, p. 163, on being able 'to impose the objective way of thinking upon egocentric space'.
[69] See Evans, 1980.

appreciation of the fact that he or she is one subject among others and, similarly, that his or her present subjective point of view is one among many.

This proposal also promises to contribute to an account of delusional thinking in schizophrenia. The appreciation that I am one subject among others provides me with a way of understanding disagreement and a way of 'reconciling discordant information', as Langdon puts it (ms p. 34). So a breakdown in this appreciation could leave schizophrenic patients less able to adopt a critical stance towards their own view of the world, less able to engage with others in a joint assessment of the real truth of the matter,[70] and so more prone to delusional thinking.

9. Twelve questions about psychological understanding and social skills

Many issues have been raised in this volume, but there is no single issue that dominates over all others. There are many questions to which we would like to know the answers. But no question seems to have a unique claim to be pivotal for future research. In this final section, we shall briefly review and connect some of the issues that seem to us to be important, deep, and fascinating. Along the way, we shall pose a dozen questions.

9.1 Correlation and causation

Many of the chapters describe correlational studies that explore the relationships between aspects of psychological understanding and communicative and social skills. We have half a dozen questions about these studies. In general, correlational studies do not settle questions about causation. But, in some cases, the correlational studies presented in this volume do include evidence for causal relationships – for example, Astington presents evidence from a longitudinal study to support the claim that the direction of causation is from development of false-belief understanding to certain features of pretend play. So our first question is this:

- Do correlational studies provide any evidence of a causal relationship between false-belief understanding and pragmatic aspects of language use or social competence?

There might, of course, be a causal relationship even though correlational studies provide no evidence of it. If there is in fact a causal relationship in the direction from false-belief understanding to some aspects of social competence, then we follow up with this question:

- Does the development of false-belief understanding operate as a distal cause of these social skills or is it a proximal cause, directly implicated in the exercise of the skills at a given time?

Some earlier studies have failed to find even correlations between psychological understanding and pragmatic aspects of language use or social skills, once language ability is taken into account. Tager-Flusberg assesses psychological understanding in children with autism using a test battery that is broader than just first-order false belief tasks in two ways. It covers a greater developmental span and it assesses understanding of a wider range of psychological phenomena. We have a question about this test battery:

- Is one or the other or are both of these dimensions of broadening crucial for Tager-Flusberg's findings of correlations with a pragmatic aspect of language use (namely,

[70] For several discussions of the role of this kind of breakdown in schizophrenic delusions, see the special issue of *Philosophy, Psychiatry, and Psychology*, volume 8, number 2/3, June/September 2001, 'On understanding and explaining schizophrenia', edited by Christoph Hoerl.

staying on topic) and a measure of social competence (Vineland socialisation score)?[71]

Keenan suggests that attributional accuracy tasks could provide another developmentally extended measure of psychological understanding. But accurate attributions of mental states often require much more than just basic conceptual understanding; they may require substantial empirical knowledge about how the psychological world works. If people with autism have an impaired ability even to represent mental states such as beliefs and desires, then it may be that their performance on attributional accuracy tasks will vary less than their performance on tests of basic understanding because of a floor effect. So we ask:

- If we were to use attributional accuracy tasks as a measure of psychological understanding in people with autism, should we expect to find correlations with social competence and pragmatic aspects of language use?

Dissanayake and Macintosh find that, in children diagnosed with Asperger's disorder, experimentally tested psychological understanding outruns real-life social skills. They consider more than one kind of explanation for this finding. The mismatch might be the result of the relatively narrow range of tasks used to assess psychological understanding or it might indicate that these children succeed on false belief tasks without real psychological understanding. So we ask, fifth:

- If psychological understanding in people with Asperger syndrome were to be assessed using Tager-Flusberg's broader test battery, or using a test of attributional accuracy as proposed by Keenan, how would it then be related to social skills?

False-belief understanding in deaf children from hearing families is delayed by comparison with both hearing children and 'native signers'. This finding is consistent with Harris's suggestion that participation in conversation as such is crucial for the development of belief understanding. But it is also consistent with Peterson's rather different proposal that participation in conversation about mental states is crucial for the development of an understanding of beliefs, desires and feelings. We might be able to resolve this disagreement in favour of either Peterson or Harris by using information about the time course of understanding of mental states other than belief, and especially understanding of desire, in normally developing children, in children with autism, and in deaf children of hearing parents. But our question is more basic:

- Might it be that Peterson and Harris are both right – that conversation about mental states facilitates the development of understanding of all mental states and that participation in conversation also plays a special role in the development of belief understanding?

9.2 Anti-social behaviour

Although we started out from the truistic-seeming idea that psychological understanding is fundamental to social functioning, it turns out not to be a promising line of thought to suppose that anti-social behaviour is a consequence of impaired psychological understanding. Psychological understanding is best thought of as a collection of neutral tools that can be used for good or ill. Several of the chapters describe investigations of the additional factors that are at work in anti-social behaviour.

[71] With this question in mind, it is interesting to note that, even with one test in the battery that assesses understanding of desire, there is no relationship between psychological understanding as assessed by the test battery and the use of desire words. In contrast, there is a relationship with use of cognition words to refer to cognitive mental states.

Blair reports that psychopathic individuals have a specific impairment in their response to distress cues and in their recognition of sad and fearful emotional expressions on faces. Sutton reports that bullies score more highly than other children on false belief tests, even when the tasks require recognition and understanding of emotions. This suggests that the explanation of anti-social behaviour in bullies is different from the explanation of anti-social behaviour in psychopathic individuals. To confirm this, we ask:

- Does Sutton's data show that bullies do not have the specific impairment in recognition of sad and fearful emotional expressions that is characteristic of psychopathic individuals?

McIlwain provides a vivid account of Machiavellianism as 'a personality style characterized by an average or above average competence in theory of mind, but lacking in certain forms of empathy' (McIlwain, this volume, ms p. 3). Repacholi, Slaughter, and their colleagues report that children with high Kiddie-Mach scores did not show impaired empathy. So our eighth question is:

- Are there individual differences within the Machiavellian personality style with some Machiavellians impaired in empathy and others not?

Where there is no impairment to empathy, anti-social behaviour may still be forthcoming; and (as Blair stresses), even where empathy is impaired, this is not sufficient for anti-social behaviour. So, in either case, additional factors – such as cynical beliefs, attributional biases, or executive function deficits – must be involved. This seems to suggest that there may be a good deal of variation within the category of individuals who are psychopathic. So we have a further question about individual differences:

- Are there individual differences in psychological understanding and empathy amongst psychopathic individuals, and what might be the consequences of these differences for remediation?

9.3 Simulation theory

Over the last fifteen years or so, the simulation theory of psychological understanding has been seen as an alternative to both the theory theory and the modular theory.[72] But in the eleven chapters that we have reviewed, the simulation theory is scarcely mentioned.

According to the simulation theory, our everyday folk psychological practices of attribution, interpretation, and prediction centrally involve identification in imagination with the other person. In the eyes of many, though certainly not all, simulation theorists, the position draws strength from the plausibility of the idea that identification with another person in imagination is involved in first-person, or 'from the inside', emotional understanding and empathy. Mental simulation is supposed to be involved in all psychological understanding, but first-person emotional understanding and empathy are often regarded as especially favourable cases for the simulation theory. However, the chapters on anti-social behaviour tend to drive a wedge between psychological understanding, on the one hand, and empathy, on the other. So our tenth question is:

- Does the investigation of anti-social behaviour – particularly, the double dissociation between psychological understanding and empathy – pose a challenge for the simulation theory?

If the dissociation between psychological understanding and empathy does deprive the simulation theory of psychological understanding of the support that it was thought to gather from the case of empathy, this might motivate a hybrid theory. Indeed, the dissociation might provide a principle for formulating such a theory. The idea would be

[72] For some of the seminal papers in this debate, see Davies and Stone (eds), 1995.

to combine a theory theory of psychological understanding with a simulation theory of empathy and the moral emotions. The mechanistic underpinnings of psychological understanding might include a theory of mind module. The mechanistic underpinnings of empathy might include the violence inhibition mechanism.

9.4 Psychological understanding and schizophrenia

Gerrans and McGeer offer a critical assessment of 'a theory of cognitive function which aims to [meet the challenges of horizontal and vertical integration] for both autism and schizophrenia' – a theory that draws on the work of Baron-Cohen and Frith.[73] Frith introduces his theory of mind hypothesis of schizophrenia as being similar to the theory of mind hypothesis of autism. But here we need to recall that impaired attributional accuracy is not the same thing as impaired conceptual understanding of mental states. In the case of autism, the theory of mind hypothesis is that there is an impairment of the basic representation and conceptual understanding of mental states. But people with schizophrenia do not have a late-onset impairment of that same kind for they are able to represent and understand mental states. People with schizophrenia may have an impaired theory of mind in the sense that they have false views about how the psychological world works. But they do not have an impaired theory of mind in the same sense that people with autism have an impaired theory of mind.

In any case, it seems clear that the theory of mind impairment in schizophrenia – whatever exactly may be its nature – cannot provide a cognitive unification of the symptoms of the disorder. So, Gerrans and McGeer argue, if the symptoms are to be unified, this must be a neurobiological unification; and for those purposes it is irrelevant whether psychological understanding constitutes a genuine cognitive domain.[74] This certainly removes one possible motivation for thinking that psychological understanding is a genuine domain. But removal of a motivation for saying something is not yet a reason for saying the opposite. So the question whether psychological understanding has the status of a genuine, or a merely virtual, domain seems to remain open.

When we are considering schizophrenia, and particularly the theory of mind hypothesis of schizophrenia, we have to regard theory of mind, or psychological understanding, as encompassing more than just basic conceptual understanding of mental states. It includes also the substantive empirical knowledge about the psychological world that is needed for accurate attributions of mental states. In this inclusive sense, psychological understanding draws on the ability to construct certain kinds of representations, on knowledge about how the psychological world works, and on the ability to carry out inferences. It does not seem very likely that all this is underpinned by a single cognitive module. So, if a genuine cognitive domain is one that is subserved by a single dedicated module, then it is independently not very plausible that psychological understanding is a genuine domain.[75] Our question here does not concern the correctness of this claim about psychological understanding but rather its theoretical connection with the unification of the symptoms of schizophrenia:

- If it turns out that the symptoms of schizophrenia can be unified neurobiologically but not cognitively, what does this tell us about the architectural reality or virtuality of the capacity for psychological understanding?

[73] Ms p. 446; see Baron-Cohen, Leslie and U. Frith, 1985; Baron-Cohen, 1995; C. Frith, 1992.

[74] An evaluation of the significance of this claim about the possibility of unifying the symptoms of schizophrenia would need to take account of the fact that the symptoms do not always occur together, but dissociate.

[75] For the notion of a module as dedicated to a domain, that is, as domain-specific, see Coltheart, 1999.

Langdon finds a relationship, in patients with schizophrenia and high-schizotypal non-clinical adults, between false-belief understanding and a particular kind of visual perspective-taking. If our aim were just to unify these two performance deficits then it would be natural to suggest that, at least in these subjects, impaired false-belief understanding is the result of an impaired ability to take on in imagination the point of view of another person – that is, an impaired ability to engage in mental simulation. But Langdon's main concern is to understand the connection between performance deficits on theory of mind tasks, including false belief tasks, and delusional thinking. To this end, she suggests that poor false-belief understanding and poor visual perspective-taking co-occur in psychotic and psychosis-prone individuals because of a breakdown, in these individuals, of the ability to represent subjective experience as contingent on one among many subjective points of view. Such a breakdown of the appreciation that one is one subject among many may help to account for delusional thinking because it may leave a schizophrenic patient less able to engage with other people in order to reconcile conflicting views of reality.

These suggestions raise a host of interesting issues. But, in order to move towards our final question, we shall consider just one. There is an apparent contrast between Langdon's results and results with children with autism, deaf children and blind children. But, in these latter experiments, the visual perspective-taking task does not share the crucial features of Langdon's task, which asks a question about how an array would *look* if the *viewer* were to adopt a different position. Indeed, Langdon asks the subject to *imagine* adopting that different position. So, our final question links our end back to our beginning:

- How should we expect normally developing children, children with autism or Asperger syndrome, and deaf children to perform on Langdon's visual perspective-taking task; and how should we expect their performance to be related to their psychological understanding?

Experiments might confirm that children with autism, and other children with impaired psychological understanding, are able to perform the visual perspective-taking task. If so then, as Langdon argues, this may help to explain the most striking difference between schizophrenia and autism; namely, the presence in schizophrenia but absence in autism of delusions.

Other possible outcomes would teach us something important about psychological understanding, imagination, and point of view. And if the visual perspective-taking task were included in test batteries for people who engage in anti-social behaviour – bullies, Machiavellians, and psychopathic individuals – then we could also find out whether visual perspective-taking patterns with psychological understanding, with empathy, or independently of both.

The eleven chapters that we have reviewed both indicate and instantiate the extraordinary richness, fertility, and promise of contemporary research on individual differences in theory of mind. The implications of this research extend beyond the psychology of typical and atypical development, into other disciplines including philosophy and psychiatry. Perhaps our dozen questions may provide a not-wholly-disconnected sample and summary of important, deep, and fascinating issues that have been raised.

1. Do correlational studies provide any evidence of a causal relationship between false-belief understanding and pragmatic aspects of language use or social competence?

2. Does the development of false-belief understanding operate as a distal cause of these social skills or is it a proximal cause, directly implicated in the exercise of the skills at a given time?

3. Is one or the other or are both of these dimensions of broadening crucial for Tager-Flusberg's findings of correlations with a pragmatic aspect of language use (namely, staying on topic) and a measure of social competence (Vineland socialisation score)?

4. If we were to use attributional accuracy tasks as a measure of psychological understanding in people with autism, should we expect to find correlations with social competence and pragmatic aspects of language use?

5. If psychological understanding in people with Asperger syndrome were to be assessed using Tager-Flusberg's broader test battery, or using a test of attributional accuracy as proposed by Keenan, how would it then be related to social skills?

6. Might it be that Peterson and Harris are both right – that conversation about mental states facilitates the development of understanding of all mental states and that participation in conversation also plays a special role in the development of belief understanding?

7. Does Sutton's data show that bullies do not have the specific impairment in recognition of sad and fearful emotional expressions that is characteristic of psychopathic individuals?

8. Are there individual differences within the Machiavellian personality style with some Machiavellians impaired in empathy and others not?

9. Are there individual differences in psychological understanding and empathy amongst psychopathic individuals, and what might be the consequences of these differences for remediation?

10. Does the investigation of anti-social behaviour – particularly, the double dissociation between psychological understanding and empathy – pose a challenge for the simulation theory?

11. If it turns out that the symptoms of schizophrenia can be unified neurobiologically but not cognitively, what does this tell us about the architectural reality or virtuality of the capacity for psychological understanding?

12. How should we expect normally developing children, children with autism or Asperger syndrome, and deaf children to perform on Langdon's visual perspective-taking task; and how should we expect their performance to be related to their psychological understanding?

Table 1: Twelve questions

References

Astington, J.W. 1996: What is theoretical about the child's theory of mind?: A Vygotskian view of its development. In P. Carruthers and P.K. Smith (eds), *Theories of Theories of Mind* (Cambridge: Cambridge University Press), 184–99.

Astington, J.W. and Jenkins, J.M. 1995: Theory of mind development and social understanding. *Cognition and Emotion*, 9, 151–65.

Baron-Cohen, S. 1995: *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

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

Bartsch, K. and Estes, D. 1996: Individual differences in children's developing theory of mind and implications for metacognition. *Learning and Individual Differences*, 8, 281–304.

Blair, R.J.R. 1995: A cognitive developmental approach to morality: Investigating the psychopath. Cognition, 57, 1–29.

Bowler, D. 1992: 'Theory of mind' in Asperger's syndrome. Journal of Child Psychology and Psychiatry, 33, 877–93.

Bryant, B.K. 1982: An index of empathy for children and adolescents. *Child Development*, 53, 413–25.

Chomsky, N. 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Coltheart, M. 1999: Modularity and cognition. *Trends in Cognitive Sciences*, 3, 115–20.

Currie, G. and Sterelny, K. 2000: How to think about the modularity of mind-reading. *Philosophical Quarterly*, 50, 145–60.

Davies, M., Coltheart, M., Langdon, R. and Breen, N., Monothematic delusions: Towards a two-factor account. *Philosophy, Psychiatry and Psychology*, 8, 133–58.

Davies, M. and Stone, T. (eds) 1995: *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell Publishers.

de Villiers, J.G. 2000: Language and theory of mind: What are the developmental relationships? In S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (Second Edition). Oxford: Oxford University Press, 000–00.

Eilan, N., McCarthy, R. and Brewer, B. (eds), 1993/1999: *Spatial Representation: Problems in Philosophy and Psychology*. Oxford: Oxford University Press. (First published by Blackwell Publishers, 1993.)

Ellis, H.D. and Gunter, H.L. 1999: Asperger syndrome: A simple matter of white matter? *Trends in Cognitive Sciences*, 3, 192–200.

Evans, G. 1980: Things without the mind: A commentary upon chapter 2 of Strawson's *Individuals*. In Z. van Straaten (ed.), *Philosophical Subjects: Essays Presented to P.F. Strawson*. Oxford: Oxford University Press, 76–116. Reprinted in *Collected Papers*. Oxford: Oxford University Press, 1985, 249–90.

Evans, G. 1982: *The Varieties of Reference*. Oxford: Oxford University Press.

Frith, C. 1992: *The Cognitive Neuropsychology of Schizophrenia*. Hove, E. Sussex: Erlbaum (UK) Taylor and Francis.

Frith, U. and Happé, F. 1999: Theory of mind and self-consciousness: What is it like to be autistic? *Mind and Language*, 14, 1–22.

Garfield, J.L., Peterson, C.C. and Perry, T. 2001: Social cognition, language acquisition and the development of the theory of mind. *Mind and Language*, 16, 494–541.

Goldie, P. 2000: *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

Goldman, A.I. 1993/1995: Empathy, mind, and morals. *Proceedings and Addresses of the American Philosophical Association*, 67, 000–00. Reprinted in M. Davies and T. Stone (eds), *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell Publishers, 1995, 185–208.

Grice, H.P. 1989: *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Happé, F. and Frith, U. 1996: Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology*, 14, 385–98.

Happé, F., Brownell, H. and Winner, E. 1999: Acquired 'theory of mind' impairments following stroke. *Cognition*, 70, 211–40.

Happé, F.G.E. 1991: The autobiographical writings of three Asperger syndrome adults: Problems of interpretation and implications for theory. In U. Frith (ed.), *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press, 207–42.

Happé, F.G.E. 1993: Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, 48, 101–19.

Happé, F.G.E. 1995: Understanding minds and metaphors: Insights from the study of figurative language in autism. *Metaphor and Symbolic Activity*, 10, 275–95.

Harris, P.L. 1996: Desires, beliefs, and language. In P. Carruthers and P.K. Smith (eds), *Theories of Theories of Mind* (Cambridge: Cambridge University Press), 200–20.

Heider, F. 1958: *The Psychology of Interpersonal Relations*. London: Wiley.

Hume, D. 1958: *A Treatise of Human Nature*. Oxford: Oxford University Press.

Ickes, W. 1993: Empathic accuracy. *Journal of Personality*, 61, 587–610.

Ickes, W. (ed.) 1997: *Empathic Accuracy*. New York: Guilford Press.

Jackson, N. and Coltheart, M. 2001: *Routes to Reading Success and Failure*. Hove, E. Sussex: Psychology Press.

Jenkins, J.M. and Astington, J.W. 1996: Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, 32, 70–8.

Lalonde, C.E. and Chandler, M. 1995: False belief understanding goes to school: On the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion*, 9, 167–85.

Langdon, R. and Coltheart, M. 2001: Visual perspective-taking and schizotypy: Evidence for a *simulation*-based account of mentalising in normal adults. *Cognition*, 82, 1–26.

Langdon, R., Davies, M. and Coltheart, M. 2002: Understanding minds and understanding communicated meanings in schizophrenia. *Mind and Language*, 17, 68–104.

Leekam, S.R. and Prior, M. 1994: Can autistic children distinguish lies from jokes?: A second look at second-order belief attribution. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 35, 901–15.

Maccoby, E. E. and Martin, J. A. 1983: Socialization in the context of the family: Parent-child interaction. In P.H. Mussen and E.M. Hetherington (eds), *Handbook of Child Psychology Volume 4: Socialization, Personality, and Social Development* (Fourth Edition). New York: Wiley, 1–101.

Nichols, S. 2001: Mindreading and the cognitive architecture underlying altruistic motivation. *Mind and Language*, 16, 425–55.

O'Connor, T.G. and Hirsch, N. 1999: Intra-individual differences and relationship-specificity of mentalizing in early adolescence. *Social Development*, 8, 256–74.

Perner, J., Ruffman, T. and Leekam, S.R. 1994: Theory of mind is contagious: You catch it from your sibs. *Child Development*, 65, 1228–38.

Peterson, C.C. and Siegal, M. 2000: Insights into theory of mind from deafness and autism. *Mind and Language*, 15, 123–45.

Raffman, 1999: What autism may tell us about self-awareness: A commentary on Frith and Happé. *Mind and Language*, 14, 23–31.

Sperber, D. 2000: Metarepresentations in an evolutionary perspective. In D. Sperber (ed.), *Metarepresentations: A Multidisciplinary Perspective*. Oxford: Oxford University Press, 117–37.

Strawson, P F. 1959: *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen.

Stroud, B. 1977: *Hume*. London: Routledge and Kegan Paul.

Sutton, J. and Keogh, E. 2000: Social competition in school: Relationships with bullying, Machiavellianism and personality. *British Journal of Educational Psychology*, 70, 443–57.

Sutton, J., Reeves, M. and Keogh, E. 2000: Disruptive behaviour, avoidance of responsibility and theory of mind. *British Journal of Developmental Psychology*, 18, 1–11.

Sutton, J., Smith, P. K. and Swettenham, J. 1999: Social cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, 17, 435–50.

Tallal, P. 1988: Developmental language disorders. In J.F. Kavanagh and T.J. Truss Jr. (eds), *Learning Disabilities: Proceedings of the National Conference*. Parkton MD: York Press, 181–272.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W.M. and Merzenich, M.M. 1996: Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271, 81–4.

Wellman, H.M. 1990: *The Child's Theory of Mind*. Cambridge, MA: MIT Press.

Wellman, H.M., Cross, D. and Watson, J. 2001: Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–84.