

Social Cognition, Language Acquisition and The Development of the Theory of Mind

JAY L. GARFIELD, CANDIDA C. PETERSON AND TRICIA PERRY

Abstract: Theory of Mind (ToM) is the cognitive achievement that enables us to report our propositional attitudes, to attribute such attitudes to others, and to use such postulated or observed mental states in the prediction and explanation of behavior. Most normally developing children acquire ToM between the ages of 3 and 5 years, but serious delays beyond this chronological and mental age have been observed in children with autism, as well as in those with severe sensory impairments. We examine data from studies of ToM in normally developing children and those with deafness, blindness, autism and Williams syndrome, as well as data from lower primates, in a search for answers to key theoretical questions concerning the origins, nature and representation of knowledge about the mind. In answer to these, we offer a framework according to which ToM is jointly dependent upon language and social experience, and is produced by a conjunction of language acquisition with children's growing social understanding, acquired through conversation and interaction with others. We argue that adequate language and adequate social skills are jointly causally sufficient, and individually causally necessary, for producing ToM. Thus our account supports a social developmental theory of the genesis of human cognition, inspired by the work of Sellars and Vygotsky.

How are we to decide whether to take reason to be an essentially private thing that can, however, turn on a public display when it chooses to do so, or, like conversing, to be an essentially social skill, which can, however, be retained a while through periods of solitary confinement? (Annette Baier, *The Commons of the Mind*, p. 5).

This research was supported by an Australian Research Council Special Initiatives Collaborative Grant (1997–1999) supplemented by institutional grants from The University of Tasmania, Flinders University, The University of Adelaide, Macquarie University and the Australian National University (JLG), by Smith College (JLG and TP) and by the University of Queensland (CP). We are grateful to the following people for research assistance and for assistance in organising the conferences in which these ideas incubated, and/or for important contributions to the discussions in those conferences: Greg Currie, Martin Davies, Kim Sterelny, Ian Ravenscroft, Gerard O'Brien, Phillip Gerrans, Max Coltheart, Philip Petit, Sandra Kellett. We also thank James Peterson, Jill deVilliers and Peter de Villiers, Lynne Rudder Baker, Jake Bridge, Max Coltheart, Kathryn Pyne Addelson and Catherine Sophian for comments on earlier drafts and the Propositional Attitudes Task Force, especially Ernie Alleva, Nalini Bhushan, Joe O'Rourke and Harold Skulsky for sharp discussion. We also thank audiences at the Australasian Association of Philosophy 2000 conference and at the University of Hawai'i for valuable comments and two anonymous referees for *Mind and Language* for suggestions that have improved the clarity of the presentation.

Address for correspondence: Prof. C.C. Peterson, School of Psychology, University of Queensland, Brisbane, Queensland, Australia 4072.

Email: candi@psy.uq.edu.au

Mind & Language, Vol. 16 No. 5 November 2001, pp. 494–541.

© Blackwell Publishers Ltd. 2001, 108 Cowley Road, Oxford, OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

The squirrel does not infer by induction that it is going to need stores next winter as well (Wittgenstein, *On Certainty* § 287).

The problem of thought and language thus extends beyond the limits of natural science and becomes the focal problem of historical human psychology, i.e. of social psychology. Consequently, it must be posed in a different way (Vygotsky, *Thought and Language*, p. 51).

From the very first days of the child's development his activities acquire a meaning of their own in a system of social behaviour and, being directed towards a definite purpose, are refracted through the prism of the child's environment. The path from object to child and from child to object passes through another person. This complex human structure is the product of a developmental process deeply rooted in the links between individual and social history (Vygotsky, *Mind in Society*, p. 30).

1. Introduction

The question we really care about is this: What is the character and origin of our knowledge about our own and others' minds? We don't mean anything fancy or technical by the word 'knowledge' here, nor by the word 'mind'. We are concerned simply with the cognitive achievement—remarkably widespread and uniform in our species (though, notoriously and importantly not universal among us)—that enables us to report our propositional attitudes (such states as belief, desire, hope, fear, etc, henceforth PA's) and to attribute such attitudes to others, and, more importantly, to use such postulated or observed states in the prediction and explanation of behaviour. That we do observe, report, and speculate about such states, and that we use attributions of such states in our normal social intercourse, in planning and in cooperation is undeniable and uncontroversial. Everything else is controversial, and every plausible thesis regarding the explanation of this achievement has been asserted and has been denied at some time or other.

But given the current state of play in the many interlocked debates about the 'Theory of Mind' (ToM) as it has come to be called,¹ we cannot approach this question directly without answering several others first:

¹ We will use the term 'Theory of Mind' abbreviated 'ToM' as a broad umbrella term to denote whatever knowledge guides propositional attitude attribution and the explanation and prediction of behaviour by means of inner states and processes. In particular, our use of the term, unless otherwise indicated, is meant to be neutral between positions according to which ToM is an explicitly articulated scientific theory (Gopnik and Meltzoff, 1997; Gopnik, 1996), those according to which it is a highly implicit set of representational capacities (Karmiloff-Smith, 1992) and according to which it is not theory-like at all (Astington, 1996; Butterworth, 1994; Shatz, 1994) as well as between modular (Baron-Cohen, 1995; Baron-Cohen and Swettenham, 1996) and non-modular views (Gopnik, 1996; Harris, 1996) of the representation and use of that knowledge. We will address all of these debates below, but regard the term as sufficiently entrenched that it can be regarded as a primitive entry in the cognitive science lexicon.

1. What is the nature of that knowledge and just how is it represented? Is it a full-blown theory or is it instead more like an assemblage of skills? Are we asking about explicitly articulated lore or implicit cognitive capacities?
2. Is the knowledge in question modularised or is it highly interwoven with the rest of our knowledge and cognitive capacities? And if modularised, is it an acquired or an innate module? If acquired, how? And by the way, what does 'modularity' in the context of debates about ToM acquisition really mean?
3. Perhaps closest to our hearts, what is the relationship between ToM knowledge and the mastery of language and social skills? Is ToM, as many (Baron-Cohen, 1995; Boucher, 1996; Carruthers, 1996; Mitchell, 1994; Segal, 1996) have argued, a necessary condition of the development of social intelligence, cooperation and mentalistic language, and the causal condition of these cognitive developments? Or is it, as others (Astington, 1996; Garfield, 1998; Harris, 1996; Sellars, 1998; Shatz, 1994; Siegal and Peterson, 1994; de Villiers and de Villiers, 1999; and Smith, 1996) have argued, the consequence of the development of specific social and/or linguistic competencies?
4. Is the acquisition mechanism for ToM essentially social and experiential, or is it neurobiologically maturational, driven by an innate dynamic?

We think that by asking and answering these clusters of questions, we can answer the big question, and explain the character and acquisition of ToM. And we believe that despite the cacophony characterising the current debates, enough data are available to settle these questions. At least more or less. In this paper we will argue that ToM is acquired through the acquisition of social and linguistic competencies, and does not precede them as an autonomous body of knowledge. We will therefore argue that the acquisition of ToM is essentially social in character, and that the body of knowledge represented by ToM is inextricably bound up with broader knowledge about persons and their lives. Nonetheless, we will argue that there are innate mechanisms that subserve the processes that mediate the acquisition of ToM, and which do so quite specifically, and that elements of the deployment of ToM in on-line processing are modular in an important sense. This modularisation, we will argue, is best seen as an achievement, and as a characteristic of cognitive processes and not of knowledge or its representation. We will also suggest that ToM is best seen as a set of skills and dispositions, and not as a set of theses represented in some inner code, and so as theoretical in only a limited sense. On our view, then, the development of language and the development of a set of social skills are prior to, jointly causally sufficient, and individually causally necessary for the acquisition of ToM, in contradistinction both to strongly modular theories of the genesis of ToM and 'theory theory' accounts.

While language and the relevant social skills may be largely innately determined, and strongly modular, we argue that ToM does not constitute an innately determined module, and that its modularity is considerably weaker.

In developing this big picture we take ourselves to be responsible for accounting for the broad range of data concerning the development of and deficits in theory of mind and related attention and perceptual skills in normal, perceptually impaired, autistic and Williams syndrome children. But we also take ourselves to be responsible for locating the account of ToM and its development in the larger landscape of cognitive science and epistemology. The account we offer must make sense of our self-knowledge and our knowledge of other minds, and must harmonise with our more general accounts of the nature of early learning and social development. We will hence have one eye on relevant data and problem cases and another on epistemology and the meta-theoretic foundations of cognitive science.

To be sure, the questions we address about the origin and aetiology of ToM are important and of considerable theoretical interest in their own right (and may even have significant clinical implications). But they also have ramifications for larger theoretical concerns, and we shall be concerned to draw lessons for epistemology and the foundations of cognitive science.

Questions about epistemic access to the minds of others and to our own minds have been central to epistemology in the last century, and have gained momentum as epistemology has become more closely allied with empirical psychology. Those who are influenced by the work of Sellars (1956 reprinted in 1998) and Wittgenstein (1956) argue that our conceptions of, and access to, our own mental states and processes, and those of others, are acquired with, and mediated by, public language and a public discourse. These views, enormously influential over the second half of the Twentieth century, challenge the notion that a strong nativist, or a strongly modular, theory could possibly be correct, even as they also challenge the idea that introspection and the attribution of mental states to others is possible in the absence of theory of, or at least a sophisticated conceptual framework through which to understand, the mind. The ToM debate allows us to explore the degree to which these compelling approaches to epistemology and the philosophy of mind are supported or undermined by empirical psychological research and hence to get some valuable empirical leverage on some of the most important epistemological debates of the Twentieth century.

Contemporary debates about the structure of theory in cognitive science often address the nature of knowledge representation; the status of modularity and its relationship to on-line performance and maturation; and the relation between the respective contributions of innately determined processes and social processes in psychological growth and development. At stake in this debate are also questions about individualism in the ontology of mind and in the methodology of psychology. We see this specific discussion as a contribution to these larger investigations, as well and so as bearing on the biggest

questions concerning the nature of mind, knowledge, and the relation of the individual mind to its social context.

We begin by canvassing some of the principal arguments for the innateness of the mechanisms subserving the acquisition of ToM and for the modularity of these mechanisms. We will argue that many of the most prominent arguments in the literature overstate just how much must be innate, as well as the degree to which it must be modular. We will argue that the data regarding the distribution of ToM deficits and regarding the pattern of normal acquisition of ToM are best explained by appeal to general purpose innate language acquisition mechanisms and social intelligence modules, together interacting with a social environment in which the knowledge and capacities specific to ToM are developed. In the next four sections of the paper we review a range of empirical data from studies of the acquisition of ToM by normal, autistic and deaf children, arguing that the data support neither a strong nativist nor a strongly modularist interpretation. We consider the special problems posed by Williams syndrome, and argue that the developmental pattern displayed by children in this category is also best explained by our mixed innate general mechanisms/social learning model. We then argue directly for our model of the acquisition of the ToM and show that the very arguments advanced by strong nativists and modularists should have led them to our position in the first place. We conclude with a discussion of the relationship between the model we advance and the programs of Vygotsky, Gibson and Karmiloff-Smith, developing general morals for the understanding of cognition in a social context.

2. Arguments for Innate Mechanisms in the Acquisition of ToM and for the Modularity of ToM

Arguments for the innateness of ToM are often bound up with arguments for its modularity. And the arguments are often roughly the same. This is not surprising, of course, since many innate cognitive structures are modular (language acquisition mechanisms, visual scene parsing mechanisms, etc...) and since many (e.g. Fodor, 1983) take innateness to be a core characteristic of modularity. (See also discussions in Garfield, 1987.) But these two properties should be distinguished. As Garfield (1994) and Karmiloff-Smith (1992) argue, modules can often be 'assembled'. That is, they may, instead of being innately specified, or the result of innately specified developmental processes, be acquired through learning or social/environmental interaction. Moreover, of course, there are cognitive processes that are to some degree innately specified but non-modular (the capacity for abductive or inductive learning, for instance). We therefore separate these issues analytically, even though some of the arguments will overlap. We begin with arguments for the innateness of ToM, recalling that 'innateness' is used in our somewhat broader technical sense.

The first argument for the innateness of ToM is that its acquisition by normal children is characterised by a relatively fixed developmental sequence: Many researchers find that normal children monitor eye gaze and adjust their gaze to fix another's attention on an object by twelve months (Butterworth and Jarrett, 1991; Baron-Cohen, 1995); that they report their desires, attribute desires to others and explain behaviour by means of desires by 2 years (Gopnik and Slaughter, 1991), understand and use propositional attitude locutions by 3 years (de Villiers and de Villiers, 1999) and pass the familiar false-belief tests by 4 years (Baron-Cohen *et al.*, 1985; Light, 1993; Lewis and Mitchell, 1994; Karmiloff-Smith, 1992). The fact that the developmental pattern is so regular has been taken by many to indicate that ToM is innate.

There are, however, some *prima facie* reasons for hesitancy in drawing this conclusion. The first is empirical; the second more principled. On the basis of empirical research conducted in several non-Western cultures, Vinden (1996, 1999) finds that this developmental pattern is not culturally universal. For example, her administration of standard ToM tests to a group of children aged 4 to 8 years from the isolated Junin Quechua language group of Peru reveals that a majority of the oldest children are unable to respond correctly to standard questions probing their understanding of their own and other people's false beliefs, and there is no statistically significant improvement of children's performance with increasing age. Similarly, in a further study of children and adolescents from four cultural groups, Vinden (1999) finds that those from Western culture (Europe, North America and Australia) significantly outperform those from each of the four non-Western groups in their levels of success on standard ToM tasks. Indeed 6-year-olds from the Tolai culture of Papua New Guinea are no better than chance, unschooled Mofu children from Northern Cameroon do not begin to pass ToM tests until they are between the ages of 7 and 10 years, and Tainae teenagers from isolated jungle regions of Papua New Guinea continue to perform at chance accuracy on questions about thinking through the ages of 14 and 15 years.

We will discuss these results further below, but for now we note that the explanation of the difference in developmental patterns across cultures could advert either to linguistic differences, cultural differences regarding the understanding of behaviour and its causes, or both. Moreover, a fixed developmental pattern, whether within or across cultures, could reflect the innateness of the *processes and structures subserving* ToM, coupled with the relative uniformity of the relevant social and environmental parameters, and not the innateness of ToM, *per se*, with such variation as there is reflecting variation in the latter dimension. In fact we will argue for just such an interpretation.

The second major argument for the innateness of ToM is that full-blown ToM has obvious developmental precursors, such as the shared attention mechanism noted above, the ability to discriminate maternal voice, the disposition to attend differentially to conspecific faces, to speech sounds from the native language, etc, which are obviously innately determined. Given that

these are each causally necessary for, and some even partially constitutive of ToM, it is plausible that the acquisition of each of these capacities is at least in part a stage of the innately determined acquisition of ToM. Again, though, while this argument may raise the plausibility of an innateness hypothesis, we note that it is not compelling. For, as we argue above, even if each of these causally necessary capacities is itself innate, this underdetermines the innateness of ToM. To establish that stronger conclusion, one would have to demonstrate that they are *sufficient* as well. We will present evidence below to show that they are *not*.

Phylogenetic evidence has been taken as a third argument for innateness (Hobson, 1994; Mitchell, 1994; Mitchell, 1997). There is considerable evidence that great apes, including chimpanzees, bonobos and gorillas attribute mental states to one another in planning behaviour and stratagems, and that they make use of information regarding visibility, available information, etc., in attributing these states (Whiten and Byrne, 1988; Premack and Woodruff, 1978; Povinelli *et al.*, 1991). Since it is plausible that these capacities are innate in the apes, and since it is plausible that we share much of the innate cognitive endowment of these phylogenetic cousins, it is plausible that such mechanisms are innate in us. Moreover, since many of the above-noted precursors to full-blown ToM are also present in these species, and since, despite a social environment very different from ours and despite the absence of language, they develop *a* ToM, albeit, perhaps, less sophisticated than ours, it would appear that these innate precursors are indeed *sufficient* for the development of ToM.

Once again, we urge that this argument be treated with caution. We note first that impressive as the successes of the apes are in ToM tasks, their failures are also spectacular. (Povinelli *et al.*, 1990). Indeed the pattern of successes and failures, when taken together is at least compatible with, and probably *favours* an explanation in terms of (complex) understanding of the regularities governing the behaviour and perceptual capacities of conspecifics, as well as a number of innately determined capacities and behavioural/cognitive dispositions subserving social interaction (perhaps constituting a social intelligence module). Moreover, the ape 'ToM', even on the most charitable reading, is considerably impoverished as compared with that with which a four-year-old *Homo sapiens* is endowed. So, even *were* we to *accept* that the apes have a ToM *and* that it is innately specified, we would still not have the conclusion either that the *more sophisticated human* ToM is that of the apes or that *human* ToM is innately specified.

The final, and perhaps most compelling, argument for the innateness of ToM is the fact of double dissociation of ToM from general intelligence. A host of researchers (Baron-Cohen, Leslie and Frith, 1985; Baron-Cohen, 1991; Mitchell, Saltmarsh and Russell, 1996) have noted that children with Downs syndrome develop ToM at relatively normal mental ages while autistic children with comparable or higher IQ's frequently fail to develop ToM. Moreover,

ToM development can appear normal despite a gross impairment of most other cognitive functions, as in studies we will review of children with Williams syndrome. So, it appears that the development of ToM is driven by its own autonomous dynamics, independent of social or general intelligence. We review these data in detail below. Much of our empirical argument, and the analysis of the pattern of results reported in the literature below, will be devoted to explaining why this evidence does not support the strong nativist position, and why it does support the much more modest partial nativism and multi-dimensional account of ToM development we will articulate. For the moment we simply note that these dissociations, impressive as they are, are not the whole story. Evidence from the development of ToM in deaf children (Peterson and Siegal, 1995, 1997, 1999; de Villiers and de Villiers, 1999), that we will examine closely in a later section, suggests that the crucial association is between ToM and language development, together with social experience and conversation. The dissociations noted in autism literature are, we will argue, parasitic on this more fundamental relation.

The principal arguments for the modularity of ToM point to the characteristic features of modular phenomena (Fodor, 1983; Garfield, 1987, 1994): innateness, domain-specificity, mandatoriness, speed, characteristic breakdown patterns and informational encapsulation. Proponents of a strong modularity thesis accept Fodor's claim that these traits cluster and together constitute modularity, and moreover that ToM constitutes a module in just this sense. Baron-Cohen (1995), Baron-Cohen and Swettenham (1996) and Carruthers (1996) advance this view. We will defend a weaker modularity claim, arguing that the Fodorian properties dissociate, and that ToM knowledge and competence exhibits some, but not all of these properties.²

² We note here (and this will become much more explicit below, that our concept of modularity is different from that employed in orthodox, Fodorian modularity theory. We adopt a view consonant with that of Garfield (1994) and Karmiloff-Smith (1992) according to which modules need not be innate, but can be acquired, and indeed assembled from other modules, perhaps innate, perhaps acquired. To the extent that innateness is essential to modularity, our position is anti-modularist. But we see no reason to accept the claim that innateness is essential to modularity, and think that there is good reason to employ the concept of modularity as we develop it here. We will discuss Coltheart's (1999) alternative to a Fodorian conception of modularity according to which domain specificity is the only essential feature of a module below, and will argue that it is in fact best to see speed and mandatoriness as the crucial features of modules with other modular features subserving those. Some (Thomas and Karmiloff-Smith, 1998) argue that the term 'modularity' has been re-defined so frequently and is now used in so many ways that it has become useless. We disagree. It often happens in science that a theoretical term is introduced to denote a phenomenon that is not completely understood. An attributive understanding of the semantics of such a term might lead one to think that every time our views about the phenomenon in question change the meaning of the word changes, and that divergences in views entail that disputants are not using the term in question in the same sense. But this is the wrong way to think about the semantics of theoretical terms. A term, such as 'electron' or 'modularity' is introduced to denote a particular phenomenon, about which we acknowledge that our knowledge is incomplete and that many of our beliefs may be wrong. The term denotes

We have already considered the arguments for the innateness of ToM. If the principle arguments for the innateness of ToM are undermined, and if innateness is rejected as a necessary condition of modularity, any argument from the innateness of ToM to its modularity is discredited. We do not think that all modules are innate and we will argue that ToM is most likely an *acquired* module. Let us now consider modularity as a phenomenon and ToM as an instance of that phenomenon more carefully.

Cognitive modules plausibly evolve because they are good for something, and that something is overwhelmingly likely to be fast, mandatory processing. It is good for beings who live in real time to have an array of fast, mandatory interfaces with the real world, e.g. gaze direction detector, motion detectors, devices for detecting the sexual receptiveness of conspecifics and for detecting stereotypic threats. Without them, we become panther chow or outcasts, which in the long run amounts to the same thing—failure to reproduce. Speed and mandatoriness are hence, from an adaptational and functional point of view, what drive such features as encapsulation (*ceteris paribus*, the more encapsulated a process is the faster it is, and the less choice about when it comes into play) and domain specificity (the more narrowly dedicated a process is, again, *ceteris paribus*, the faster and more automatic it can be). But in each case, *only ceteris paribus*. *Mutatis mutandis* this is true of neural localisation.³ We therefore think of modules as self-contained packages that do some particular thing quickly and mandatorially. We treat the other modularity features as contingently related to these, and treat questions about their presence in a module as always empirical. Mentalising is, in normal humans over the age of three (and fragmentarily before that) fast and mandatory. This is disputed nowhere in the standard literature on ToM. It is hence at least to some extent, in our sense, modular.

Now, Fodor (1983) argued that innate specification was a central feature of modularity, and we agree that many of the processes *subserving* mentalising are innately specified. But the claim that *all* modules are innately specified is absolutely wrong. Reading is modularised, as is expert perception in domains such as chess and music. All are fast, mandatory, relatively informationally encapsulated, and may even be neurally localised. And all are domain specific

that phenomenon regardless of whether it happens to have the precise set of properties we initially ascribe to it. That is why we can have scientific debates not only about the properties of electrons, but about those of modules. Now, to be sure, it might be that we discover that nothing has even a weighted most of the properties most theoreticians associate with the term in question (as was the case with 'phlogiston') in which case we decide that the term is empty of reference and move on. But if we—or indeed any of our principal interlocutors in the modularity debates—are even approximately right about modularity, then, while modules do not have the properties ascribed to them by Fodor, the term has considerable utility in that many of the classic modularity properties continue to attach either essentially or typically to the structures denoted by 'module'.

³ We do not address neural localisation arguments specifically in this study. In our view there are no compelling data suggesting either that ToM is localised or that it is not.

wherever they are instantiated. But *none of* these is innately specified. Indeed, for a social, communicative species occupying a malleable and challenging environment, it makes good evolutionary sense to develop the capacity to acquire new modules as needed, or to modify substantially innately specified modules. Only such an organism could develop new fast, mandatory processes and so adapt easily to changing environments. This raises important issues about how acquired modules are acquired, and how and to what degree particular acquired modules ride piggyback on innately specified modules. We will return to these questions below.

Encapsulation is matter-of-degree-ish, and comprises several distinct properties including imperviousness to top-down processing, shallow output, opacity to introspection, or isolation from general extra-modular influence. These properties dissociate from one another and vary from module to module in degree of realisation. It is implausible that a ToM module would be encapsulated to any significant degree in any of these senses, and there is no evidence presented even by the friends of modularity to suggest that it is (despite the fact that many of the more basic capacities underlying ToM are plausibly somewhat encapsulated, such as gaze monitoring). ToM reasoning is simply too dependent on general knowledge about the goals, attitudes and information available to those about whom we are reasoning, and it is, well, too much a matter of *reasoning* to be encapsulated. As Gopnik and Meltzoff (1997) argue, ToM reasoning seems to exploit general cognitive mechanisms and a wide range of information, and in this respect is largely similar to other theoretical reasoning.

Domain specificity is what really drives the ToM debates. Dissociation studies, cross-species evidence and developmental uniformity all provide compelling evidence that for humans and their close evolutionary relatives there is something *special* about the psychological domain—we develop a particularly sophisticated battery of skills, propensities and knowledge especially for navigating that domain. We agree, and so agree that ToM is modular in *this* sense as well. Given the adaptive significance of success in this domain it is tempting then to argue that specific innate mechanisms would be selected for use in this domain, and that the speed, mandatoriness, etc, are realised by an encapsulated, innately specified module. But the argument from domain specificity either to encapsulation or to innateness is, of course, at best nondemonstrative, and at worst, simply fallacious. (See also Currie and Sterelny, in press.)

Noting that Fodor (1983) did not intend his cluster of module characteristics to be treated as definitional criteria or necessary conditions for use of the concept, Coltheart (1999) takes a more ambitious line himself by actually proposing a definition of modularity. According to Coltheart 'a cognitive system is modular when and only when it is domain-specific' (p. 115). On this basis, Coltheart urges that questions about whether or not particular modular systems also possess other Fodorian properties like innateness should be treated as purely empirical. Thus, were we to conclude that there is a ToM

module in Coltheart's sense, the question regarding whether or not it is innate, encapsulated, etc., would only be settled by data of the kind we canvass below.

As Coltheart correctly points out, demonstrating that a capacity is domain-specific says *nothing* about the extent to which that capacity is autonomous. It may achieve that domain specificity through recruiting other processes, which themselves may or may not be domain-specific. This is bound up with the larger question of domain individuation and leads to our worries about Coltheart's criterion and our preference for speed and mandatoriness as the fundamental criteria of cognitive modularity. Our knowledge about Microsoft Word is domain specific. It is useless for anything else. And it can be conceived quite naturally as a cognitive ability—one manifested as we write this paper. But it is not for all that modular. Now of course Coltheart would argue that neither is it a cognitive system. But by what criterion? There seems to be no non-question-begging way to rule out such pseudo-domains and pseudo-modules unless there is an independent way of individuating domains and modules.

Moreover, what constitutes a domain for the purposes of a claim of domain specificity is only determinable *a posteriori*, and may be malleable as a module developed for one domain is recruited or extended to another. Letter recognition skills specific to the Roman alphabet may come to operate with speed and mandatoriness in Hebrew or Tibetan; the recognition of the meaningfulness and the power of language may be recruited as a mechanism for interpreting inner states. Phenomena that might appear to belong to different domains sometimes come to be seen as falling in the same domain for a mind, depending on how a mind processes information regarding them. Or phenomena that appear to be co-resident could indeed turn out to belong to different domains. For example, face recognition and object recognition, which might pretheoretically appear to be comprised by the same cognitive domain and to be subserved by the same cognitive functions, have been shown to be in fact quite separate domains (Coltheart, 1999). But those very different functions are themselves fast, mandatory special-purpose cognitive operations which can be identified independently of the domain(s) in which they operate. This is not to say that most modules are not domain-specific in an important sense. Here we agree with Coltheart about the *centrality* of domain-specificity to modularity. We disagree only about whether it is fundamental. We see domains as emerging from modules, not modules as conforming to pre-existing domains. So, while we will urge that ToM is in an important sense a domain-specific module, we will urge that both its domain and its specificity to that domain emerge from the recruitment of more general-purpose cognitive processes, some of which derive from the operation of very different modules whose domains cross-cut this one.

We thus urge that there is room for a moderate position both with regard to the innateness question and with regard to the modularity question, and we will defend such a moderate position. That is, we will argue that while some of the processes that subserve ToM are innately specified, and modular

in some senses, the acquisition of ToM is dependent as well on social and linguistic accomplishment, and that it is modular in only a weak sense. Hence our developmental model will be social and ecological as opposed to being individualistic; our account of what develops will be that of a set of skills and dispositions and not articulated theory; our epistemology will be externalist and empiricist and not Cartesian. We now turn to the empirical picture.⁴

3. Review of Empirical Data on Deaf, Blind and Autistic Children

3.1 Autism

Baron-Cohen, Leslie and Frith's (1985) discovery of much poorer comprehension by autistic subjects than by either normally developing children or mental-aged-matched retarded children with Down's syndrome of the behavioural consequences of a person's entertaining an objectively false belief stimulated a great deal of further work using a set of carefully contrived empirical methods for assessing children's understanding of beliefs and other mental states. These have come to be known as *standard false belief tests*. Owing to their elegant simplicity and careful controls to overcome problems unrelated to mentalising, such as inadequate language comprehension, excessive memory demands and random guessing, they have gained the status of 'litmus tests of theory of mind' (Baron-Cohen, 1995, p. 59).⁵

A large body of research using these tests to replicate Baron-Cohen *et al.*'s (1985) findings now conclusively shows both that ToM concepts are exceptionally difficult for individuals with autism, and that the deficits connected with autism are specific to concepts of mental representation. Whereas autistic children, adolescents and adults typically fail both changed location and misleading appearance tests of false belief understanding at mental ages well beyond four years, which is the normal age of acquisition, mentally retarded children of comparable or lower mental age often succeed on the same tasks, ruling out general cognitive deficits as an explanation for autistic children's delays (e.g. Baron-Cohen *et al.*, 1985; Perner, Frith, Leslie and Leekam, 1989). Indeed, when Happé (1995) combined an analytic review of the results of 27 separate studies of autistic children's ToM performance that were published between 1985 and 1993 with her own further study of a sample of 70 individuals with autism (mean age = 12.25 years) on two false belief tasks (the Sally-

⁴ We hence agree with Currie and Sterelny (in press) both that ToM is weakly modular and that it is an *acquired*, as opposed to an *innate* module, though acquired in part through the operation of more strongly modular, innate modules. As will become clear, however, we part company in our explanation and characterisation of that module, its means of acquisition, and its relation to other allied cognitive structures.

⁵ Baron-Cohen *et al.*'s (1985) original procedure known as the 'Sally-Ann' task, has itself become seminal as a method for assessing ToM. It tests children's capacities to infer the search behaviour resulting from the false beliefs of an actor who lacks information that the children themselves possess.

Ann changed location task and the misleading container task (Smarties test), in comparison with normally developing and mentally retarded control groups, she obtained strong evidence of a false belief deficit that was specific to autism. Collectively sampling more than 300 relatively able subjects with autism (all with mean verbal mental ages of 5 years and over), the results of these studies consistently revealed severe difficulties on false belief tasks among autistic children, adolescents, and young adults that were not echoed among retarded children or normal developers of similar mental age. For example, pass rates for autistic subjects in 14 of studies with samples predominantly in their teens ranged from only 15 to 60 percent, with a mean of just 33 percent passing. In Happé's own study, only 20 percent of autistic 12-year-olds passed both false belief tasks, as compared with 56 percent of normal 4-year-olds and 59 percent of 12-year-olds with mental retardation.

Furthermore, this dramatic failure by individuals with autism to grasp the concept of false belief is found to be specific to representations that arise in the mental domain. Concepts of behaviour, of visual perception and, significantly, of photographic representation, are all seen to develop normally in many autistic individuals who are incapable of passing ToM tests. For example, Baron-Cohen, Leslie and Frith (1986) find that autistic children who succeed readily in arranging pictures to record sequences of physical events and overt behaviours are selectively incapable of performing the same task when the stimuli depict mental states, and Baron-Cohen (1991) discovers unimpaired understanding of simple emotions (happiness and sadness as outcomes of situations) in autistic people who fail standard false belief tests. Similarly, Reed and Peterson (1990) find that autistic children and adolescents are as capable as matched groups of mentally retarded adolescents and normal preschoolers of making accurate inferences about visual perception. They can both infer invisibility based on blocked line of sight, and identify the varying percepts of viewers observing the same scene from different vantage points. However, despite these abilities, these same autistic children routinely fail corresponding tests that differ only in that the alternative perspectives are cognitive rather than perceptual. Thus these subjects can neither infer ignorance (as opposed to invisibility) based upon blocked informational access, nor false belief based on the misleading mental input available to another mind.

Studies by Leekam and Perner (1991), Leslie and Thaiss (1992) and Peterson and Siegal (1998) of autistic children's understanding of false photographs provide a particularly striking demonstration of the specificity of these children's deficits to concepts of mental representation. Even though the procedures, instructions and syntactic forms of questioning used in the false photographic and false belief versions of their respective sets of tasks are almost identical, significantly better performance is displayed in each study on the photographic than the belief version by autistic children, in contrast to normally developing control groups. (Children in these latter groups generally find false belief tests either as easy as, or significantly easier than, tests of similarly false representation by a camera).

Verbal mental age, as assessed using a standardised test of receptive vocabulary, is a significant predictor of ToM success for both normally developing and autistic children. However, the actual levels of verbal mental age required for passing false belief tasks are likely to be higher for autistic groups. For example, Happé (1995) finds that whereas about 25 percent of normally developing children pass multiple ToM tasks by a verbal mental age of 3.50 years, and this rises to 80 percent at 4.50 years, no autistic subject passes with a verbal mental age of 5.50 years or less, and even at a verbal age of 9.00 years, a pass rate of only 50 percent is achieved. On the other hand, all four of the autistic subjects with verbal mental ages above 11.50 years pass both tasks, suggesting that the ToM concepts that normal developers acquire between verbal and chronological ages of 3 and 4 years are eventually mastered by a highly functioning minority of individuals with autism, though at significantly more advanced levels of chronological and linguistic maturity than those required for normal acquisition.

This finding, as we shall see, provides evidence for our conclusion that linguistic and social skills jointly support the development of ToM. Autistic children suffer from severe impairments in both of these domains. So, we argue, it is not surprising that they are severely impaired in the development of ToM. If ToM were an autonomous module selectively impaired in autistic children, then these children would never develop it. If, on the other hand, ToM, however modularised it may be, develops as a consequence of the development of social intelligence,⁶ and linguistic capacity, then, to the extent that one or the other of these capacities develops, some ToM skills should be demonstrated. The fact that this is observed in high-functioning people with autism is evidence for our view. But far more striking evidence is provided by the pattern of development of ToM in deaf and other sensorily impaired children.

3.2 Deafness

When a child is born profoundly deaf into a hearing family, there are likely to be many departures from the normal courses of development of language, conversation and social understanding. Even when hearing parents make extensive efforts to learn sign language, it is rare for them to attain the same level of proficiency as a native speaker. According to Vaccari and Marschark: 'Over 90 percent of deaf children have hearing parents, the majority of whom either do not know sign language or have relatively little skill in that domain' (1997, p. 793).

In contrast to parents who are deaf native signers, hearing parents typically report difficulties in communicating with their deaf children even about familiar everyday routines and have extreme difficulty sharing their thoughts,

⁶ Following Baron-Cohen (1991) we use the term 'social intelligence' synonymously with 'social cognition' to refer to 'those aspects of the cognitive system that are used in understanding the social world' (p. 302).

memories, intentions, and beliefs (Meadow, 1975). In many hearing families with a deaf child, any signs or communicative gestures that are produced by parent or child are restricted to topics in the immediately perceptible visual field, leaving parents and offspring alike unclear about one another's needs, desires, beliefs and capabilities (Vaccari and Marschark, 1997). Hearing mothers of deaf children are found to discuss their emotions and intentions rarely, if at all, and may adopt a didactic role which discourages playful or inquisitive conversational exchange (Courtin and Melot, 1998).

Consequently, most deaf adults who eventually become fluent users of a sign language like signed English, ASL (American Sign Language) or Auslan (Australian Sign Language) acquire this language belatedly after varying periods of restricted conversation in their hearing families of origin. For example, Power and Carty (1990), in a survey of deaf native speakers of Auslan, discover that: 'in 90 percent of cases Auslan is learnt not from parents within a family setting, but from other deaf students, usually in school' (p. 223). This means that 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 may have no obvious visual referent. This is consistent with research showing that 'deaf child of hearing parents may have no language in the sense of a code shared by many users' (Charrow and Fletcher, 1974, p. 436) until school entry at the age of 5 or 6 years.

Conversations about unobservable thoughts and feelings are apt to be selectively curtailed. Morford and Goldin-Meadow (1997) study four profoundly deaf children who, despite not being exposed to a useable conversational language, manage to express themselves at home by means of idiosyncratic systems of gestures known as 'homesign' (p. 240). Only one of them makes spontaneous references to fantasy, hypothetical ideas or future events in conversation. Initiations of communication about the non-present by their caregivers are even less frequent. According to Marschark (1993): 'Deaf children are less likely than hearing children to receive explanations from their parents concerning emotions, reasons for actions, expected roles and the consequences of various behaviours' (p. 60).

Hence, severely and profoundly deaf children who grow up in hearing families may encounter the same degree of difficulty as autistic children in engaging in conversations with family members about false or imaginary beliefs and other abstract mental states, though for very different reasons (Tager-Flusberg, 1993). A normally intelligent and sociable deaf child with no symptoms of autism and no impairments apart from auditory handicap is likely to be blocked from conversation, especially about propositional attitudes, owing to the simple lack of a fluently shared common language.

Peterson and Siegal (1995) tested a sample of 26 signing prelingually deaf children of hearing parents on the same two-trial Sally-Ann 'litmus' test of false belief understanding that is routinely passed by normally developing children by age four

and discovered a high failure rate. Indeed, a majority of deaf children aged 8 to 13 years (65 percent) failed this task, despite being intellectually normal and fluent in sign language at the time of testing. There is no statistically significant difference between the performance these belatedly signing deaf children from hearing families and the average levels of performance observed among autistic children of similar chronological and nonverbal mental age (Happé, 1995). The similarity to groups with autism, and significant delay behind hearing children, is replicated in three subsequent investigations of fresh samples of late signing Australian deaf children (Peterson and Siegal, 1997, 1998, 1999).

Russell *et al.* (1998) observe similar delays in a further study, in which the same Sally-Ann test, with its stringent controls for language comprehension, memory and guessing, is administered by a fluently signing interpreter to a sample of signing Scottish deaf children from hearing families. This study confirms the hypothesis that there are gains in deaf children's understanding with increasing age. Consistent with Peterson and Siegal's observations of Australian deaf children, only 17 percent of the Scottish deaf children aged 5 to 7 years are successful, whereas 60 percent of the deaf adolescents (aged 13 to 16 years) pass the test. These results indicate that when deaf children grow up in hearing families without any other member who is fluent enough to converse freely about beliefs and other intangible mental states, the development of a ToM is delayed, but not completely precluded. Eventually, a majority of late-signing deaf children are likely to acquire an understanding of false belief, possibly as a result of conversing with signing deaf classmates in school.

However, (unless they are native signers) profoundly deaf children lack early access to fluent conversation about mental states and these data consistently show that they are also correspondingly delayed in developing ToM, just as our theory predicts. Furthermore, the delays observed among deaf children also resemble those displayed by autistic children in being specific to mentalistic concepts, rather than pertaining to a false representation more generally. In two experiments, Peterson and Siegal (1998) administered two tests each of false belief understanding and of the understanding of false photographic representation to matched groups of children from three populations: a) autistic children; b) normal preschoolers; and c) signing deaf primary school children from hearing families. The children with deafness, like their peers with autism, perform at chance levels on false belief tests, while succeeding readily on the mechanical versions of each task, which employ identical procedures and question syntax, differing only in that they require reasoning about nonmental representations such as photographs or drawings. Normally developing 4-year-olds, by contrast, achieve near perfect performance on both types of tasks, whereas ordinary 3-year-olds are no better than chance on either of them.

Peterson and Siegal (2000) review the results of 11 separate studies of false belief understanding in signing profoundly deaf children published between 1995 and 1999. These studies include children from Australia, France, England, Scotland and the United States, the vast majority of whom acquired signing

belatedly upon school entry. Taken collectively, the populations of severely and profoundly deaf children tested in these 11 studies are impressively varied, having been drawn from a wide variety of family circumstances, sign language communities, approaches to deaf education, and preferred communication modalities. Overall, the results of these studies support Peterson and Siegal's (1995) original findings. Deaf children from exclusively hearing households who eventually become fluent signers are seriously delayed in mastering the concepts of false belief that give evidence of a ToM, notwithstanding normal levels of nonverbal intelligence and freedom from the severe social and affective impairments with which autism is identified (Peterson and Siegal, 2000).

Interestingly, however, the results of this collective body of empirical evidence on deaf children's performance of ToM tests suggest that it is not deafness *per se*, but rather deafness in conjunction with a family in which all other members are hearing (and hence, typically containing no signers with native fluency), that is responsible for the delays observed in deaf children's development of false belief understanding. Profoundly deaf offspring of signing deaf parents (second generation deaf signers), along with those who have another native speaker of sign language in their immediate household (e.g. a signing deaf grandparent or an older deaf sibling who has become a fluent signer at school), can be dubbed 'native signers' owing to their access, throughout their growing up, to a fluent native conversational partner with whom they are able to share a common first language, such as ASL or Auslan.

The studies of false belief understanding that include native signers in their samples reveal significantly higher levels of performance by these children with fluently signing conversational partners than by late-signing deaf children from hearing families of comparable age and nonverbal intelligence (Peterson and Siegal, 1999; Rimmel, Bettger and Weinberg, 1998). Indeed, though the numbers of very young native signers in these samples are often so small as to preclude statistical analysis, it appears that native signers may develop the concepts of false belief that enable success on tasks like the Sally-Ann and Smarties at the same age as normally hearing children do, near the fourth birthday.

It is certainly quite probable that many of the native signers who perform at or near the ceiling on tests of false belief understanding that were administered when they averaged 8 years (Rimmel *et al.*, 1998) or 9 years of age (Peterson and Siegal, 1999) would have mastered these concepts many years earlier as a result of fluent conversations in sign language with native signing family members. Indeed, Meadow, Greenberg, Erting and Carmichael (1981) report that native signing deaf children converse as fluently in sign about non-present ideas, objects and events with their signing deaf relatives as hearing children with their hearing parents in spoken language.⁷ Since our theory pre-

⁷ While this does not amount to clear evidence of the possession of ToM, in virtue of the fact that such conversation and ToM skills covary (such conversation is absent or severely reduced in autistic children and in deaf children of hearing parents), it is certainly suggestive.

dicts that children with normal social intelligence, normal social experience and normal linguistic development will develop ToM normally, but that impairment in any one of these areas will impair ToM development, the fact that these native signing deaf children develop ToM normally, while their linguistically or socially deprived deaf colleagues do not, provides powerful empirical support for our account. Deaf children who are native signers do possess a linguistic fluency with which to discuss mental states and have normal social experiences. Thus the finding that they develop ToM at the same age as normal hearing children provides crucial evidence for our theoretical framework, according to which ToM requires both social interaction and the linguistic skill to converse freely with peers and family members about unobservables like mental states.

Peterson and Siegal (2000) review experiments investigating the development of ToM in 'oral-aural' or vocally trained children who, often with lesser levels of hearing loss than their signing deaf peers, were taught to speak and comprehend spoken language with the assistance of hearing aids. (Though not always successful in children with serious hearing losses, oral training does enable some deaf children to participate in family conversation through the spoken modality). Oral deaf children's ease of mastering ToM may depend upon the level of language skill that these children achieve during the preschool and early primary school period. Even when oral training is eventually successful, language development is frequently delayed, and this may be associated with commensurate delays in acquiring concepts of false belief (de Villiers *et al.*, 1997). On the other hand, when hearing losses are less severe, early amplification with external hearing aids or cochlear implants sometimes enables oral-aural deaf children to achieve fluent spoken communication with hearing family members at an early age. This could explain Peterson and Siegal's (1999) finding of significantly better false belief performance by orally trained deaf pupils in Total Communication Australian primary schools than by their late-signing deaf classmates from hearing families.⁸

These observed variations in rates of ToM development among deaf children as a function of their language modality, and their corresponding access to fluent and varied communication with family members, are consistent with accounts of the growth of mental state understanding that emphasise conversation and social experience, since the deaf children who are the earliest to develop concepts of false belief also have the greatest access to these potentially stimulating influences.

⁸ Of course, exposure to bilingual education may also have helped the oral group, who had sufficient spoken language skills to cope with false belief testing in a purely oral modality, yet were mostly also fluent enough in sign to converse freely with their signing severely and profoundly deaf classmates.

3.3 Blindness

Children who are congenitally blind are also likely to have difficulty communicating with sighted family members, especially when it comes to conversing about intangibles like false beliefs and other abstract mental states. Blind children are often slow to acquire language and may persist in having such pragmatic difficulties as the confusions over conversational pronoun reversal that Tager-Flusberg (1993) also observes in autistic children long after basic spoken language skills have been acquired (Fraiberg, 1977). Blindness also deprives infants and toddlers of visual exposure to their family members' facial expressions of emotion, gaze direction (Baron-Cohen, 1995) and other non-verbal indicators of mental state. Thus a sociocultural explanation for the growth of a ToM as the product of early social experience and conversational interaction predicts delays for blind children that may be on a par with those observed among late-signing profoundly deaf children.

Using two Misleading Container tests, McAlpine and Moore (1995) examined concepts of false belief in a total of 16 visually impaired children with a mean age of 6 years. Partially sighted children, with visual acuities of 20/300 or better, perform as well as normally developing children on both tasks. A total of 80 percent of these children pass both false belief tasks and 90 percent pass at least one. However the remaining 6 children, with acuities of only 20/400 or less, display problems with a ToM. Two-thirds of them fail both misleading appearance tests, and those who pass any task are more than 9 years old. With due recognition to the limitations imposed by such a small sample, McAlpine and Moore (1995) propose that severe blindness may inhibit the development of a ToM, while not completely precluding it. Thus 'children with severe visual impairments are delayed but not deficient in acquiring an understanding of false belief' (p. 354). They speculate further that a ToM might develop in blind children by the age of 12 years, though there were too few older children in their sample to provide statistically reliable support for this suggestion.

Minter, Hobson and Bishop (1998) tested a larger group of totally blind children on false belief tasks of two different types. Their sample, like McAlpine and Moore's, is relatively young, consisting of 21 children ranging in age from 5 to 9 years with a mean of 6 years 11 months. Since birth, these children had either no vision at all or only minimal light perception. One false belief test involves Misleading Containers while the other is a variant of Sally-Ann involving false beliefs about the locations of objects.

Minter *et al.*'s (1998) findings support those of McAlpine and Moore by showing delays in ToM development among blind children. As among late-signing deaf children (Peterson and Siegal, 1995) and high-functioning children with autism, a majority of blind children fail at least one task, as compared with only 10 percent of an age-matched sighted control group.

Peterson, Webb and Peterson (2000) also observe deficits in young blind children's performance on four tests of false belief understanding, two involv-

ing the changed location of objects and the others, the misleading appearances of containers. At the age of 6 years, only 14 percent of their totally blind and severely visually impaired samples pass all four tasks and percentages correct on each task individually range from 29 percent to 50 percent, this latter figure being the level of chance performance. Some improvement is noted between the ages of 7 and 10 years, though only half in this age group display a sound grasp of ToM by passing all four of the false belief tasks. Eventually, between the ages of 11 and 12 years, success rates rise to those typically displayed by 4- to 5-year-old sighted children. Seventy percent of blind 11- and 12-year-olds pass all four false belief tasks and all manage to pass at least one.

The results of a multiple regression analysis confirm that whereas blind children's ToM performance improves significantly with increasing age, there is no significant difference at any age between totally blind and severely visually impaired children. As Reed and Peterson (1990) observe among autistic children, there is likewise a disjunction between comprehension of visual perception and failure to understand comparable concepts in the mental domain. The vast majority of blind children even in the youngest group perform perfectly on measures of visual perspective-taking, indicating that their difficulties with ToM are not a consequence of an inability to comprehend other people's perceptual access to information.

3.4 Individual Differences in Rate of ToM Development in Normal Children

In addition to the severe delays that are generally observed in the development of ToM by children with autism, deafness and blindness, recent empirical studies also reveal reliable individual differences of smaller magnitude among normally developing preschoolers in rates of mastering concepts of mind. The bases for these contrasts between normal children who are precocious or slow to acquire mindreading may also prove instructive in adjudicating among different theoretical accounts of ToM.

Family conversation about the mind is one such variable, as revealed in Dunn's (1994, 1996) intensive longitudinal studies of family conversations in the households of normally developing preschoolers with one sibling, in the USA and in the UK. Dunn's findings reveal marked variations in the frequency in which children discuss belief, intentions and feelings with their mothers and siblings during play, conflict and everyday routines and suggest the frequent conversations about the mind can accelerate the growth of ToM. For example, using a longitudinal methodology, Dunn, Brown, Slomkowski, Tesla and Youngblade (1991) find that the breadth and depth of the conversational exchanges involving mental state information spontaneously arising between 33-month-olds and their mothers and siblings is a significant predictor of the children's false belief understanding some 7 months later. Children who are able to explain story characters' behaviour in terms of false belief at 40 months had more frequent family conversations as 2-year-olds about emotions, desires

and psychological causality (e.g. 'Why don't you like to eat ice-cream before dinner?') than failers, even when matched for age and overall verbal fluency.

Sibling constellation is another predictor of individual variation in ToM acquisition. Only-children have no access at home to child conversation partners, whereas sibling children can converse and socialise with brothers and/or sisters as well as with their parents. Thus differences in rates of ToM development among normal children with no developmental, neurological or sensory problems, who differ simply in sibling constellation, would be consistent with our model, but would be harder to reconcile with a nativist account.

In fact, a number of recent empirical studies suggest that the presence of siblings in their immediate families may accelerate children's development of false belief understanding. Perner, Ruffman and Leekam (1994) test 76 children aged 3 and 4 years on a narrative changed-location test of false belief understanding and discover a linear improvement in performance with increasing family size. The effect is substantial. Only about half the number of only-children pass the task as children with two or more siblings, and this represents a difference as large as that between 3-year-olds and 4-year-olds, the year of steepest improvement in normal ToM performance. Children with just one sibling score intermediately, doing significantly better than only-children, but worse than those with two siblings or more.

Jenkins and Astington (1996) ask whether differences in rates of language development between only-children and those with one or more older or younger siblings might contribute to these variations in ages of mastering concepts of belief. Their sample consists of 68 children from university-affiliated daycare or nursery schools, comprising 22 only-children, 16 eldest children from families of two, 22 youngest children from families of two to four, and 8 middle children with both younger and older siblings. Their results confirm Perner *et al.*'s (1994) findings of better false belief performance by children from larger families, and this effect remains statistically significant even after the influences of chronological and verbal mental ages are controlled. But birth order exerts no statistically significant influence upon false belief understanding, either in the main analyses, or in a separate comparison between eldest and youngest siblings in families of exactly two children. Jenkins and Astington conclude that: 'It is the number of siblings that the child has that is important for the development of false belief rather than whether these siblings are older or younger and how far distant in age they are' (p. 75).

Lewis, Freedman, Kyriakidou, Maridaki-Kassotaki and Berridge (1996) examined the development of false belief understanding in two samples of Greek children in relation to their range of access to siblings, other child (e.g., cousins) and adult kin (parent, live-in grandparents, etc). A series of multivariate logistic regression analyses reveal that: 'the four variables that jointly best predict false belief performance are: a) the number of adult kin who live in close proximity; b) the number of older siblings of the child; c) the number of younger siblings of the child; and d) age' (p. 2936). The highest levels of

false belief performance in the sample are displayed by children whose sibling constellation is apt to offer the greatest variety of conceptual perspectives, namely: middle-born children who are able to converse with both younger and older siblings at home. In a second experiment, Lewis *et al.* (1996) assessed false belief understanding in relation to the subjects' actual levels of contact on the previous day with siblings and other child and adult members of their extended families. Results support varied social interaction with mixed-age playmates and conversational partners as responsible for early growth of ToM, prompting the conclusion that: 'theory of mind is not simply passed from one sibling to another by a process of social influence. It seems more likely that a variety of knowledgeable members of his or her culture influence the apprentice theoretician of the mind' (p. 2930).

Ruffman, Perner, Naito, Parkin and Clements (1998) report the results of four separate studies of false belief involving a total of 265 normally developing children aged 3 to 5 years in Japan and the UK. These studies show that children's rates of mastering false belief understanding increase in proportion to the variety of ages and conversational perspectives provided by the sibling constellation, with the presence of older brothers and sisters at home proving especially beneficial for early ToM. A close mutual understanding is likely to grow up among siblings in a family, and this is likely to motivate their interest in, and awareness of, the contents of one another's' minds. Perhaps this shared awareness, fostered through play, conflict and conversation with siblings, explains the rapid development of concepts of false belief, as confirmed in five studies with varied populations of children from varied cultures, educational and family backgrounds. Sibling children develop ToM more quickly than only-children, and also have more varied social experiences at home. Yet other research suggests that only-children develop language faster than a child with many closely-spaced siblings. (Zajonc and Hall, 1986). Furthermore, the sibling benefit for ToM is evident even after children's differing levels of linguistic maturity are partialled out. (Jenkins and Astington, 1996; Ruffman *et al.*, 1998). Thus these data support our suggestion that the varied social interactions that can foster social intelligence in child from a large family are jointly necessary, along with language, for the early growth of ToM.

As noted above in the context of Happé's (1995) studies of autistic children, variations in language ability among normal developers are also reliable correlates of individual differences in ToM. Preschoolers who rapidly master the vocabulary, syntax and semantics of their native language are also apt to display early mastery of false belief. Furthermore, this association between rapid language development and advanced false belief understanding is observed even after controlling for individual differences in nonverbal intelligence and cognition. (Jenkins and Astington, 1996).

Astington and Jenkins (1999) tested a sample of 59 3-year olds longitudinally over a period of 7 months and discovered that precocious linguistic skill fosters ToM but not vice versa. At later testing points, early mastery of the

vocabulary, syntax and semantics of language continues to predict advanced false belief understanding even after the effects of earlier ToM development have been statistically controlled. But earlier ToM fails to predict later language test performance once earlier language is controlled, leading Astington and Jenkins to conclude: 'it is not just that children need language skills to display their theory of mind in false belief tasks. Language plays a fundamental role in theory of mind development' (p. 1319).

4. Implications of These Data for the Strong Nativist Position

According to Frith, Happé and Siddons (1994, p. 110), 'In the normally developing child the computational capacity to represent mental states has an innate neurological basis. In the autistic child, neurological damage to a circumscribed system of the brain has occurred.'

The neurobiological version of the strong nativist position postulates that damage to an innately specified, modular, brain mechanism (called ToMM for 'theory of mind mechanism': Leslie, 1994) explains autistic children's failure to develop a ToM, as well as their other diagnostically significant social, linguistic and cognitive impairments. A number of different neuroanatomical candidates have been proposed as the locus of the problem. For example, Baron-Cohen (1995) suggests that ToMM is located in the orbito-frontal cortex, based on three lines of evidence: (1) Brain damaged adult patients with orbito-frontal lesions are sometimes described as deficient in 'social judgement' (p. 92); (2) Neuroimaging studies reveal increased blood flow and neural activation in the orbito-frontal cortex when volunteers perform tasks involving mental state concepts (e.g. selecting words like 'dream', 'imagine', 'hope' and 'pretend') but not when performing control tasks (e.g. selecting words like 'teeth', 'blood', 'walk' and 'eat'); (3) Animal studies show that lesions in the orbito-frontal cortex produce changes in social behaviour that are often accompanied by a loss of social status.

Of course, speculation even as to the existence of a neuroanatomical ToMM, let alone its locus, remains controversial. Though not all exponents of the nativist, modular account would agree in pinpointing the orbito-frontal cortex as the specific area that is deficient or damaged in autism, most theoretical variations on neurobiological explanatory themes do share a number of core assumptions, including the presuppositions that damage to the 'specialised cognitive mechanism which subserves the development of folk psychological notions' (Leslie and Thaiss, 1992, p. 237) is (a) innate, (b) a cause rather than a consequence of the linguistic and social impairments that autism also entails, and (c) a problem unique to the autistic disorder, rather than a correlate of mental retardation, language delay, emotional disturbance or other clinical conditions or developmental abnormalities.

While allowing that some experiential input may be required in order to trigger and guide the maturational development of the ToM model, the role

accorded to social experience in neurobiological accounts is generally quite minimal, a view inspired by Chomsky's (1986) model of a language acquisition device that is innately preprogrammed to recognize language and test linguistic input against a constrained and preordained set of test hypotheses about language structure.

According to the strong nativist account of ToM development, the deficits in both receptive vocabulary and social involvement that are also observed in connection with autism (Happé, 1995; Currie and Sterelny, in press) are consequences of damage to an innate ToM module. This damage is deemed to be the original cause that leads children with autism subsequently to have problems relating to others and mastering the lexical, syntactic and pragmatic complexities of language. As Happé (1995) explains:

In the autistic child...according to the theory of mind deficit account of autism, the failure to "mind-read" may significantly hamper word learning by interrupting processes of joint attention, reference and ostentation (Frith and Happé, 1994). For this reason, a measure of number of words known may inadvertently be a measure of theory of mind in children with autism (p. 85).

By contrast, cultural, socio-linguistic or socio-experiential accounts, (Lillard, 1997) while likewise quite varied, agree in their reversal of the direction of causality that is assumed by the nativist position. According to the socio-linguistic view, development of mental-state understanding follows from language development and social interaction. By observing people and engaging with them in play and conversation, normally developing preschoolers are exposed to others' mental states. They may also be exposed to deception, disagreement, pretence, and miscommunication when other people's expressed beliefs or thoughts conflict with their own. Eventually, these social experiences are seen to culminate in the construction of an effective ToM. According to these accounts, the delayed development of ToM in autistic children derives from the triad of social aloofness, language difficulties and deficient imagination. These deficits, according to this analysis, not only define autism as a diagnosed clinical condition, but also cause the autistic child's delays in ToM by limiting access to other people's thoughts, intentions, beliefs, and other mental states. In other words, selective deficits in the understanding of mental states may come about *because* a child with autism remains socially aloof from family conversational partners, has too few linguistic skills able to converse with family members about false or abstract ideas, and lacks the imaginative capacity that is necessary in order to fully appreciate someone else's beliefs. A deficit of pragmatic communication skill has been reliably identified with autism, beginning in infancy as absence of directive pointing and shared attention, and evolving into impairments of narrative discourse and inability to introduce new topics into conversation (Bruner and Feldman, 1993).

When these findings are combined with the results we review regarding the delayed onset of ToM in deaf and blind children, and particularly the degree to which that delay depends upon social interaction and language development in these groups, it is clear that the data from autism cited in support of the conclusion that ToM is subserved by an autonomous, innate, module that becomes selectively impaired in autism in fact provide no support whatever for that conclusion. Furthermore, the fact that there is no dissociation between ToM development and language acquisition lends additional support to our view. We now turn to a different source of evidence providing more positive support for our account.

5. Data Favourable to a Sociolinguistic Developmental Model

In order to demonstrate more directly the connection between social interaction, language acquisition and the development of ToM, we turn to observational studies of naturally-occurring child-parent interactions in a range of relevant populations. The pattern that will emerge is this: Children who have the opportunity to develop social knowledge, to interact socially and to develop language develop ToM. Children with sensory, social or environmental deficits that restrict any of these variables are delayed in their ToM development. Tager-Flusberg (1993), comparing spontaneous mother-child dialogues in households in which the child has either autism or Down's syndrome, finds significantly fewer references to mental states and appeals for joint attention in dyads including a child with autism, despite comparable levels of talk about topics outside of the cognitive domain, including simple emotions. She notes that, unlike their peers with Down's syndrome, autistic children 'never spoke about cognitions as they relate to behaviour or contrast with reality' (p. 169).

None of the mentally retarded children in Tager-Flusberg's (1993) sample commit pronoun reversal errors (e.g. saying 'you' to describe the self as a speaker), while all of the autistic children do at least some of the time, suggesting confusion over the pragmatic roles of speaker and listener in a conversational exchange. In addition, autistic children ask fewer questions and are less contingent or connected with their interlocutor's perspective in conversations (that is, less likely to expand, continue or oppose a topic their mother introduces). Their dialogues are also especially striking for their total absence of references to cognitive mental states (believe, dream, forget, guess, trick, wonder, pretend, etc). Tager-Flusberg concludes that: 'One of the primary functions of language, to serve as a major source of knowledge, is impaired in autistic children even in the prelinguistic period. It is this impairment which links deficits in joint attention, later problems with communication, and the understanding of belief' (p. 153).

Bruner and Feldman (1993) also note that one of the consistent differences between autistic children and mentally retarded or normally developing comparison groups of similar mental age and vocabulary size is a difficulty with

pragmatic conversation. Autistic children have difficulty sustaining a conversational exchange with another person. As Bruner and Feldman (1993) explain: 'In dialogue, autistic speakers seem unable to extend the interlocutor's previous comment' (p. 274). In addition, autistic children appear to suffer a narrative deficit which leaves them unable to construct a coherent story line and deprives them of the ability to 'make new comments on a topic in discourse' (p. 275). Owing to these pragmatic skills deficits, autistic children are likely, as Tager-Flusberg's (1993) observations have shown, to miss out on the kinds of dialogues with other members of their family in which the intentional stance is taken and in which they are given insight into the workings of their conversational partners' minds.

Hence, autism as a clinical disorder may serve to restrict conversational access to the kinds of social, linguistic and emotional inputs necessary for the development of a ToM in normal children. According to Smith (1996), 'Young incipient mind-readers need to be supported in the ontogenetic development of mind-reading skills. We need to consider as prerequisites both individuals who can develop mind-reading, and enculturation within a community that mind-reads' (p. 353). Children may entertain a wide variety of explanations for human behaviour initially on the basis of their direct observations of the world, only proposing a mentalistic source if they happen to grow up in a social environment where people converse freely about psychological states and in a culture that ascribes human action to such mentalistic causes as false belief.

It is difficult to adjudicate between innate neurobiological accounts and conversational experiential accounts of ToM development on the basis of evidence from autism alone. As we have seen, the empirical fact that autistic children are delayed in developing an understanding of mental states, like false belief, is undeniable. There is also firm evidence to link autistic children's ToM deficits to their slow development of language (Happé, 1995). But the direction of causality cannot be determined from these data alone. Given the behavioural definition of autism (Frith, 1989) as a triad of impairments in language, imagination and socialisation, one possibility is that these symptoms of the condition serve to block, or at least delay, the child's access to the kinds of conversation and other social experience that would nurture the development of ToM, perhaps especially because these difficulties exclude autistic children from linguistically complex and abstract family discussions about mental states like imaginary thought, false beliefs, or mistaken memories. The selective dearth of these kinds of dialogue in Tager-Flusberg's (1993) recordings of the conversation of families with autistic children is consistent with this possibility.

The development of ToM in other groups of children who, despite delays in acquiring language, are not congenitally socially aloof, or impaired in the cognitive capacity for imagination that completes the diagnostic triad of features distinguishing autism as a clinical condition (Frith 1989) helps to clarify the picture. The data we have just reviewed concerning normally intelligent

deaf children and blind children whose acquisition of language is delayed, as is ToM, provide convincing evidence for a social/linguistic account of ToM development.⁹ Dunn's (1994, 1996) observations that families where mental states are frequently discussed produce preschoolers who are precocious 'mind-readers' provide additional support for a sociolinguistic, as opposed to a nativist, explanation. Finally, we note that de Villiers and de Villiers (1999) demonstrate convincingly that by far the most significant variable in predicting success on ToM tasks is the production and comprehension of sentences containing propositional complement clauses. These data, taken in the context of the literature regarding the development of language and of ToM in deaf and blind children, show that it is precisely the familiarity with discourse in which propositions figure as objects ('Sally says/thinks that [proposition]') that enables children to propose mental states in which propositions figure as objects.

We have now reviewed a wide range of empirical data demonstrating that the strongly nativist-modularist position, which appears to be supported by data from autistic children, fails to predict the pattern of ToM development across populations of deaf and blind children. That pattern suggests a social and linguistic account of the development of ToM. The view that ToM develops on the basis of social interaction and language acquisition in turn predicts that the presence of siblings and other family members, as well as the richness of parent-child conversation, should enhance the development of ToM. Indeed, as we have seen, these predictions are confirmed. We now turn to a population that has been cited by the strong modularists and innatists as providing special difficulties for any non-modularist/non-innatist proposal—children with Williams syndrome. We will argue that, far from posing a problem for our account, the pattern of ToM development observed in Williams syndrome accords exactly with our predictions.

6. Does Williams' Syndrome Present a Special Problem?

According to Karmiloff-Smith, Tyler, Voice, Sims, Udwin, Howlin and Davies (1998) 'Williams' Syndrome is a rare contiguous gene disorder ... caused by a hemizygous submicroscopic deletion at chromosome 7q11.23 including both the elastin gene and the LIM-Kinase I gene' (p. 343). In terms of psychological functioning, Williams' syndrome is marked by an uneven

⁹ It is worth acknowledging that our model does presuppose that language acquisition, *per se*, is modular and that its structural features are largely innately determined. While this is still somewhat controversial, we place our bets with a broad nativism here (while not endorsing any specific model of language acquisition proposed within that broad paradigm). But even if this (in our view well confirmed) hypothesis concerning language acquisition were to be rejected, we nonetheless present here powerful evidence that language acquisition, however it is achieved, is causally necessary for the development of ToM. The only part of our argument hostage to the innatist, modularist theory of language acquisition is the specific account of the development of the ToM module we offer below.

profile of cognitive skills and deficits. Language is relatively well preserved in the face of serious retardation of non-linguistic mental abilities, along with deficits in adaptive behaviour that are generally so severe as to preclude the handling of money, simple decision-making, or even functioning independently during childhood or adult life. (Bellugi, Marks, Bihrlé and Sabo, 1988). The condition thus presents a theoretical opportunity to clarify interconnections among, or dissociations of language, social intelligence, and such aspects of 'cool' cognition as abstract problem solving.

Williams' syndrome children and adolescents generally do badly on many logical and spatio-mathematical tasks such as Towers of Hanoi and Piagetian quantity conservation¹⁰ (Bellugi *et al.*, 1988). Their social behaviour is often abnormal, even relative to control groups with equivalent levels of mental retardation (Einfeld, Tonge and Florio, 1997), though in contrast to individuals with autism, they generally display high levels of motivation for, and interest in, social contact with peers and adults. Williams' syndrome children are often characterised as 'over-affectionate', 'indiscriminately friendly', 'inappropriately happy' and 'preferring adult to peer company'. Einfeld *et al.* (1997) hence suggest that the syndrome involves the desire to interact and an ability to express feelings. Thus Williams' syndrome children have neither the lack of affective contact (Hobson, 1993) nor the social aloofness that are defining features of autism.

Studies of the performance of individuals with Williams syndrome on standard ToM tasks therefore present another interesting test case. Karmiloff-Smith, Klima, Bellugi, Grant and Baron-Cohen (1995) study this in a sample of 18 patients with Williams' syndrome, aged 9 to 23 years. Six experiments, are reported, each involving different ToM tasks. The first assesses abilities to use the direction of a pictured person's eye gaze to infer desire ('Which candy does he want?') or goal ('Which candy is he about to take?'). Results indicate that Williams' syndrome subjects have little difficulty with this task, performing near ceiling, and on a par with a normally developing 4-year-old control group. Experiments 2 and 3 involve first-order test of false belief understanding based on the changed locations of objects and the misleading appearances of containers. Williams' syndrome subjects display 94 percent accuracy on these tasks, showing strong evidence of a ToM capacity.

Experiments 4 and 5 test a higher level of ToM understanding, using two different second-order tests involving stories. In the first of these, a scenario is presented in which two protagonists, A and B, each possess accurate information. However, B believes that A lacks this crucial knowledge thus requiring subjects to understand B's false belief about A's state of mind. In normal development, this task is not routinely passed until the age of 7 or so. Only 31

¹⁰ This undermines the 'executive function' hypothesis for ToM, as Williams' syndrome children fail high executive function tasks but have intact ToM (Carruthers, 1996).

percent of the subjects with Williams' syndrome succeed, all with a verbal mental age of greater than 9 years. The other second-order ToM task, used in Experiment 5 with 8 of the Williams' syndrome subjects, involves a somewhat simpler narrative about a birthday present, and the attribution of true knowledge rather than false belief. In contrast to poor levels of performance on a similar task displayed in previous studies of children with autism, 88 percent of the Williams' syndrome subgroup passes this test.

Finally, Experiment 6 involves the understanding of metaphor and sarcasm. Approximately half of the Williams' syndrome group displays an accurate understanding of both of these linguistic devices by explaining the intentions of story characters who use them (e.g. a mother who tells her forgetful child 'your head is made of wood'). This frequency of perfect performance exceeds that reported for autistic subjects in previous research. Patterns of partial success also distinguish these two clinical groups. When autistic subjects pass only a single task, they often find metaphor easier than sarcasm, owing to an overriding tendency to take sarcastic statements literally. But Williams' syndrome subjects display an opposite pattern, often having difficulty with metaphoric constructions, possibly because their cognitive deficits make it difficult for them to represent abstract meaning relations.

Tager-Flusberg, Boshart and Baron-Cohen (1998) likewise find that adults with Williams' syndrome do relatively well on an advanced test of ToM involving the recognition of the emotions expressed by the eye region of the human face. Photographs of eyes are shown with the rest of the face masked, and the subjects must select a word (eg 'compassionate', 'flirtatious') to describe the emotion being depicted. The Williams' syndrome sample (aged 17 to 37 years) performs at a significantly higher level than an age-and-IQ-matched sample of retarded adults with Prader-Willi syndrome, and are above the level of success predicted by chance, though significantly below an age-matched normal adult group, prompting the conclusion that: 'adults with Williams' Syndrome are quite good at reading both simple and more complex mental state information from the eye region of the face' (p. 635).

This evidence suggests that Williams' syndrome children have fewer problems with ToM tasks than do autistic individuals, despite often comparable or lower levels of nonverbal intelligence. This disability, apparently representing the other half of the double dissociation between general intelligence and ToM, has been taken by proponents of a strongly modular account of ToM to demonstrate domain specificity in both the ToM capacity and its impairment in the case of autism. But we hope to have by now demonstrated the availability of another, equally straightforward and better-confirmed explanation for the performance of the Williams' syndrome children: The fact that language, communication and social motivation develop relatively normally in children who have Williams' syndrome explains completely their adept grasp of a ToM. On our account, after all, the development of ToM is just the product of social experience and linguistic competence, and *not* of the neuro-

biological maturation of an innate ToM module. These preconditions for ToM development are absent in autistic children and present in Williams' syndrome children. There is hence no surprise in the fact that ToM capacity is also absent in the former and present, to perhaps a surprising degree, in the latter.

On the other hand, Tager-Flusberg and Sullivan (1999) argue convincingly that the picture with respect to Williams' syndrome is more complex. They find that Williams' syndrome children perform no better than mental-age-matched controls on false belief tasks, while outperforming these control children (and autistic children) on tasks requiring the detection of affect and its use in explanation. They thus argue that social intelligence, in fact, comprises two dissociable components. They call these 'social-cognitive' and 'social-affective' and argue that only the latter is spared in Williams' syndrome. These results suggest that full-blooded ToM requires *both* the social-affective motivations and competencies that autistic children lack, but which are present in deaf children, and the linguistic skills that are absent in both populations. Again, this pattern of deficit would be predicted by our account, according to which both sets of capacities are necessary for ToM development. We now present a model consonant with this suggestion.¹¹

7. Joint Necessity for Social Intelligence and Communicative Competence/Experience for the Acquisition of Mentalising Abilities

In summary, then, the evidence from a large body of recent research suggests that children with autism are seriously delayed in acquiring a ToM. However, such delays are not limited to the autistic disorder with its triad of impairments in language, imagination and social relatedness. Instead, congenitally deaf or blind children of normal intelligence and social responsiveness are likely to experience delays in ToM development when their sensory handicaps combine with their family situations to preclude the free and open exchange of information with family members about true and false beliefs, and other mental states. However, if mentally retarded children who have Williams' syndrome develop some aspects of ToM at close to a normal rate, this could easily be explained by their spared linguistic capacity and social motivation, notwithstanding general intellectual impairment. Taken together, this body of evidence supports our conclusion that social and language development are each crucial to the development of ToM.

¹¹ Some (Frith, Morton and Leslie, 1991) suggest the possibility that a more general meta-representational capacity plays a role in the development of both ToM and of language and that that more general capacity is damaged in autism. While this hypothesis would account for the correlation between ToM deficits and deficits in metarepresentational language in this disorder, it is inconsistent with the evidence that children with autism do pass false photo tests, which also require meta-representation. Therefore, we conclude that it is far more plausible that it is specifically *language* that provides this input for a developing ToM.

De Villiers and de Villiers (1999) agree that linguistic competence is a critical precursor to the development of ToM. They demonstrate convincingly that the mastery of the syntactic structures enabling the assertion of the 'that'-clauses that underpin propositional attitude ascriptions is a necessary condition of passing ToM tasks, and indeed that such syntactic mastery reliably precedes, by a few months, success on standard ToM tasks. But they do not address sufficiency, and given the classical modularity (innateness, encapsulation, mandatoriness, neural localisation, domain specificity) of language acquisition, we see a theoretical danger here. One might be tempted to conclude from their results that ToM acquisition is simply a stage in the maturation of the language module, requiring no more from the environment than mere 'triggering' of the domain-specific processes that subserve language acquisition.¹² This would return us to a strong innatist position, albeit one in which ToM is not regarded as an autonomous module, but rather as a submodule of the language module.

To draw this conclusion would be wrong for several reasons: First, there is no evidence that high-functioning individuals with autism, who still fail standard ToM tasks, fail to master the use of 'that'-clauses in 'says that' locutions. Second, though, and of more theoretical moment, this would introduce into the language module a curious anomaly: It would privilege one domain—that of naïve psychology—as the only one for which specific *empirical* knowledge is encoded in the language module. While this is not an *a priori* impossible configuration for a mind, it would be an odd one. This observation receives additional support from the fact that no known post-acquisition trauma impairing the function of the language module also impairs ToM performance. We conclude that while de Villiers and de Villiers are correct in asserting that a certain level of linguistic maturation is *necessary* for the development of ToM, a longer developmental story needs to be told, in which a larger suite of cognitive capacities is brought into play.

Currie and Sterelny (in press) agree that the ability to attribute psychological states to ourselves and to others is closely allied with but distinct from a social intelligence module, but they reverse our order of explanation:

Mind-reading and the capacity to negotiate the social world are not the same thing, but the former seems to be necessary for the latter; people with autism are extremely restricted in their comprehension of mental states and they have comparable difficulties in negotiating the social world. And no wonder: while not every social fact is a mental fact, and not every social property is definable in mentalistic terms, our basic grip on the social world depends on our being able to see our fellows as motivated by beliefs and desires we sometimes share and sometimes do not. ... So

¹² We emphasise that de Villiers and de Villiers (1999) neither adopt nor endorse this position (personal communication).

it is no surprise that in evolutionary and developmental theories of social intelligence, the primary focus has been on our capacity for detecting, thinking about and responding to the mental states of conspecifics... Clearly such social understanding is deeply and almost exclusively mentalistic (pp. 1–2).

If we are correct, this gets things exactly backwards. It is not the case that mind-reading is a necessary condition for social negotiation. Rather, the learning of the basic skills for interacting with others that enable the social world to be negotiated is a necessary condition of learning to mind-read, together with the acquisition of language. We, like our primate kin, are wired to interact socially and to do so by perceiving and reacting to social cues relevant to cooperation, competition, nurturance and reproduction, and this is the lesson of Baron-Cohen's (1995) results concerning the early development of gaze monitoring and shared attention mechanisms. We, unlike our primate kin, are also wired for language.¹³ Put these two sets of wiring into a conducive environment, and as we learn more about our conspecifics and acquire our native languages, we learn to posit, to use in explanation, and to report, mental states. But to suggest that we could do the latter independent of the former is gratuitously to posit a rich, articulated conceptual structure without any basis for its acquisition, and without evidence for the inferential capacities it implicates. What we need instead is an account of how we can bootstrap from innately determined capacities, together with skills acquired through development and learning into the ability to represent mental states and to mind-read, and an account that explains not only normal, but pathological development.¹⁴ (See also Hobson, 1994; Whiten, 1994, 1996 and Smith, 1996.)

¹³ There are, of course, competing positions regarding the degree and nature of innateness of each of these faculties, even among those who describe themselves as nativists. One might argue at one extreme, for instance, that rich declarative knowledge structures are innate, or at the other that only very general constraints on what can be learned are innate. We suspect that different accounts are true of different domains in which innate knowledge plays a role in cognition. None of the present account depends on any specific account of innateness either in the domain of language learning or in that of social intelligence, and we defend no specific position on the relevant spectra, as the juries seem to be out.

¹⁴ As Sophian (personal communication) points out, if one were to reject the modularity and innateness of language acquisition, one could read the evidence the other way: Language acquisition depends upon the acquisition of the theory of mind. (This would have a nice Gricean ring as well.) We think that there are two reasons not to go this way: First, as we note above, the evidence for the modularity and innate determination of much of language acquisition is too overwhelming. Second, the evidence regarding the acquisition of ToM in deaf children suggests that the acquisition of language is causally necessary for ToM development, and not vice versa, since language appears to be the independently manipulated variable.

8. Why Even Modularists and Other Nativists Should Have Expected a Social Dimension: Which Module Should We Expect to be Innate?

We now return to questions about the modularity of ToM, *per se*. We are certainly hostile neither to a modular view nor to the claim that *in some sense*, ToM is innately determined. In fact we take the evidence we have surveyed and the argument we mount to support a moderate modularism and a moderate innatism. In this section we consider a range of possible positions regarding the modularity and innateness of ToM and locate our own position in that conceptual space. We begin by considering the two most extreme views: an orthodox parameter-setting model of modularity of the kind Baron-Cohen (1995) has defended; and the position that ToM simply emerges from domain-general reflective thought, a position championed by Gopnik (1996b).

The parameter setting model of modularity has been attacked decisively by Scholl and Leslie (1999): Parameter-setting only makes sense in a domain where there are a number of available parameters in an otherwise relatively uniform theory, and can only receive empirical support where an otherwise relatively uniform developmental pattern demonstrates regular clusters representing alternative developmental pathways within that larger pattern. The classic case, of course, is the domain of language acquisition where Universal Grammar (UG) and the Language Acquisition Device (LAD) represent the uniformity, and the various options for natural language grammar the parameters. But in the case of ToM, Scholl and Leslie argue, there just don't seem to be the available parameters. Nor does the structure of the task domain share that of language learning that makes parameter setting make sense. There is simply too much cross-species uniformity in the end-state of the system.¹⁵

At the other end of the spectrum is Gopnik's ultra-theory-theory. It requires too much and explains too little. It requires an implausible degree of inductive and hypothesis-testing competence in three- and four-year-olds, an implausible universal, uniform and rather selective fixation on the mind as a domain of study, and fails to explain why ToM would be selectively impaired in those like high-functioning autistics who are nonetheless theoretically competent in other domains (that is, it doesn't account for the false picture/false belief performance dissociation).

Inductive reasoning and hypothesis-testing—if they are to be distinguished as cognitive achievements from mere liability to operant and classical conditioning—are sophisticated intellectual accomplishments, often the subject of deep contestation by scientists and philosophers. They are also relatively domain-independent, and, if we correct for differences in specific empirical

¹⁵ Of course Vinden's (1996, 1999) evidence suggests that this uniformity is not complete. Still, both the scope of the uniformity in development (note how rare deviations are) and the striking uniformity in the endstate (no counterexamples) suggest that these odd cases do not support a parameterization model.

expertise in different domains, those who are good at induction and hypothesis-testing in one domain are good at it in others. If we were to accept Gopnik's picture, even if we were to grant her claim to the surprising precocity of children as scientists and theoreticians, the only way that we could account for the precocity in the development of ToM with respect to other domains of theory would be to posit a highly selective and universal fixation on the mind as an object of study despite all else that there is of interest to the child in the world. Perhaps the mind is almost universally of special interest to children, and thus attracts their special scientific talent. But then *that* fact cries out for explanation and for independent confirmation, and Gopnik gives neither. Finally, Gopnik's model has no explanation for the pathologies of ToM. Why should it be that high-functioning autistics, despite what Gopnik's model must regard as their skill at developing sophisticated theories in other domains fail to attend to the mind, or fail to work out a satisfactory theory of it? And it seems a bit crude to characterise these massive pathologies simply as an unusual choice of scientific interest.¹⁶

Our model harmonises with Karmiloff-Smith's more moderate approach to modularity as exemplified in her (1992) representational redescription (RR) model of cognitive development. Her model, according to which subsequent stages of cognitive development involve the recruitment of more sophisticated and powerful representational capacities to redescribe the representations and

¹⁶ Jake Bridge (unpublished) replies that Gopnik can indeed account for the selective impairment of ToM in autistic children by appealing to the fact that autistic children are missing innate 'starter' conceptions of mind, on the basis of which other theories are developed. And indeed Gopnik (1996a) says as much. This response, however, forces an uncomfortable dilemma for Gopnik's position: On the one hand, the conceptual structures relevant to ToM that are plausibly innate (EDD, SAM, etc) are best characterised as *capacities* or as behavioural dispositions, and not as theoretical views. For one thing, they are not used in explanation in any sense. For another (as Howe, unpublished) has noted) they are shared by apes to whom we would never want to ascribe a theory in Gopnik's sense. So it seems that what autistic children lack but normal children have is not so much an initial theory as an initial set of perceptual/attentional dispositions. But then Gopnik must say whether these are domain-specific and mandatory, determining the development of a domain-specific cognitive structure or not. If they are, her position collapses into a modularity theory, as there is then no structural difference between her account and, say, a Chomskyan account of language acquisition. But if not, there is no reason why astute young theoreticians could not bring their theoretical prowess to bear on the mentalistic phenomena about them, Gopnik (as Bridge also notes) might reply that there is a reason that theoretical prowess is not brought to bear by autistic children on the mentalistic domain: Their initial impairment might blind them to the relevant data that would determine theory construction, refutation and change. But this is a hard case for Gopnik to make, inasmuch as the relevant data are the behaviours of those around the child! The autistic child, to be sure, ignores these data, but that is what is to be explained! *Why* does the child ignore these data? If the answer is an innate impairment of some early capacity, we are back to the earlier dilemma. If instead the proponent of this strong version of the theory-theory claims that the impairment is one of theory-construction capacities, another dilemma ensues: Why just in this domain? If the capacities are domain-specific, we are back to modularity. If they are domain-general, we have no real explanation at all.

objects represented at earlier developmental levels accounts nicely for the apparent modularity of our ToM capacities, and hence for its relative domain specificity and susceptibility to selective impairment, for its characteristic pattern of development in normal cases, but also importantly for its dependence on the development of increasingly sophisticated representational capabilities. Modularity, on this view, issues sometimes from innately specified special-purpose cognitive processes, but sometimes from the development of specific special-purpose systems of representation through recruiting resources developed for other, perhaps more general, cognitive purposes. This harmonises elegantly with the results we report here and lends further support to their implication that the development of ToM is dependent upon the development of linguistic skills and experience in social interaction, and hence that ToM deficits may be at least to some extent the results rather than the causes of the other deficits found in autism. Let us explore this idea in more detail.

We suggest that the development of language and of interpersonal communicative skills is not simply the acquisition of the ability to externalise what is already internally represented. Instead, to develop language, including both its pragmatic and syntactic dimensions, is itself to develop a representational system and a set of representational and interpretative practices which themselves are preconditions of the attribution of mental states to others and to ourselves. Moreover, these linguistic and social capacities provide the epistemological entrée into the ToM. The consequent model of development would involve an interplay between internal and external representational media. This is, of course, in part to take issue with those like Baron-Cohen (1995) and Segal (1996) who argue for an innately determined and entirely *internal* (in roughly the sense that term has in internalist/externalist debates about content attribution) program of representational development. But it is not to abandon nativism completely or even implausibly. For language itself develops via an innately specified module, and humans are, if specialised for nothing else, specialised for language. The picture of representational development and of the development of the ToM suggested jointly by the work of Karmiloff-Smith (1992) and empirical studies of the acquisition of ToM by deaf and blind children, then, is that the biologically determined but socially parameterised language module is crucial to the development of the ToMM, just as it is to that of the reading module. Normal language and communicative development would then be necessary precursors to the normal development of a ToMM, and not its effect, and indeed we have seen that this is the case.

We can see how the RR model helps us here: Basic, innate modules, such as those identified by Baron-Cohen and his colleagues, are sufficient to generate a pattern of attention monitoring and behavioural coordination. Such monitoring and coordination is quite possibly demonstrated by Povinelli (1990) and others among the great apes. And this kind of monitoring and coordination is precisely what is absent in autistic children. But when language arrives on the scene, a more powerful representational medium allows a rede-

scription of these representations in the language of the propositional attitudes, and this, in conjunction with the lore developed through social interaction, permits a full-blown theory of mind to develop of a kind unique to post Sally-Ann humans.

This approach also explains why it is that the PA's, *per se*, are so central to the ToM. It is a striking fact that our folk psychology relies so heavily on attitude ascriptions, as opposed to explanation in terms of moods, emotions, mental sets, etc. Now of course one possible explanation for this reliance on the PA's would be that we just get it right: Our PA's just are the right explanans for the relevant explananda. But this is also improbable as an account of human psychology and is at least controversial. Various things beside PA's cause behaviour, including emotion, mood, and mental illness. And scientific psychology has posited a wide range of explanans, of varying kinds. If children were just doing unconstrained cognitive science, we would expect a greater diversity of theoretical constructs to emerge. If, on the other hand, our ToM rides piggyback on language as a representational system we would *also* expect that inner states directed on propositions would be those we would posit, and that those who never quite come to grips with the structure of language would never quite be able to use it in this theoretically sophisticated manner. This is because the PA's take as their objects *propositions*, and it is natural that familiarity with *overt* attitudes (e.g. sayings) toward propositions would lead to the positing of *inner* attitudes towards objects of the same kind, *viz* the PA's.¹⁷

At this point we can explain the dissociation of the false photo from the false belief task in the language-impaired in a more general way. False photos, while they encode information, encode it *pictorially*, as opposed to *propositionally*. If there is no impairment of the understanding of cameras and of imagistic representation, we would expect no deficit on the false photo task. But if there is an impairment of language processing, we would expect that a set of theoretical entities that can only be modelled on a linguistic analogy and to which we have access only through linguistically informed introspection and third-person speculation, would not be available in explanation.

We hence propose that language is epistemically prior to the PA's, and hence causally necessary for the emergence of a ToM. But this means neither that mentalising is non-modular nor that it does not have other crucial causal preconditions. Among these other causal preconditions we might well expect

¹⁷ By 'riding piggyback' we mean this: As de Villiers and de Villiers (1999) have shown, the development of competence in the use of constructions involving sentential complement clauses is a precondition of passing ToM tasks and predicts success better than any other variable. Karmiloff-Smith (1992) presents a plausible model (the Representational Re-description theory) according to which much cognitive development is driven by the re-description of representations operative at one level of processing so that those representations can be used as data in higher level processing and as templates for the development of new representations. We propose that just such a process is necessary (but not sufficient) for the development of ToM.

to find such things as the Eye-Direction Detector (EDD) and Shared-Attention Mechanism (SAM) identified by Baron-Cohen and his collaborators as well as Level I representational mechanisms of the kind identified by Karmiloff-Smith (1992).¹⁸ That is, we should be thinking about the development of the ToMM as involving the progressive recruitment of a suite of processes and capacities, including the monitoring of the behaviour of conspecifics and social processes as well as linguistic capabilities. To the extent that these are somewhat independent and perhaps dissociable from one another we might expect to find multiple possible sources of mentalising deficits, such as deafness, blindness, only-childhood, language acquisition delay or disorder, etc.

This perspective allows us better to situate and interpret Baron-Cohen's and his colleagues' results as well as those of Povinelli and others on great apes' mentalisation. The SAM's, EDD's and basic attention monitoring skills of apes represent important precursors—modular precursors—to the development of ToMM. But they are not therefore ToMM at an early stage of development. ToMM only fully emerges when it is joined with a language faculty sufficiently matured to enable the epistemic and conceptual resources necessary to complete the picture (de Villiers and de Villiers, 1999). And this requires appropriate socialisation as well as maturation.

We can thus bring together a range of insights from a number of theorists who disagree about many of the fundamental issues in this debate. The RR hypothesis of Karmiloff-Smith, the primacy of the attention-monitoring modules discovered by Baron-Cohen and his colleagues, and the crucial role of language anticipated by Vygotsky and Sellars and defended by Peterson and Siegal and de Villiers and de Villiers, can be unified to systematise the range of puzzling groupings and dissociations that these data present. The result is a more dynamic picture of modular development than might otherwise be suspected, and one involving far more interaction between modules and between the organism and the social and linguistic environment than one might suspect. This only reinforces the suspicion that the dichotomy between natural development and enculturation is illusory in a species that is naturally cultural.

At this point we note the affinity of our approach to that of Vygotsky (1962), both in respect of the model we propose, and in respect of the general perspective we adopt towards understanding cognition. In other words, our approach is Vygotskian both in the developmental sequence we envision and in the dynamic we see driving that development. Vygotsky argues that language evolves initially out of the infant's instinctive vocal noisemaking (crying, digestive sounds etc.) into a non-symbolic social coordination device when

¹⁸ During Karmiloff-Smith's Level one of the RR model, the child 'focuses on external data to create "representational adjunctions" ...[which] neither alter existing stable representations nor are brought back into relation with them. Once new representations are stable, they are simply added, domain specifically, to the existing stock, with minimal effect on what is already stored' (1992, p. 18).

parents come to treat these noises as a form of communication. Later, 'when speech begins to serve the intellect and thoughts begin to be spoken' (1962, p. 43), overt language becomes a support for complex cognitive activity (the 'egocentric speech' phase). The final stage is when the logic of language becomes internalised as discursive thought. Vygotsky's argument is both empirical and conceptual. He emphasises, as we do, that the process of learning to think is a process of skill acquisition, and that the social environment supplies both the initial reason to acquire this skill, and the necessary supports to enable its acquisition. Only later, with sufficient linguistic mastery, is it possible to think autonomously and hence to think about thought.

We note particularly Vygotsky's insistence that the transition from infra-lingual to language user marks the transition from animal cognition to fully human thought. We agree. (See also Garfield, 1988, 2000, forthcoming.) And that transition enables the child to conceptualise, to posit, and to use in explanation the concepts implicated in ToM. It does *not* and *cannot* presuppose them. Vygotsky is tantalisingly close to developing an account of ToM *per se*. After the development of internalised speech, or discursive thought, he says:

Next comes the stage which we might call 'naive psychology', by analogy with what is called 'naive physics'. ...

This phase is very clearly defined in the speech development of the child. It is manifested by the correct use of grammatical forms and structures before the child has understood the logical operations for which they stand. The child may operate with subordinate clauses... long before he really grasps causal, conditional or temporal relations. He masters syntax of speech before syntax of thought (Vygotsky, 1962, p. 46).

de Villiers and de Villiers (1999) would agree. And so would we. But Vygotsky emphasises, as do we, that the linguistic underpinnings of ToM are only half the story:

The fact is that maturation, *per se*, is a secondary factor in the development of the most complex, unique forms of human behaviour. The development of these behaviours is characterised by complicated, qualitative transformations of one form of behaviour into another. The concept of maturation as a passive process cannot adequately describe these complex phenomena. Nevertheless...in our approaches to development we continue to use the botanical analogy in our description of child development... Recently, several psychologists have suggested that this botanical model must be abandoned.

In response to this kind of criticism, modern psychology has ascended the ladder of science by adopting zoological models as the basis for a new general approach to understanding the development of children. Once the captive of botany, child psychology is now mesmerised by zoology... (1978, 19–20).

Seen from this perspective, we have been arguing that theorists as divergent as Gopnik and Meltzoff on the one hand, and Baron-Cohen and Leslie on the other, are mired in a botanical model of development. Smith and Currie and Sterelny hold the zoological view. We, with Vygotsky, urge a more social approach. Vygotsky puts the necessity of this added dimension this way:

Thus, with the help of speech children, unlike apes, acquire the capacity to be both the subjects and objects of their own behaviour (Vygotsky, 1978, p. 26).

From the very first days of the child's development his activities acquire a meaning of their own in a system of social behaviour and, being directed towards a definite purpose, are refracted through the prism of the child's environment. The path from object to child and from child to object passes through another person. This complex human structure is the product of a developmental process deeply rooted in the links between individual and social history (*Ibid.*, p. 30).

It is hence only in the context of a social matrix, Vygotsky argues, that we can make sense of human cognitive development. We, of course, agree. ToM has its ontogenetic origins not only in internal developmental dynamics, but also in social interactions. The attribution of propositional attitudes and their use in explanation is initially achieved by the child as part of a social process, and only later is internalised. And as Peterson and Siegal (1995) have shown, there is a zone or proximal development wherein children can pass ToM tasks with appropriate conversational support, but where they fail if that support is absent. Only after mastery is achieved in the social context, these data suggest, can the child proceed alone.

For this reason, we think, it makes good conceptual, psychological and evolutionary sense to say that normal pre-ToM children, *pace* Currie and Sterelny (in press), are capable of acquiring and displaying social skills, subserved by a social intelligence module, despite lacking the concepts comprised by ToM. On our analysis, children at this stage are able to perceive a wide variety of socially meaningful objects and properties in their social environments. This range of perceptibles then expands considerably with the advent of language and the theoretical resources it scaffolds. Impairment of the social intelligence faculty should then, we argue, be seen as a kind of *perceptual* impairment, making the information available to the normal child opaque to the autistic child, hence precluding ToM development.¹⁹ Impairment of the linguistic faculty alone would leave social skills, including innate social perceptual skills intact, but would make it impossible for those skills to develop the linguistic

¹⁹ 'Mindblindness' is hence, ironically, *exactly* the word needed to describe the deficits present in autism, as it refers to a perceptual impairment, *as opposed to* a theoretical deficit

mediation permitting the construction of a ToMM, issuing not in a perceptual, but a theoretical impairment. Vygotsky puts the point this way:

The child begins to perceive the world not only through his eyes, but also through his speech. As a result, the immediacy of 'natural' perception is supplanted by a complex mediated process; as such, speech becomes an essential part of the child's cognitive development (Vygotsky, 1978, p. 32).

There is a significant difference between perceptual and theoretical impairments in this domain, and either kind of impairment will lead to substantial performance deficits.

Finally, this suggests a solution to the 'convergence problem' that neither requires the kind of strongly innatist hypothesis advanced, for instance, by Baron-Cohen nor the hypothesis that children independently converge on the one true theory of mind advanced by Gopnik. The problem is this: There is striking cross-cultural uniformity on the theory of mind on which children converge, despite cultural diversity regarding surrounding metaphysics, epistemology, cosmology, etc.²⁰ There seem to be two plausible explanations: First, one might argue that this, like the uniformity in the development of depth perception or the uniformity in universal grammar is explained by an innate, fully determined, special purpose unitary cognitive module. Second, one might suggest that the facts are simply so obvious, and kids so smart that everybody figures out how the mind works by the age of five. We have argued that both are erroneous. Our broadly Sellarsian-Vygotskian framework suggests an intermediate explanation: normal children converge on a single theory of mind because of the uniformity of a number of independent modules coupled with the consequent relative uniformity (that is, uniformity in the relevant respects) of social interaction patterns across cultures. Convergence is hence neither surprising, nor simply data-driven, nor the consequence of the unfolding of a single capacity.

9. Conclusion

We set out to answer the following clusters of questions about development of ToM:

- (1) What is the nature of that knowledge and just how is it represented?

²⁰ Once again, we note that despite the variation Vinden demonstrates in trajectory towards ToM, the endstate appears uniform. The variation in trajectory, we speculate, is to be explained by either specific features of the relevant languages (probably involving the way they handle complementation) or peculiarities of the social structures and interaction patterns or the explicit folk psychology of those cultures.

- Is it a full-blown theory or is it instead more like an assemblage of skills? Are we asking about explicitly articulated lore or implicit cognitive capacities?
- (2) Is the knowledge in question modularised or is it highly interwoven with the rest of our knowledge and cognitive capacities? And if modularised, is it an acquired or an innate module? If acquired, how? What does 'modularity' in this context really mean?
 - (3) What is the relationship between ToM knowledge and the mastery of language and social skills? Is ToM a necessary condition of the development of social intelligence, cooperation and mentalistic language, and the causal condition of these cognitive developments? Or is it the consequence of the development of specific social and/or linguistic competencies?
 - (4) Is the acquisition mechanism for ToM essentially social and experiential or is it neurobiologically maturational, driven by an innate dynamic?

To the first cluster our answer is programmatic. We suggest that, in virtue of the way in which and the age at which ToM is typically acquired it would be at least gratuitous to characterise this knowledge as the explicit representation of an articulated theory. Here, then, we part company decisively with Gopnik and her colleagues. Rather it seems to us that ToM is best conceived as a complex of interpersonal and linguistic skills involving, *inter alia*, the skill of metarepresentation and inference.

In response to the second question, we argue that ToM is indeed modular, but weakly so. It is neither encapsulated nor innately determined. It is, on the other hand, underwritten by domain specific, fast and mandatory processes which themselves ride piggyback on more strongly modular processes.

Most of our argument has been directed to the third and fourth questions, which constitute the heart of the ToM debates. We have argued that ToM, far from being an innately specified, free-standing module, is tightly interwoven with linguistic and social skills and knowledge. The development of social intelligence and language are no doubt more classically modular and more innately specified. And it makes a great deal of sense to say that part of our innate endowment is hence a higher-order disposition to acquire ToM in virtue of being endowed with these more fundamental cognitive capacities and in virtue of the social and linguistic environment in which they inevitably unfold.

The consequences of this position for epistemology, the philosophy of mind and for methodology in cognitive science are immediate and significant: Our self-knowledge and our knowledge of other minds, as Sellars argued, are of a piece, and are socially and linguistically mediated. Neither is possible without language, and each depends as well on non-linguistic social practices. Despite the fact that this knowledge is so mediated and is gradually acquired along

with the development of language and social experience, however, its outlines are partially innately determined through the structure of the social intelligence and language faculties. Neither our own minds, nor those of others are self-presenting phenomena. Nonetheless, while there is an important sense in which our knowledge of the mind is theory-laden, it is not therefore *theoretical*. Our acquisition of ToM is not achieved through general-purpose theoretical activity, and our on-line ToM performance is more perceptual than theoretical. The mind we perceive is a socially-determined mind, whose genesis is only partially driven by innate dynamics, and so whose ontology cannot be wholly individualistic. Cognitive science, therefore, cannot eschew in either its theory, or in its method, attention to the social world as a determinant of the most fundamental features of human cognition. In this respect Vygotsky's central methodological insights are vindicated. We conclude that for social beings such as *Homo sapiens*, the study of even the most universal features of psychology cannot ignore the social processes that determine development. We have evolved not as isolated minds solving survival problems in a hostile environment, but rather as social organisms every facet of whose existence and development is scaffolded by a rich social environment in which a wealth of knowledge is made available. For humans, its availability is most readily assured through the medium of a shared natural language. To ignore this fact in psychological theory is to ignore what is the most distinctive characteristic of human psychology.

*Departments of Philosophy
Smith College and University of Tasmania*

*School of Psychology
University of Queensland*

*Department of Philosophy
Smith College*

References

- Astington, J. 1993: *The Child's Discovery of the Mind*. Cambridge: Harvard University Press.
- Astington, J. 1996: What is theoretical about the child's Theory of Mind?: A Vygotskian view of its development. In Carruthers and Smith, 1996.
- Astington, J. and Jenkins, J.M. 1999: A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, 35 (5), 1311–1320.
- Baier, A. 1997: *The Commons of the Mind*. Chicago: Open Court.

- Bailey, A., Phillips, W. and Rutter, M. 1996: Autism: Towards an integration of clinical, genetic, neuropsychological and neurobiological perspectives. *Journal of Child Psychology and Psychiatry*, 37 (1), 89–126.
- Baldwin, D. and Moses, L. 1994: Early understanding of referential intent and attentional focus: Evidence from language and emotion. In Lewis and Mitchell, 1994.
- Baron-Cohen, S. 1991: The Theory of Mind deficit in autism: How specific is it? *British Journal of Developmental Psychology*, 9, 301–314.
- Baron-Cohen, S. 1995: *Mindblindness*. Cambridge: MIT Press.
- Baron-Cohen, S., Leslie, A. and Frith, U. 1985: Does the autistic child have a Theory of Mind? *Cognition*, 21, 37–46.
- Baron-Cohen, S. and Swettenham, J. 1996: The relationship between SAM and ToMM: Two hypotheses. In Carruthers and Smith, 1996.
- Bellugi, U., Marks, S. Bihrlé, A. and Sabo, H. 1988: Dissociation between language and cognitive function in Williams' Syndrome. In D. Bishop and K. Mogford (eds.), *Language Development in Exceptional Circumstances*. New York: Churchill Livingstone, 177–184.
- Bennett, M. 1993: *The Development of Social Cognition: The Child as Psychologist*. New York: Guilford Press.
- Boucher, J. 1996: What could possibly explain autism? In Carruthers and Smith, 1996.
- Bridge, J. unpublished: Taking aim at Neurath's Boat. University of Massachusetts Philosophy Department, 2000.
- Bruner, J. and Feldman, C. 1993: Theories of Mind and the problem of autism. In S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (eds.), *Understanding Other Minds: Perspectives from Autism*. Oxford: Oxford University Press, 267–291.
- Butterworth, G. 1994: Theory of Mind and the facts of embodiment. In Lewis and Mitchell, 1994.
- Butterworth, G. and Jarrett, N. 1991: What minds share in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9, 55–72.
- Carey, S. and Spelke, E. 1996: Science and core knowledge. *Philosophy of Science*, 63, 515–533.
- Carruthers, P. 1996: Autism as mind-blindness: An elaboration and partial defence. In Carruthers and Smith, 1996.
- Carruthers, P. and Smith, P. 1996: *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Charrow, S. and Fletcher, M. 1974: English as a second language for deaf children. *Developmental Psychology*, 10, 463–470.
- Coltheart, M. 1999: Modularity and cognition. *Trends in Cognitive Sciences*, 3, 115–120.
- Courtin, C. and Melot, A. M. 1998: Development of Theories of Mind in deaf children. In M. Marschark (ed.), *Psychological Perspectives on Deafness*. Mahwah, NJ.: Erlbaum, 79–102.
- Currie, G. and Sterelny, K. in press: How to think about the modularity of mind-reading. forthcoming in *Philosophical Quarterly*, 2000.
- Deleau, M. 1997: L'Attribution d'états mentaux chez les enfants sourds et entendants:

- une approche du rôle de l'expérience langagière sur une théorie de l'esprit. *Bulletin de Psychologie*, 427, 48–56.
- Dennett, D. 1987: *Brainstorms*. Montgomery, VT: Bradford Books.
- de Villiers, J., Curran, L., Phillip, W. and DeMunn, H. 1997: Acquisition of the quantificational properties of mental predicates. 22nd Annual Boston University Conference on Language Development.
- de Villiers, J. and de Villiers, P. 1999: Linguistic determinism and the understanding of false beliefs. In P. Mitchell and K. Riggs (eds.), *Children's Reasoning and the Mind*. New York: Psychology Press.
- Dunn, J. 1994: Changing minds and changing relationships. In C. Lewis and P. Mitchell (eds.), *Origins of an Understanding of Mind*. Hove, E. Sussex: Erlbaum, 297–310.
- Dunn, J. 1996. Family conversations and the development of social understanding. In B. Bernstein and J. Brannen (eds.), *Children, Research and Policy: Essays for Barbara Tizard*. Washington: Taylor and Francis. pp 81–95.
- Dunn, J., Brown, J., Slomkowski, C., Tesla, C. and Youngblade, L. 1991. Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, 62, 1352–1366.
- Einfeld, S., Tonge, B. and Florio, T. 1997: Behavioral and emotional disturbance in individuals with Williams' Syndrome. *American Journal of Mental Retardation*, 102, 45–53.
- Fraiberg, S. 1977: *Insights from the Blind*. London: Souvenir.
- Frith, U. 1989: *Autism: Explaining the Enigma*. Oxford: Oxford University Press.
- Frith, U., and Happé, F. 1994: Theory of Mind and self-consciousness. What is it like to be autistic? *Mind & Language*, 14, 1–22.
- 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.
- Frith, U., Happé, F. and Siddons, F. 1994: Autism and the Theory of Mind in everyday life. *Social Development*, 3, 108–123.
- Gale, E., de Villiers, P. and de Villiers, J. 1996: Language and Theory of Mind in oral deaf children. *Proceedings of the Boston University Conference on Language Development*. Boston: Cascadilla Press.
- Fodor, J. 1983: *The Modularity of Mind*. Cambridge: MIT Press.
- Garfield, J. 1987: *Modularity in Knowledge Representation and Natural Language Understanding*. Cambridge: MIT Press.
- Garfield, J. 1989: The Myth of Jones and the mirror of nature: reflections on introspection. *Philosophy and Phenomenological Research*, 50, no 1, 1–26.
- Garfield, J. 1994: Modularity. In S. Guttenplan (ed.), *The Blackwell's Companion to the Mind*, London: Basil Blackwell.
- Garfield, J. 1998: Comments on Peterson: Language, modularity and the development of the Theory of Mind module. ARC Workshop on Modularity, Cognitive Simulation and the Theory Theory, Hobart.
- Garfield, J. 2000: Thought as language: A metaphor too far. In *Protosociology*.
- Garfield, J. forthcoming: The Meanings of 'meaning' and 'meaning': Dimensions of the sciences of mind. Forthcoming in *Philosophical Psychology*.

- Gibson, J.J. 1979: *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gopnik, A. 1996a: Theories and modules: Creation myths, developmental realities and Neurath's Boat. In Carruthers and Smith, 1996.
- Gopnik, A. 1996b. The scientist as child. *Philosophy of Science*, 63, 485–514.
- Gopnik, A. and Meltzoff, A. 1997: *Words, Thoughts and Theories*. Cambridge: MIT Press.
- Gopnik, A. and Slaughter, V. 1991: Young children's understanding of changes in their mental states. *Child Development*, 62, 98–110.
- Gopnik, A., Slaughter, V. and Meltzoff, A. 1994: Changing your views: How understanding visual perception can lead to a new theory of Mind. In Lewis and Mitchell 1994.
- Happé, F. 1995: The role of age and verbal ability in the ToM performance of subjects with autism. *Child Development*, 66, 843–55.
- Happé, F. and Frith, U. 1994: Theory of Mind and self consciousness. *Mind & Language*, 14, 1–22.
- Harris, P. 1996: Desires, beliefs and language. In Carruthers and Smith, 1996.
- Haugeland, J. 1998: Mind embodied and mind embedded. In Haugeland, 1998, *Having Thought*. Cambridge: Harvard University Press.
- Hobson, R.P. 1993: *Autism and the Development of Mind*. Hove, E. Sussex: Erlbaum.
- Hobson, R.P. 1994: Perceiving attitudes, conceiving minds. In Lewis and Mitchell, 1994.
- Howe, M. unpublished: Smart behaviorists: Precursors of ToM in our primate kin. University of Massachusetts Department of Philosophy, 2000.
- 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, 70–78.
- Karmiloff-Smith, A. 1992: *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge: MIT Press.
- Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant J. and Baron-Cohen, S. 1995: Is there a social module? Language, face processing and Theory of Mind in subjects with Williams' Syndrome. *Journal of Cognitive Neuroscience*, 7, 196–208.
- Karmiloff-Smith, A., Tyler, L., Voice, K., Sims, K., Udwin, O., Howlin, P. and Davies, M. 1998: Linguistic dissociations in Williams' Syndrome. *Neuropsychologia*, 36, 343–351.
- Leekam, S.R. 1995: Children's understanding of mind. In M. Bennett, 1993.
- Leekam, S.R. and Perner, J. 1991: Do autistic children have a metarepresentational deficit? *Cognition*, 40, 203–218.
- Leslie, A. and Thaiss, L. 1992: Domain specificity and conceptual development: Neuropsychological evidence from autism. *Cognition* 43, 225–51.
- Lewis, C., Freedman, N.H., Kryiakidou, C., Maridaki-Kassotaki, K. and Berridge, D.M. 1996: Social influences on false belief access: Specific sibling influences or general apprenticeship. *Child Development*, 67, 2930–2947.
- Lewis, C. and Mitchell, P. 1994: *Children's Early Understanding of Mind: Origins and Development*. Hillsdale: Lawrence Earlbaum Associates.

- Light, P. 1993: Developing psychologies. In M. Bennett (ed.), *The Child As Psychologist: An Introduction to the Development of Social Cognition*. London: Harvester Wheatsheaf.
- Lillard, A. 1997: Other folks' theories of mind and behavior. *Psychological Science*, 8, 268–274.
- Marschark, M. 1993: *Psychological Development of Deaf Children*. New York: Oxford University Press.
- McAlpine, L.M. and Moore, C.L. 1995: The development of social understanding in children with visual impairments. *Journal of Visual Impairment and Blindness*, 89, 349–358.
- Meadow, K.P. 1975: The development of deaf children. In E.M. Hetherington (ed.), *Review of Child Development Research, Vol. 5*. Chicago: University of Chicago Press, 441–508.
- Meadow, K.P., Greenberg, M., Erting, C. and Carmichael, H. 1981: Interactions of deaf mothers and deaf preschool children: Comparisons with three other groups of deaf and hearing dyads. *American Annals of the Deaf*, 126, 454–68.
- Minter, M., Hobson, R.P. and Bishop, M. 1998: Congenital visual impairment and 'Theory of mind'. *British Journal of Developmental Psychology*, 16, 183–196.
- Mitchell, P. 1994: Realism and early conception of mind: A synthesis of phylogenetic and ontogenetic issues. In Lewis and Mitchell, 1994.
- Mitchell, P. 1997: *Introduction to Theory of Mind: Children, Autism and Apes*. London: Arnold Publishers.
- Mitchell, P., Saltmarsh, R. and Russell, H. 1996: Overly literal interpretations of speech in autism: Understanding that messages arise from minds. *Journal of Child Psychology and Psychiatry*.
- Morford, J.P. and Goldin-Meadow, S. 1997: From here and now to there and then: The development of displaced reference in Homesign and English. *Child Development*, 68, 420–435.
- Perner, J., Frith, U., Leslie, A. and Leekam, S. 1989: Exploration of the autistic child's Theory of Mind: Knowledge, belief, and communication. *Child Development*, 60, 689–700.
- Perner, J., Leekam, S. and Wimmer, H. 1987: Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125–137.
- Perner, J., Ruffinan, T. and Leekam, S. 1994: Theory of Mind is contagious: You Catch it from your sibs. *Child Development*, 65, 1228–38.
- Peterson, C. and Siegal, M. 1995: Deafness, conversation and the Theory of Mind. *Journal of Child Psychology, Child Psychiatry and Allied Disciplines*, 36, 459–474.
- Peterson, C. and Siegal, M. 1997: Psychological, physical and biological thinking in normal, autistic and deaf children. In Wellman, H. and R. Inagaki, *New Directions in Child Development*. Hillsdale: Lawrence Erlbaum Associates.
- Peterson, C. and Siegal, M. 1998: Changing focus on the representational mind. *British Journal of Developmental Psychology*, 16, 301–320.
- Peterson, C. and Siegal, M. 1999: Representing inner worlds: Theory of Mind in autistic, deaf and normal hearing children. *Psychological Science*, 10, 126–129.

- Peterson, C. and Siegal, M. 2000: Insights into Theory of Mind from deafness and autism. *Mind & Language*, 15, 123–145.
- Peterson, C., Webb, J. and Peterson, J. 2000: Factors influencing the development of a Theory of Mind in blind children. *British Journal of Developmental Psychology*, 18, 431–447.
- Povinelli, D., Nelson, K. and Boysen, S. 1990: Inferences about guessing and knowing by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 104, 203–210.
- Povinelli, D., Parks, K. and Novak, M. 1991: Do rhesus monkeys (*Macaca mulatta*) attribute knowledge and ignorance to others? *Journal of Comparative Psychology*, 105, 318–325.
- Power, D. and Carty, B. 1990: Cross-cultural communication and the deaf community in Australia. In C. Hendrick and R. Holton (eds.), *Cross-Cultural Communication and Professional Education*. Adelaide: Flinders University Centre for Multicultural Studies.
- Premack, D. and Woodruff, G. 1978: Does the chimpanzee have a Theory of Mind? *Behavioural and Brain Sciences*, 1, 515–526.
- Ruffman, D. 1999: What autism may tell us about self-awareness: A commentary on Frith and Happé. *Mind and Language*, 14, 23–31.
- Reed, T. and Peterson, C.C. 1990: A comparative study of autistic subjects' performance at two levels of visual and cognitive perspective-taking. *Journal of Autism and Developmental Disorders*, 20, 555–67.
- Rommel, E., Bettger, J.G. and Weinberg, A. 1998: The impact of ASL on Theory of Mind. Paper presented at TIRSL 6. Washington, DC.
- Ruffman, T., Perner, J., Naito, M., Parkin, L. and Clements, W. 1998: Older (but not younger) siblings facilitate false belief understanding. *Developmental Psychology*, 34, 161–174.
- Russell, P.A., Hosie, J.A., Gray, C., Scott, C. and Hunter, N. 1998: The development of Theory of Mind in deaf children. *Journal of Child Psychology and Psychiatry*, 39, 903–910.
- Scholl, B. and Leslie, A. 1999: Modularity, development and 'Theory of Mind'. *Mind & Language*, 14, 131–153.
- Segal, G. 1996: The modularity of Theory of Mind. In Carruthers and Smith, 1996.
- Sellars, W. 1998: *Empiricism and the Philosophy of Mind*. Cambridge: Harvard University Press. Originally published in *Minnesota Studies in the Philosophy of Science* 1, H. Feigl and M. Scriven, eds., Minneapolis: University of Minnesota Press, 1956.
- Shatz, M. 1994: Theory of Mind and the development of social-linguistic intelligence in early childhood. In Lewis and Mitchell, 1994.
- Siegel, M. and Peterson, C.C. 1994: Children's Theory of Mind and the conversational territory of cognitive development. In Lewis and Mitchell, 1994.
- Smith, P. 1996: Language and the evolution of mind-reading. In Carruthers and Smith, 1996.
- Steeds, L., Rowe, K. and Dowker, A. 1997: Deaf children's understanding of beliefs and desires. *Journal of Deaf Studies and Deaf Education*, 2:3, 185–195.
- Tager-Flusberg, H. 1993: What language reveals about the understanding of minds in

- children with autism. In S. Baron-Cohen *et al.* (eds.), *Understanding Other Minds*. Oxford: Oxford University Press.
- Tager-Flusberg, H., Boshart, K. and Baron-Cohen, S. 1998: Reading the windows of the soul: Evidence of domain-specific sparing in Williams' Syndrome. *Journal of Cognitive Neuroscience*, 10, 631–639.
- Tager-Flusberg, H. and Sullivan, K. 1999: Are children with Williams Syndrome spared in the Theory of Mind? Unpublished ms. University of Massachusetts.
- Thomas, M. and Karmiloff-Smith, A. 1998: Quo vadis modularity in the 1990s. *Learning and Individual Differences*, 10, 245–250.
- Vaccari, C. and Marschark, M. 1997: Communication between parents and deaf children: Implications for social-emotional development. *Journal of Child Psychology and Psychiatry*, 38, 793–801.
- Vinden, P. 1996: Junin Quechua children's understanding of mind. *Child Development*, 67 (4), 1707–1716.
- Vinden, P. 1999: Children's understanding of mind and emotion: A multi-culture study. *Cognition and Emotion*, 13 (1), 19–48.
- Vygotsky, L. 1962: *Thought and Language*. Cambridge: MIT Press.
- Vygotsky, L. 1978: *Mind in Society: The Development of Higher Psychological Processes*, M. Cole, V. Johns-Steiner, S. Cribner and E. Souberman, eds. Cambridge: Harvard University Press.
- Wellman, H.M. 1990: *The Child's Theory of Mind*. Cambridge: MIT Press.
- Whiten, A. 1994: Grades of mindreading. In Lewis and Mitchell, 1994.
- Whiten, A. 1996: When does smart behaviour-reading become mind-reading? In Carruthers and Smith, 1996.
- Whiten, A. and Byrne, R. 1988: Tactical deception in primates. *Behavioural and Brain Sciences*, 11, 233–244.
- Wittgenstein, L. 1956: *Philosophical Investigations*. New York: MacMillan.
- Wittgenstein, L. 1969: *On Certainty*. New York: Harper.
- Zajonc, R. and Hall, E. 1986: Mining new gold from old research. *Psychology Today*, 2, 46–51.