© Blackwell Publishers Ltd. 1999, 108 Cowley Road, Oxford, OX4 1JF, UK
and 350 Main Street, Malden, MA 02148, USA.
Mind & Language, ISSN: 0268-1064
Vol. 14. No. 1 March 1999, pp 131–153.

Modularity, Development and 'Theory of Mind'

BRIAN J. SCHOLL AND ALAN M. LESLIE

Abstract: Psychologists and philosophers have recently been exploring whether the mechanisms which underlie the acquisition of 'theory of mind' (ToM) are best characterized as cognitive modules or as developing theories. In this paper, we attempt to clarify what a modular account of ToM entails, and why it is an attractive type of explanation. Intuitions and arguments in this debate often turn on the role of *development*: traditional research on ToM focuses on various developmental sequences, whereas cognitive modules are thought to be static and 'anti-developmental'. We suggest that this mistaken view relies on an overly limited notion of modularity, and we explore how ToM might be grounded in a cognitive module and yet still afford development. Modules must 'come on-line', and even fully developed modules may still develop *internally*, based on their constrained input. We make these points concrete by focusing on a recent proposal to capture the development of ToM in a module via *parameterization*.

1. Introduction

The currency of our mental lives consists largely of propositional attitudes, even when we are interpreting the behaviours of others. If you see a person running to catch up with a just-departing train, for example, you interpret the person as an intentional agent, who *believes* that there is a just-departing train, and who *wants* to get on it. It has been suggested that this capacity—termed a 'theory of mind' (ToM)—arises from an innate, encapsulated, and domain-specific part of the cognitive architecture, in short a *module* (e.g. Baron-Cohen, 1994, 1995; Leslie, 1987, 1991, 1994a, 1994b; Leslie and Happé, 1989; Leslie and Thaiss, 1992; Leslie and, Roth 1993). Several recent authors, in contrast, have argued against the modularity view, often when making the case that ToM is better explained by appeal to developing *theories* (e.g.

We would like to thank Cheryl Browne, Tim German, Ron Mallon, Adam Morton, Shaun Nichols, Richard Samuels, Gabriel Segal, Stephen Stich, Helen Tager-Flusberg, Polly Tremoulet, and especially Zenon Pylyshyn for stimulating conversation, and for helpful comments on earlier drafts. An earlier version of this paper was read at the 1998 annual meeting of the *Society for Philosophy and Psychology*.
Address for correspondence: Brian J. Scholl, Rutgers Center for Cognitive Science, Psychology Building Addition, Busch Campus, Rutgers University–New Brunswick, 152 Frelinghuysen Road, Piscataway, NJ, 08854–8020, USA.
Email: scholl@ruccs.rutgers.edu.

Bartsch and Wellman, 1995; Gopnik, 1996; Gopnik and Meltzoff, 1997; Gopnik and Wellman, 1992, 1994; Perner, 1991; Wellman, 1990).

In this paper we attempt to clarify what a modular account of ToM entails, and why it is an appealing type of explanation. We focus on the following question: What *kind* of cognitive module would be required in order to explain ToM? In particular, what kind of development would such a module need to undergo in order to account for the relevant data? Intuitions and arguments in the dispute between modules and theories often turn on the role of development: ToM appears to develop, whereas cognitive modules are thought to be static and 'anti-developmental'. This picture, we will argue, is mistaken: modularity and development, far from being antithetical to each other, can actually be intimately related. We suggest that ToM may be grounded in a cognitive module and yet still afford development.

We make these points concrete by focusing on a recent proposal by Gabriel Segal (1996) to capture the development of ToM in a module via *parameterization*. Segal distinguishes between *synchronic* modules (which reflect a static capacity), and *diachronic* modules (which attain their character from the environment via parameters, as in the case of 'Universal Grammar'), and suggests that ToM may result from the latter sort of module. This seems plausible at first blush due to the developmental nature of ToM, but we will suggest that it's not such a clear-cut issue. This discussion will refine our understanding of how development and modularity may interact in the case of ToM, and will help clarify our understanding of how a high-level representational and developing competence like ToM could in fact be grounded in a cognitive module.

We begin with a brief discussion of the basic notions involved in both ToM and modularity. Then, in section 2, we discuss developmentally motivated objections to the modularity view, and why they misconstrue the nature of modular accounts of the acquisition of ToM. In sections 3 and 4 we make this general discussion concrete by discussing a specific proposal as to how ToM might be explained in modular terms via the notion of *parameterization*. This discussion will motivate several general concluding points, concerning the relationships between ToM, development, and environmental input.

1.1 Theory of Mind

A theory of mind refers to the capacity to interpret, predict, and explain the behaviour of others in terms of their underlying mental states. It is an ability that all normal humans enjoy, and seems to manifest itself in early childhood. This capacity is inherently 'metarepresentational', in that it requires one not only to employ propositional attitudes, but to employ them *about* propositional attitudes, for example having beliefs about (others') beliefs (e.g. Leslie, 1987; Perner, 1991; Pylyshyn, 1978). Also, the capacity persists even when an embedded belief is false: someone might reason that you were running towards the train because you believed it was departing, even though it wasn't. As such, a ToM has often been thought to require its owner to have acquired the *concept* of belief.

This raises the obvious developmental question: *how* do we acquire such a concept, and thereby acquire ToM? The obvious empiricist reply—that we learn it, perhaps from our parents—is suspect, due to the extremely abstract nature of mental states. After all, it's not as if a parent can simply point and say 'See that? That's a *belief*'. Children become competent reasoners about mental states, even though they cannot see, hear, or feel them. This puzzle has caused other thinkers to opt for a more rationalist reply—that the capacity to acquire ToM has a specific innate basis (e.g. Fodor, 1992; Leslie, 1987, 1994a). And since ToM seems to apply in only a rather specific domain (viz. the cognitive properties of intentional agents), it has been proposed that the specific innate capacity takes the form of an architectural *module*.

1.2 The Modularity of Mind

We assume that the notion of modularity (Fodor, 1983) is familiar, and will only give a quick gloss here. The essence of architectural modularity is a set of restrictions on information flow. The boundaries of modules in this sense are (either one-way or two-way) informational filters: Either some of the information inside the module is not accessible outside the module (the so-called 'interlevels' of processing), or some of the information outside the module is not accessible inside the module. (It is important to keep in mind, however, that informational restrictions of this sort are always a matter of *degree*.)

By the first criterion, a central system does not have access to the processing ('interlevels') of the module, but only its output. Thus, for example, if early vision comprises a module (as argued, for example, by Pylyshyn, in press) whose output is a certain type of representation (perhaps descriptions of visual *surfaces*; Nakayama, He, and Shimojo, 1995), then no other extra-modular cognitive processes have access to the computations by which such surface-descriptions are constructed, but only to the surface-descriptions themselves. By the second criterion, the modularized processes have no access to any external processing or resources. The standard examples here are visual illusions. In the Müller-Lyer illusion, for example, you still fall prey to the mistaken percept that one line is longer than the other, even though you may have convinced yourself (say, by use of a ruler) that they are in fact the same length. This is to be explained by appeal to informational encapsulation: the processes which the module uses to construct your percept are encapsulated with respect to your beliefs. A cognitive function is architecturally modular to the degree that it is informationally 'encapsulated' in this way.

These restrictions on information flow engender a number of other 'symptoms' of modularity:

(1) Modules are *domain-specific*: they only operate on certain kinds of inputs—'specialized systems for specialized tasks'.

(2) Modules process in a *mandatory* way, such that their operation is not entirely under voluntary control.
(3) Modules are typically *fast*, perhaps due in part to the fact that they are encapsulated (needing to consult only a circumscribed data base) and mandatory (not needing to waste time determining whether or not to process incoming input).
(4) Modules offer highly constrained 'shallow' outputs, which themselves often undergo further processing down the line.
(5) Modules may often (though need not) be implemented in fixed specialized portions of neural architecture (cf. Scholl, 1997).
(6) Modules, and the abilities they support, may be selectively impaired by neurological damage.

Architectural modules have most of these symptoms, and often exhibit all of them. In employing the concept of modularity here, we second Fodor's statement that 'whenever I speak of a cognitive system as modular, I shall . . . always mean to some interesting extent' (Fodor, 1983, p. 37). Again, whenever modularity exists, it is always a matter of degree. For a more complete discussion of these symptoms, and of modularity in general, see Fodor (1983). For concise summaries of the controversy which currently surrounds modularity, see Carston (1996) and Garfield (1994).

Our application of the notion of modularity to the domain of ToM results in the following claim: *ToM has a specific innate basis*.[1] Each part of this claim is crucial. ToM has a *specific* innate basis in that the processes that determine the essential character of ToM do not apply to other cognitive domains, and can be selectively impaired. On some 'theory-theories', in contrast, the processes underlying ToM apply generally to all knowledge, and indeed are identical to the processes which underlie scientific reasoning (e.g. Gopnik and Meltzoff, 1997).[2] ToM has a specific *innate* basis in that the essential character of ToM is given as part of our genetic endowment. On some 'theory-theories', in contrast, we *learn* ToM in exactly the way that we learn about scientific theories. Finally, our claim is not that the entirety of ToM is modular, but only that ToM has a specific innate *basis*. We expand upon this last point below in our discussion of '*early ToM*', by analogy to 'early vision'.

We will assume that the normal acquisition of ToM is at least in part due to the operation of a ToM-specific architectural module. At a minimum, ToM interpretations, explanations, and predictions seem to be relatively domain-

---

1 The claim of innateness is obviously not required of the modularity view (cf. Karmiloff-Smith, 1992), and we list it here simply as part of our conception of how modularity applies to ToM.

2 We will use the term 'theory-theory' to pick out a certain class of explanations which attempt to assimilate ToM acquisition to a general account of theory acquisition in science (e.g. Gopnik and Meltzoff, 1997). Note that modular theories are counted as 'theory-theories' in some other contexts—especially in the debate between 'theory-theories' and 'simulation theories' (e.g. Leslie and German, 1995; Nichols, Stich, Leslie, and Klein, 1996).

specific (they seem to involve specialized sorts of representations and computations), fast (they typically occur without lengthy and effortful reasoning), and mandatory (you can't decide *not* to interpret lots of situations as involving intentional agents, although you can ignore the interpretation; cf. Dasser, Ulback, and Premack, 1989; Heider and Simmel, 1944). Furthermore, Leslie and his colleagues have argued that autistic disorder consists at least partially of a ToM-specific impairment (e.g. Baron-Cohen, 1995; Baron-Cohen, Leslie, and Frith, 1985; Happé, 1994; Leslie, 1987; Leslie and Roth, 1993; Leslie and Thaiss, 1992).

2. *Modularity, Development, and the Role of the Environment*

Many researchers have offered arguments as to why modular accounts are not appropriate for explaining the origin of ToM, and nearly all such arguments seem to turn on the crucial role of development. Gopnik and Wellman (1994), for example, explicitly state that (only) developmental evidence will decide between theories and modules (p. 283; also Gopnik and Meltzoff, 1997). Modularity and development, in short, are assumed to be antithetical to each other. In this section, we point out how misleading this assumption is, and suggest that modularity and development are actually entirely compatible.

The original notion of modularity, described above, concerned the degree of informational encapsulation between entirely static cognitive mechanisms or processes. At first blush, this may seem at odds with the dynamic, developing nature of a competence like ToM. ToM is traditionally cast in developmental terms, and most of the relevant experimental research has attempted to discover exactly *when* children begin to manifest different sorts of metarepresentational competence. ToM, like language, seems to undergo a process of development, where the competence itself is being (partially) learned. And this development presumably explains the experimental results in which preschoolers are progressively able to solve harder and harder ToM problems. Modular accounts, in contrast, are cast as 'developmental only in a limited sense' (Bartsch and Wellman, 1995, p. 189), or as 'nondevelopmental' (Gauvain, 1998, p. 37); Gopnik goes so far as to state that modular accounts of any representational system are 'anti-developmental' (1996, p. 174; also Gopnik and Meltzoff, 1997, p. 54).

These uncharitable characterizations seem to depend on a rather limited sense of 'modularity', which as far as we can tell is not actually employed in any modular account of the origin of a cognitive competence. First of all, modules are not typically thought to exist full-blown in the mind of a neonate, but must be triggered by the environment during maturation. This is true even of the original domains discussed by Fodor (1983): children who are denied linguistic interaction with the environment during the critical sensitive period do not ever develop a full grammar (e.g. Lenneberg, 1967; Newport, 1990), and the early visual system will not ever develop correctly if it is denied visual input during the critical sensitive period of its development (e.g. Maurer, Lewis, and Brent, 1989; Wiesel and Hubel, 1963). Even canonical modules, in other words, must 'come online', often through intermediate states, and this is surely a variety of development.

Even aside from this, though, *there is nothing in the notion of modularity which prevents even 'matured' modules from developing!* Modularity is defined only in terms of restrictions on informational access, and there is no requirement that the processes 'inside' the module do not develop. Gopnik (1996) is thus wrong that 'The encapsulation of modules means that they are indefeasible' (p. 170). The (possibly innate) restrictions on informational access constrain the information available to an internal developmental process, to be sure, but do nothing to prevent internal development itself. Certainly there is nothing in the notion of modularity that rules out a module being 'static' and 'hardwired', but then there is nothing that rules this in either (cf. Scholl, 1997). Karmiloff-Smith (1992) has even proposed that some modules may be entirely acquired. Other modularity theorists have argued that modules are *designed* to develop internally: 'The claim that there are several innate modules is the claim that there are several innate learning machines, each of which learns according to a particular logic' (Pinker, 1997, p. 33). Modules are distinguished not in that they don't develop—surely everything develops!—but in *how* they develop.[3]

Another obvious indication that modules can develop is that certain components of allegedly modular systems—e.g. the lexicon in the context of a language module, or Marr's (1982) 3D object catalogue in the context of a visual module—start out 'empty', and proceed to do nothing *but* acquire content from the environment. Such components may still be modular, however, since access to such databases (both from the 'inside out' and from the 'outside in') is highly constrained and limited to only certain other mechanisms.

2.1 Modularity and the Universality of ToM

One hallmark of the development of a modular cognitive capacity is that the end-state of the capacity is often strikingly uniform across individuals. Although the particulars of environmental interaction may affect the precise time-table with which the modular capacity manifests itself, *what* is eventually manifested is largely identical for all individuals.

As the modular account thus predicts, the acquisition of ToM is largely uniform across both individuals and cultures. The essential character of ToM a person develops does not seem to depend on the character of their environ-

---

[3] In principle, it is even possible for a module—using only its constrained input—to 'reprogramme' itself in such a way that it *alters* its restrictions on information flow over time. But notice that even this type of development would still need to be 'cognitively impenetrable'. The internal development of a module could proceed in a 'logical, rational progression', but not because of the influence of higher-level thought.

ment at all.[4] *It is at least plausible, prima facie, that we all have the same basic ToM!* Although it requires some qualification, even this extreme claim is not obviously wrong. The point is that the development of beliefs about beliefs seems remarkably uniform and stable. See Harris (1990) for a discussion of the possible cultural universality of ToM. He notes that the issue is still open: 'There is limited data to answer this question at present, but all of it is encouraging' (p. 221). The available results 'support the claim that the same theory of mind emerges universally in the young child with approximately the same time-table' (p. 222). Similarly, Segal (1996) notes that 'the few cross-cultural studies that have been done suggest that the pattern [of ToM development] is identical across the species' (pp. 152–3). All of this is in marked contrast to the uneven and culturally dependent development of many other capacities (e.g. beliefs about astronomy; Vosniadou, 1994).

A few authors have recently alleged the existence of cross-cultural differences in ToM (e.g. Lillard, 1997, 1998), but these allegations turn out not to be relevant to the modularity account. The cross-cultural differences catalogued by Lillard (1998) explicitly include differences in religious beliefs, and beliefs in phenomena such as witchcraft, magic, and karma. As such, her view of cross-cultural ToM differences pertains only to the inessential fluorescences of mature ToM competence, rather than to its essential character in early acquisition (see the comments about 'early ToM' below). The modular account of the acquisition of ToM explains the origin of the basic metarepresentational concepts (like BELIEF, PRETENCE, and DESIRE), and not necessarily how these concepts may be employed by different extramodular cognitive processes in mature individuals. In general, Lillard (1997, 1998) seems to be looking at differences in *specific beliefs*, rather than at the *concept* of belief. Even specific beliefs *about* the concept of belief are not necessarily relevant: the concept of belief could be universally grounded in a module even though most cultures do not recognize the 'modular' account in their own folk psychology! Compare the case of 'early vision': early visual processing may work identically for everybody regardless of cultural affiliation, even though different cultures have wildly different folk beliefs about how vision works (as also catalogued by Lillard, 1998). Wellman (1998) offers a similar analysis of this recent work:

> theory of mind research is pitched, fundamentally, at attempting to capture deeper core construals . . . Such an attempt may be mistaken, but it would not be derailed by the evidence Lillard (1998) amassed . . . Resulting folk psychologies could be quite different from one another worldwide; each could be quite removed and advanced beyond, although grounded in, the initial framework assumptions of young children . . . [V]ariability in adult folk psy-

[4] For our purposes here, the 'essential character' of ToM refers to three basic concepts: BELIEF, PRETENCE, and DESIRE.

> chologies, of the sorts that Lillard has cataloged, may say little about the existence, nature, or development of core folk psychological conceptions. (pp. 34–5)

It has sometimes been suggested that the universality of ToM is also predicted by the 'theory-theory', as well as by a modularity theory, since, in science, certain theories *do* converge: 'In fact, when the assumption of common initial theories and common patterns of evidence, presented in the same sequence, does hold, scientists, like children, do converge on a common account of the world' (Gopnik and Meltzoff, 1997, p. 26). This, we are inclined to think, is simply false. Scientists in this situation *sometimes* converge on similar theories, but only after lengthy disagreement, debate, and other forms of communication leading to persuasion. For ToM, in contrast, there seems to be a total convergence, as described above, without children gathering to debate and persuade each other (cf. Stich and Nichols, 1998). The striking fact is that *all* normal children seem to develop the same ToM at roughly the same time—in marked contrast to the case of science.

2.2 Developmentally Motivated Objections to the Modularity of ToM

In any case, evidence about cultural universality and uniformity is at least relevant to the modularity hypothesis, since it concerns possible environmental effects on *what* develops (and not just on *when* it develops). It could be the case, for instance, that people in different cultures developed radically different basic concepts of agency, or even lacked such concepts altogether. What is less relevant is merely evidence of some kind of environmental impact: as noted above, there are several ways in which environmental interaction may affect a modular capacity. What's crucial isn't *that* the environment has an effect, but *how* it has an effect. In particular, we should always attend to whether the environment impacts the essential character of the developing capacity, or merely affects *when* various parts of this capacity manifest themselves. Again, a strong modular account of the origin of ToM predicts uniformity in the outcome of maturation, regardless of its timetable. (Even cultural *non*-uniformity could be accommodated in a modular framework by appeal to parameterization; see section 3 below. But the point here is that such moves are not necessary to account for the actual evidence. Even a very strong modularity position is entirely compatible with the current evidence, as we shall see.)

This distinction between environmental effects on *what* develops vs. (mere) effects on *when* it develops is not always fully appreciated. It is sometimes suggested, for instance, that effects of family-size or language abilities on the time-table of ToM acquisition are a problem for the modularity account, simply because they embody effects of the environment on development. Perner, Ruffman, and Leekam (1994), for instance, report that children with multiple siblings evince a more precocious understanding of false-belief situations than do children with no siblings, and that they solve standard false-belief tasks a few months earlier. (For earlier suggestions of such links between degrees of sibling interaction and performance in false-belief tasks, see Dunn, Brown, Slomkowski, Tesla, and Youngblade, 1991. For further replications and extensions of Perner's initial finding, see Jenkins and Astington, 1996, and Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, and Berridge, 1996.)

This, however, is only evidence of environmental impact on *when* the relevant capacities manifest themselves, as opposed to an impact on *what* gets manifested, and can easily be accounted for in the modularity framework by appeal to the environmental interaction necessary for triggering. Presumably, what multiple siblings offer the precocious child is simply more of the relevant sort of environmental input (i.e. social interaction) necessary to trigger and tune the maturation of the ToM module. As this modularity view predicts, the precociousness seems to be due to mere *exposure*, and not to explicit learning or teaching (Perner, Ruffman, and Leekam, 1994). And also as this view predicts, the phenomenon seems to be a general effect of exposure to more mature people (e.g. also to nearby adult kin), and not only to siblings (Lewis et al., 1996). Children with many siblings don't develop *different* ToMs; they just do it a bit earlier. So this type of evidence for environmental impact on the development of ToM is no problem at all for even a strong modularity hypothesis, which even predicts several facets of it.

Similar comments hold for many other sorts of environmental effects on ToM development; for example effects of language abilities or informational complexity factors or explicit training on when (or in what order) various false-belief tasks are mastered. There have been suggestions, for instance, that a relative lack of linguistic and conversational experience, for example in deaf children, can slow down the acquisition of ToM (e.g. Guéhéneuc and Deleau, 1997; Peterson and Siegal, 1995), and that explicit training on false-belief and related tasks can speed it up (Slaughter and Gopnik, 1996). But again, these are all environmental effects on the *time-table* of ToM acquisition, and are easily explained in the modularity framework.

Another developmentally motivated argument against the modularity of ToM is that modularity allegedly cannot explain 'errors' early in development: 'erroneous representations, which are later modified and restructured, are . . . difficult to explain on a purely modular account. Evolution might of course select for erroneous representations, the representation just has to be good enough to survive. But if the representational system is good enough, why, on a modularity account, would it be replaced in later development?' (Gopnik and Wellman, 1994, p. 284). This assumes without argument that such 'erroneous representations' (whatever exactly this means) *are* replaced later in development. But that is not required on a modularity theory. Modularity theories, remember, intend to capture only the origin of ToM, and not all of the uses to which ToM might be put in later development. So later bona fide theories might well *override* representations grounded in a ToM module without *overwriting* them (Stich and Nichols, 1998). (We return to this sort of argument below, in section 4, when we discuss the alleged existence of a 'rational progression' of theories of mind in development.)

To conclude this section, we want to stress two important points that must be kept in mind when evaluating developmentally motivated concerns about the modularity theory. The first point is that when we talk about ToM in the context of the modularity theory, we intend to capture only the origin of the basic ToM abilities, and not the full range of mature activities which may employ such abilities. It is certainly the case that these basic ToM abilities may eventually be recruited by higher cognitive processes for more complex tasks, and the resulting higher-order ToM activities may well interact (in a non-modular way) with other cognitive processes, and may not be uniform across individuals or cultures. The general point here is that, when considered as an unanalysed totality, ToM—like 'vision'—is going to be cognitively penetrable and quite unencapsulated. *The interesting question is whether there's any significant part of the capacity that is modular*. This appears to be the case with 'early vision' (Pylyshyn, in press), and may be the case with *'early ToM'*. This is the sense of ToM we intend throughout this paper.[5]

The modular theory of ToM is thus compatible with some degree of theory acquisition, but not from scratch: it requires a core architecture. As Leslie (1994b) has put it:

> If one takes a 'child-as-scientist' view and pictures the child as an ordinary everyday scientist, working hard to contribute additional phenomena to an existing theoretical framework, then of course one addresses the nature of that original framework [i.e. 'early ToM'] . . . If instead one recoils from initial structure and starts from an unconstrained, general core architecture, the child-as-scientist metaphor changes in a critical way. Now the metaphor requires one to picture the child as a great scientist, begetter of conceptual revolution and radical theory shift. This child-scientist produces her conceptual revolutions without the benefit of formal instruction, does so regularly, and in several different domains simultaneously. Her astonishingly successful and prolific early career is diminished only by the fact that all other children make essentially this same progress too, in essentially the same way, without effort, and, by and large, independently of IQ. (Leslie, 1994b, pp. 124–5)

The second related point to keep in mind is that ToM may well not be a single indivisible capacity, as we have been characterizing it thus far. It could be that ToM is a collection of multiple related capacities, each of which is cognitively realized by a distinct mechanism or process, the development of

[5] The importance of this point was made clear to us by Helen Tager-Flusberg (cf. Tager-Flusberg, Boshart, and Baron-Cohen, in press; Tager-Flusberg. Sullivan, and Boshart, 1997).

which can be (or fail to be) triggered independently. This could still afford modularity, as long as some of the capacities were not learned by induction through environmental interaction. (For an example of a theory which explicitly splits ToM into different capacities each of which is subserved by a distinct module, see Baron-Cohen, 1994, 1995.) Again, the interesting question is whether *any* significant part of ToM is modular.

3. *ToM as Grounded in a Parameterized Module*

So far in this paper we have suggested that concerns about development do not necessarily rule out a modular account of the acquisition of ToM. We will now make these general comments more concrete by focusing on some proposals as to the specific *kind* of module which may underlie the acquisition of ToM.

3.1 Synchronic and Diachronic Modularity

We first consider a recent proposal by Segal (1996) to capture the developmental nature of ToM in a module by appeal to parameterization (a strategy also employed by Stich and Nichols, 1998). Segal motivates his position by distinguishing between *synchronic* and *diachronic* intentional modules. In Segal's terms, *synchronic* modules explain a certain competence at a given time. Parts of 'early vision', for instance, may consist of innate modular machinery which is triggered by the environment (Pylyshyn, in press). In short: 'where there is something definite that we can do, we can ask if there is something definite within us that enables us to do it' (Segal, 1996, p. 142). Segal contrasts this with what he calls a *diachronic* module, which concerns the *development* of a certain competence—'the modular conception of development' (p. 147). One way to think of this is that a diachronic module is not merely a static capacity, but a device which *produces* another cognitive mechanism; think of it as a box which takes experience as input, and produces a module as output. Crucially, this experiential interaction is mediated by the diachronic module's *parameters*, which are essentially variables whose predetermined potential values can be set by experience.

Our mature syntax faculty may be a synchronic module. It seems to have a specific innate basis, embody both directions of restriction on information flow, and enjoy the whole gamut of peripheral symptoms of modularity (cf. the essays in Garfield, 1989). But it can't be fully formed from the start, since it depends on the environment for both lexical information and for the crucial parameter-settings which distinguish the grammars of different languages. For example, some grammars allow their speakers to employ null subjects in tensed sentences, while other languages don't. Tacit knowledge of this constraint (or lack thereof) is a part of every adult's language faculty, but depends essentially on the environment. This dilemma is solved by appeal to Universal Grammar (UG), which consists of innate structure, including an innate set of parameters along which languages can vary. The language acquisition device, then, has the job of sifting through the environmental input to determine the specific *settings* of the parameters. The language acquisition device, in other words, is a paradigmatic diachronic module. It interacts with the environment (along the parameterized dimensions) to produce the mature language faculty, itself a synchronic module. The principles of UG innately specify the general form of human languages, while the parameters incorporate just enough flexibility so that a child can adapt these innate principles to whatever natural language she is trying to acquire.

Segal (1996) intends this to be a sort of pragmatic distinction: whenever you're talking about a single time-slice of some modular capacity, it is a synchronic module; whenever you're talking about how it develops, it's a diachronic module. 'Development' here could involve learning by induction, setting the essential character of the capacity by appeal to parametrically mediated interaction with the environment, or even simple triggering and maturation. In this paper, we will co-opt the distinction and construe it in what we take to be a more interesting way, between actual sorts of psychological mechanisms. What makes some capacity synchronically or diachronically modular is not your perspective (i.e. whether or not you're interested in how it develops) but the nature of the underlying cognitive architecture. (For the duration of the paper, we will use the terms 'synchronic' and 'diachronic' in this 'architectural' sense.) We will argue in the next two sections that it is correct to stress parameterization as the hallmark of diachronic modularity, but that this does not apply to ToM. ToM can be more usefully construed (in our architectural sense) as a nonparameterized module, distinguished from a diachronic module not in that it doesn't develop but in how it develops.[6]

3.2 The Diachronic Motivation and the Role of Environmental Input

Is there any reason to prefer a diachronic over a nonparameterized account of the acquisition of ToM? The critical factor in this distinction concerns the role of the environment in ToM's development. In the nonparameterized case, the environment may still be needed to trigger or tune the development of the innate process. (As noted above, this is true even in early vision.) Only in the diachronic case, though, can the environment determine the essential character of the domain-specific knowledge which ends up in the resulting synchronic module (via the parameter settings). Think of it this way: In the nonparameterized case, the essential character of the competence can develop in only one way. It may not fully develop (or it may not develop at all), but if it does, its essential character is fixed. In the diachronic case,

---

[6] Note that Segal (1996) distinguishes a number of other sorts of modules, which cut across this synchronic/diachronic distinction. These further distinctions are also extremely useful, but won't be essential here, since they all end up being plausibly applicable to ToM.

however, there's more flexibility. Experience may interact with the diachronic module such that the essential character of the resulting ToM depends in part on the character of the environment.

Appealing to this kind of diachronic module to explain the acquisition of ToM may initially seem quite natural. The whole purpose of a diachronic module, after all, is to account for patterns of development, and as noted above (see section 2), development (a) is thought to be intrinsic to the nature of ToM, (b) is the traditional focus of experimental investigations of ToM, and (c) is thought to be problematic for modular views. Segal seems to accept this view of development and modularity, and attempts to get around such worries by appeal to parameterization.

How does the crucial notion of parameterization apply to ToM? Segal here focuses on the fact that around the age of 4 a majority of children first begin to pass standard 'false-belief tasks' (e.g. Baron-Cohen et al., 1985; Perner, Leekam, and Wimmer, 1987; Wimmer and Perner, 1983). Before this age, it has been argued that children have not yet acquired the concept BELIEF (e.g. Perner, 1991; Wellman, 1990; but cf. for example Freeman and Lacohée, 1995; Leslie, 1994a). In Wellman's and Gopnik's theories (Bartsch and Wellman, 1995; Gopnik and Meltzoff, 1997; Gopnik and Wellman, 1992, 1994), this shift is analogous to (if not an actual case of) theory change in science. For Perner, Baker and Hutton (1994), the shift is from a neologistic concept, PRELIEF, to the adult concept BELIEF. PRELIEF is an undifferentiated PRETENCE-BELIEF concept, which the 4-year-old abandons in favour of the adult's differentiated concepts only when she realizes that PRELIEF is inadequate to account for intentional behaviour. This, then, is where the parameterization comes in: 'According to the diachronic modularity theory, what has occurred is that there is, as it were, a switch in the diachronic module. The switch is labelled 'prelief-belief, and it moves from one setting to the other' (Segal, 1996, p. 151). Segal summarizes his diachronic theory in this way:

> The idea would be that the maturation of the psychology faculty is a cognitive process, rather like developing a theory on the basis of evidence. And, further, developing this theory is a matter of the setting of parameters in a diachronic module. Indeed one might even have much the same views about the development of scientific theories. Thus the process of conceptual change described by the developmental theory-theory, the move from a prelief theory to a belief theory, just is the setting of a parameter in the diachronic module. (1996, p. 152)

3.3 Some Problems For the Parameterization Hypothesis

The central aspect of the diachronic modularity of ToM is thus parameterization: 'The diachronic modularity thesis construes maturation of the psychology faculty as a process of setting parameters' (p. 151). This view seems to be growing in popularity; see Stich and Nichols (1998), for example, for another appeal to parameterization in the context of explaining the acquisition of ToM. They defend a modularity view against a 'theory' theory by repeatedly pointing out that a module with enough parameters effectively reduces to a theory.

This fact—that a module with enough parameters reduces to a theory—is not an advantage, in our view. As Stich and Nichols (1998) point out, there is little or no evidence in this case which could be explained by the theory-theory but not by the ultra-parameterized modularity theory. This inability to disconfirm the parameterized-modularity theory detracts from its explanatory appeal, and as such we are disinclined to pursue this strategy here. Rather, we employ a more interesting, stronger, *non*-parameterized conception of modularity. While a parameterized module may be necessary to account for possible results of certain exotic thought-experiments (see Gopnik and Meltzoff, 1997; Stich and Nichols, 1998), we argue that it isn't necessary to account for the actual evidence.[7]

We agree that parameterization is crucial to the notion of diachronic modularity. Diachronic modules combine innate structure with worldly experience, and the parameters are necessary to mediate this mixture. They allow the innate structure to be flexible enough, in just the crucial ways, so that it can be accommodated to the environment. This is why the notion of parameters is crucial to linguistic theory. Linguistic parameters are those dimensions along which the essential character of natural languages can vary. How *else* could all the innate linguistic knowledge be adapted to specific natural languages? Parameters allow the environment to influence the essential character of the resulting synchronic module; without parameters, the environment has no such input. If the environment did not influence the essential character of a grammar, then there would be no reason to postulate parameters.

The crucial point here is that parameters are only necessary—and useful—when the essential character of the end-state of the resulting basic competence cannot be fixed in advance. This is why there must exist (something like) a parameter for null subjects in tensed sentences: if this variable were set innately, then a whole class of learnable natural languages would be ruled out. Parameterization is only required when there are multiple different end-states which the competence must potentially account for. With only one end-state, all of the relevant information could simply be built into a maturing nonparameterized module. Developmental shifts may still occur—compare puberty—but these can be explained without appeal to parametric machinery for environmental interaction.

[7] By the way, we agree with Gopnik (1996) that a crucial and telling experiment would be to raise a child in a radically different environment whose denizens employed a radically different ToM, and then see if they developed the exotic ToM equally easily (as predicted by the theory-theory) or failed to do so, being constrained to develop *our* ToM (as predicted by the strong modularity theory). Needless to say, our hunch about how this would come out is different from Gopnik's.

From this perspective, there is no reason why ToM must be a mixture of innate structure and parametrically mediated environmental input. After all, exactly what information must come from the environment in order to specify the essential character of ToM? As we reviewed above in section 2, the answer appears to be *none*: there is no evidence that the essential character of ToM that a person develops depends on the character of their environment. At a minimum, this appears to be a plausible default hypothesis, given the current evidence: parameterization is simply not necessary to account for what we currently know about the universality and uniformity of the acquisition of ToM.

As discussed above, Segal suggests that the diachronic ToM module employs a 'prelief/belief' parameter. He appears to accept Perner's idea that 3-year-olds can employ an undifferentiated PRETEND-BELIEF concept, called 'PRELIEF'. This concept is then replaced at 4 years with the adult differentiated concepts. The details of this proposal are not essential for our purposes here, but in fact there are good reasons to reject the 'prelief' story, which we will briefly discuss.

Many writers wish to explain the fact that the majority of children begin to solve standard false-belief tasks around the age of 4 by appeal to the fact that before this age they do not have a BELIEF concept, and so cannot understand false belief. However, *pretence* also seems to implicate some of the general conditions for an understanding of false belief (see Fodor, 1992; Leslie, 1987), and children much younger than 4 seem to have the concept PRETENCE, as evidenced by their competence at pretend-play (e.g. Hickling, Wellman, and Gottfried, 1997). To some, this indicates that these younger children do have the concept BELIEF after all, which is not evidenced in their performance on standard false-belief tasks for other reasons (see our section 4 below). Perner wants to argue the other way: 3-year-olds do not have a BELIEF concept, but rather an undifferentiated PRETENCE/BELIEF concept called 'prelief', which allows them to understand what Perner calls 'acting-as-if', but 'without differentiation as to whether the acting-as-if is a case of pretence or a mistake' (Perner, Baker, and Hutton, 1994, p. 265). Perner must (and does) therefore predict that 3-year-olds cannot actually distinguish between pretence and false belief, and reports experiments which support this view (Perner, Baker, and Hutton, 1994).

However, other researchers have recently demonstrated that 3-year-olds can indeed distinguish between the two concepts. In an elegant demonstration, Freeman and Lacohée (1995) used a modified 'Smarties' task (Perner et al., 1987) in which 3-year-olds typically succeed. They then asked those children who succeeded in reporting their own previous (false) belief (that the box contained Smarties) whether they had really believed that the box contained Smarties, or if they had only been pretending. Most of the children replied that they had really believed that there were Smarties in the box. When the other children who had failed the false-belief task (despite the modifications to make it easier) were asked whether they had really believed that the box contained a pencil, or if they had only been pretending, the majority replied that they had only been pretending! Although drawing the distinction between pretending and falsely believing is not a trivial task for the 3-year-old, these results demonstrate that even 3-year-olds possess both concepts.

Even ignoring this evidence, however, the 'prelief/belief' distinction does not make for a good parameter. For present purposes it essentially amounts to a 'no-belief/belief' distinction. And *this* difference surely does not require a diachronic module and parameters: suppose rather that the concept of belief is simply housed in a nonparameterized module which matures at a rate such that it 'comes online' around age 4. The 'prelief/belief' distinction does not characterize multiple possible end-states (as do linguistic parameters), but merely a developmental sequence, which by itself does not require parameterization. Parameterizing ToM is no more necessary than parameterizing puberty.

The essential character of ToM can thus be fixed without consulting the environment (even though the environment may still be needed to trigger or tune the basic competence!). Which means that parameters aren't necessary. Which means that a diachronic module (in our architectural sense) isn't necessary. Diachronic modules are an elegant design solution: they allow complicated, perhaps unlearnable cognitive processes to be offloaded to genetically determined mechanisms, while still being able to account for (highly constrained types and degrees of) environmental variability and uncertainty. Nonparameterized modules do much the same thing, except that they have no mechanism by which to consult the environment for content. Nonparameterized modules are thus a more appropriate design solution when the environment doesn't *have* to be consulted to determine the essential character of a competence.

4. *ToM as Grounded in a Nonparameterized Module*

Segal seems to recognize that environmental input may not determine the essential character of ToM. At one point he comments that the prelief/belief parameter could be switched 'endogenously, by an internal clock or some such' (1996, p. 151). But to abandon environmental determination in this way is to give up the motivation for parameterization in the first place. Why even entertain the notion of PRELIEF? Why not assume that the essential character of ToM is realized in a nonparameterized competence which is there all along?

As we pointed out earlier, the reason that many researchers are not inclined to follow this logic seems to be the pervading sense of cognitive *development* that emerges from much of the relevant empirical research. Segal is trying to soothe such intuitions, by designing his diachronic story to capture 'the modular conception of development' (1996, p. 147). If we think that the essential character of ToM is realized in a nonparameterized module, how can we explain the striking development between 3-year-olds and 4-year-olds in performance on standard false-belief tasks?

While our purpose in this paper is only to discuss modular accounts of ToM in general, these sorts of questions have already begun to be addressed by a specific theory which does assume continuity in the essential character of ToM. To illustrate the contrasting view, then—where ToM is characterized as being rooted in a nonparameterized module—we'll briefly discuss the *ToMM* hypothesis (e.g. Leslie, 1987, 1994a, 1994b; Leslie and Roth, 1993; Leslie and Thaiss, 1992; Roth and Leslie, 1998).

ToMM—the *Theory-of-Mind-Mechanism*—is essentially a module which spontaneously and post-perceptually attends to behaviours and infers (i.e. computes) the mental states which contributed to them. It interprets situations as involving intentional agents, who are represented as holding attitudes to the truth of propositions. In short, ToMM is the innate metarepresentational basis which imparts the essential character of ToM. It incorporates innate notions/concepts of propositional attitudes such as BELIEF and PRETENCE, and makes them available to a child before general problem-solving resources have fully developed. As a result, ToMM will provide the child with early intentional insight into the behaviours of others. A specific impairment of ToMM is thought to play a key role in autism (Baron-Cohen, 1995; Baron-Cohen et al., 1985; Leslie, 1991; Leslie and Roth, 1993; Leslie and Thaiss, 1992).

How does this relate to parameters and development? Leslie explains the experimental data by focusing on the crucial distinction between competence and performance. Leslie and his colleagues (Leslie, 1994a; Leslie and Roth, 1993; Leslie and Thaiss, 1992; Roth and Leslie, 1998) argue that the 3-year-old does indeed have a metarepresentational notion of BELIEF which is simply obscured by performance limitations. Specifically, they posit the existence of Selection Processing (SP), which may be non-modular, and may *not* be ToM-specific. The standard false-belief task, then, places (at least) two demands on the 3-year-old: (a) the metarepresentation must be computed, and (b) the correct content of the belief must be selected. The former is a job for ToMM, the latter for SP. SP's job is thus essentially to *inhibit* competing possible contents for the belief (Leslie and Polizzi, 1998). This sort of inhibitory process must exist anyway, and 'SP' is thus just a name given to this general capacity.

As ToMM is the source of the child's competence, so SP is the source of the processing limitations on the child's performance. For instance, to infer the content of somebody's belief when that content is false, SP is required to select among the possible contents that ToMM makes available. ToMM always makes the current situation available as a possible and even preferred content, because (a) the current situation is a truer picture of the world, and (b) beliefs tend to be true. Thus, it is sensible to assume that beliefs correlate with the world—i.e. the way the world is *now*. Passing the false-belief task, then, requires this 'default' interpretation to be inhibited by SP, in order that the weaker false content be selected (see Leslie and Polizzi, 1998, for two detailed models of SP). Basically, then, passing the standard false-belief task requires both a conceptual competence and a performance capability, and failure might simply reflect an immature SP.

To stress the importance of this 'default inhibition', we now briefly consider another standard developmentally motivated argument against the modularity of ToM. Scientific theories (according to Gopnik) tend to change in the manner of a logically related succession, often with intermediate states. It has been alleged that the development of ToM also has such features—'a succession of conceptions of mind, each logically related to earlier conceptions and revealing several intermediate transitional and partial completions' (Gopnik and Wellman, 1994, p. 283). Modular theories, unlike, theory-theories, allegedly cannot explain this rational progression.

This argument dissipates, however, upon examination of the actual alleged successors. For Gopnik and Wellman (1994), the relevant evidence seems to be of two types: (a) children evince an understanding of mental states which (children think) do not causally interact with the world (e.g. pretence, dreams, images) before they understand those which do (e.g. belief); and (b) representational aspects of DESIRE are understood before representational aspects of BELIEF. We first note that we're unsure exactly what makes these the 'rational progressions'. (It is dangerously easy to concoct a 'just-so' story in such circumstances!) But, in any case, the particular modular theory now being discussed (i.e. ToMM/SP) also easily explains these patterns of development. Dreams, images, pretences, and desires—unlike beliefs—do *not* tend to be true as a default, and so don't require inhibition. ToMM is thus able to deal with such representations just fine, and children can make use of them without taxing an immature SP, since there is nothing that requires inhibition. Note, in addition, that when you *do* tax SP in the context of DESIRES via other means (see Leslie and Polizzi, 1998), performance predictably drops, even for older children. These particular examples of 'rational sequences' of theories of ToM are thus bad candidates for casting doubt on modularity theories.

The ToMM/SP distinction upon which this all rests is more than just a fanciful idea. Leslie and Thaiss (1992) use the ToMM/SP distinction to explain an otherwise puzzling double-dissociation between the performances of normal and autistic children on certain tasks. It turns out that normal children tend to be able to solve standard false-belief tasks a little earlier than they are able to solve analogous problems about generic representation (e.g. involving out-of-date photographs). Autistic children, in contrast, perform better than 4-year-olds on the isomorphic problems of generic-representation, even while they fail the false-belief tasks that 4-year-olds pass! Explanation: the autistic children have an impaired ToMM, but a fully functioning SP (since they are older). Thus, they fail only those tasks which recruit ToMM—i.e. those tasks which involve mental states. Three-year-olds, in contrast, have a fully functioning ToMM, but an immature SP. For further related evidence, see Charman and Baron-Cohen (1992, 1995), Freeman and Lacohée (1995, especially section 18), Roth and Leslie (1998), and Saltmarsh, Mitchell, and Robinson (1995).

Here then is an example where a cognitive module accounts for the acquisition of ToM without parameterization, even though its behavioural manifestations are developmentally locked into a regular sequence. It is at least a possibility that the essential character of ToM is a synchronic body of innate knowledge, which interacts with the environment only insofar as its development is triggered and tuned. The developing performance on false-belief tasks, on this view, reflects not a developing ToMM, but a developing SP. The performance, in other words, lags behind the competence for a while. Parameterizing ToM, then, loses even its initial developmentally based appeal: a more strongly modular theory naturally accounts for both the underlying design considerations and the developmental results without appeal to parameterization.

5. *Concluding Thoughts*

A modular capacity may be acquired in at least four importantly distinct ways: (1) Although it may not fully develop if not appropriately triggered or tuned by the environment, the essential character of the innate modular capacity is fixed. (2) The essential character of the innate modular capacity may be determined by the environment in a number of highly constrained ways, via parameter setting. The modular capacity may thus reach a number of different (but quite constrained) end-states. (3) The essential character of the modular capacity may have an innate basis which is later shaped by module-*internal* development, which makes use only of that information which is 'allowed' past the module's informational boundaries. (4) Some of the properties and contents of the capacity or skill may not have an innate basis at all—the capacity may be 'cognitively penetrable' and simply learned by induction. This capacity may thus end up in a potentially infinite number of end-states, and may only later come to exhibit some of the symptoms of modular encapsulation (cf. Karmiloff-Smith, 1992). (This last view is perhaps true of many sorts of expertise, such as bird-watching or wine-tasting.)

A debate over the modularity of a cognitive capacity is typically a debate between an undifferentiated conglomeration of options (1) and (2) on the one hand, and option (4) on the other. Option 3 seems never to have been recognized in these debates. This is especially the case with regard to ToM, where the notion of modularity is used in a number of different (and often vague) ways. (See Pylyshyn, 1985, for a general discussion of this difference in the context of cognitive development.) The distinction between synchronic and diachronic modules (Segal, 1996)—corresponding (in architectural terms) to the distinction between options (1) and (2) above—has done much to clarify this debate. The finer grain of resolution afforded by this distinction gets at the essence of what it means for a high-level representational and developing capacity like ToM to be modular. We have suggested here, though, that the centrepiece of this developmentally motivated analysis of ToM—the parameterization—may be a red herring. We may be able to account for ToM in terms of a modular mechanism without appeal to parameterization.

We have argued, however, that this does not preclude parts of ToM from being *developmental*, because of the distinction between competence and performance. And, in any case, nothing precludes modules from developing internally themselves. When we attend to this distinction between competence and performance, and when we are careful to distinguish between (a) the mature ToM as a complex aggregate of capacities, and (b) the 'early ToM' out of which the mature competence develops, we can see how modularity and development—far from being antithetical to each other—can be quite intimately connected. This in turn may pave the way for applying the notion of modularity to other high-level representational and developmental capacities.

Rutgers Center for Cognitive Science
Rutgers University

References

Baron-Cohen, S. 1994: How to Build a Baby that Can Read Minds: Cognitive Mechanisms in Mindreading. *Current Psychology of Cognition*, 13, 513–52.

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

Baron-Cohen, S., Leslie, A. M. and Frith, U. 1985: Does the Autistic Child Have a 'Theory of Mind'? *Cognition*, 21, 37–46.

Bartsch, K. and Wellman, H. 1995: *Children Talk About the Mind*. Oxford University Press.

Carston, R. 1996: The Architecture of Mind: Modularity and Modularization. In D. Green et al., *Cognitive Science: An Introduction* (pp. 53–83). Cambridge, MA: Blackwell.

Charman, T. and Baron-Cohen, S. 1992: Understanding Drawings and Beliefs: A Further Test of the Metarepresentational Theory of Autism. *Journal of Child Psychology and Psychiatry*, 33, 1105–12.

Charman, T. and Baron-Cohen, S. 1995: Understanding Photos, Models, and Beliefs: A Test of the Modularity Thesis of Theory of Mind. *Cognitive Development*, 10, 287–98.

Dasser, V., Ulback, I. and Premack, D. 1989: The Perception of Intention. *Science*, 243, 365–7.

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, 1352–66.

Fodor, J. A. 1983: *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. A. 1992: A Theory of the Child's Theory of Mind. *Cognition*, 44, 283–96.

Freeman, N. and Lacohée, H. 1995: Making Explicit 3-year-olds' Implicit Competence with Their Own False Beliefs. *Cognition*, 56, 31–60.

Garfield, J. 1989: *Modularity in Knowledge-Representation and Natural-Language Understanding*. Cambridge, MA: MIT Press.

Garfield, J. 1994: Modularity. In S. Guttenplan (ed.) *A Companion to the Philosophy of Mind*. Cambridge, MA: Blackwell, 441–8.

Gauvain, M. 1998: Culture, Development, and Theory of Mind: Comment on Lillard (1998). *Psychological Bulletin*, 123, 37–42.

Guéhéneuc, K. and Deleau, M. 1997: Deaf Children Also Have a Theory of Mind . . . But Somewhat Later. Paper presented at the VIIIth European Conference of Developmental Psychology. Rennes, France: September.

Gopnik, A. 1996: Theories and Modules: Creation Myths, Developmental Realities, and Neurath's Boat. In P. Carruthers and P. Smith (eds), *Theories of Theories of Mind* (pp. 169–83). Cambridge University Press.

Gopnik, A. and Meltzoff, A. 1997: *Words, Thoughts, and Theories*. Cambridge, MA: MIT Press.

Gopnik, A. and Wellman, H. 1992: Why the Child's Theory of Mind Really Is a Theory. *Mind and Language*, 7, 145–71.

Gopnik, A. and Wellman, H. 1994: The Theory Theory. In L. Hirschfield and S. Gelman (eds). *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge University Press, 257–93.

Happé, F. 1994: *Autism: An Introduction to Psychological Theory*. Cambridge, MA: Harvard University Press.

Harris, P. 1990: The Child's Theory of Mind and its Cultural Context. In G. Butterworth and P. Bryant (eds), *Causes of Development*. New York: Harvester Wheatsheaf, 215–37.

Heider, F. and Simmel, M. 1944: An Experimental Study of Apparent Behaviour. *American Journal of Psychology*, 57, 243–59.

Hickling, A., Wellman, H. and Gottfried, G. 1997: Preschoolers' Understanding of Others' Mental Attitudes towards Pretend Happenings. *British Journal of Developmental Psychology*, 15, 339–54.

Jenkins, J. and Astington, J. W. 1996: Cognitive Factors and Family Structure Associated with Theory of Mind Development in Young Children. *Development Psychology*, 32, 70–78.

Karmiloff-Smith, A. 1992: *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Lenneberg, E. 1967. *Biological Foundations of Language*. New York: John Wiley and Sons.

Leslie, A. M. 1987: Pretense and Representation: The Origins of 'Theory of Mind'. *Psychological Review*, 94, 412–26.

Leslie, A. M. 1991: The Theory of Mind Impairment in Autism: Evidence for a Modular Mechanism of Development? In A. Whiten (ed.), *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading*. Oxford: Blackwell, 63–78.

Leslie, A. M. 1994a: Pretending and Believing: Issues in the Theory of ToMM. *Cognition*, 50, 211–38.

Leslie, A. M. 1994b: ToMM, ToBy, and Agency: Core Architecture and Domain Specificity. In L. A. Hirschfield and S. A. Gelman (eds), *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge University Press, 119–48.

Leslie, A. M. and German, T. P. 1995: Knowledge and Ability in 'Theory of Mind': A One-Eyed Overview of the Debate. In M. Davies and T. Stone (eds), *Mental Simulation: Philosophical and Psychological Essays*. Oxford: Blackwell, 123–50.

Leslie, A. M. and Happé, F. 1989: Autism and Ostensive Communication: The Relevance of Metarepresentation. *Development and Psychopathology*, 1, 205–12.

Leslie, A. M. and Polizzi, P. 1998: Inhibitory Processing in the False Belief Task: Two Conjectures. *Developmental Science*, 1, 247–54.

Leslie, A. M. and Roth, D. 1993: What Autism Teaches Us about Metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (eds), *Understanding Other Minds: Perspectives from Autism*. Oxford University Press, 83–111.

Leslie, A. M. and Thaiss, L. 1992: Domain Specificity in Conceptual Development: Neuropsychological Evidence from Autism. *Cognition*, 43, 225–51.

Lewis, C., Freeman, N., Kyriakidou, C., Maridaki-Kassotaki, K. and Berridge, D. 1996: Social Influences on False Belief Access: Specific Sibling Influences or General Apprenticeship? *Child Development*, 67, 2930–47.

Lillard, A. 1997: Other Folks' Theories of Mind and Behavior. *Psychological Science*, 8, 268–74.

Lillard, A. 1998: Ethnopsychologies: Cultural Variations in Theories of Mind. *Psychological Bulletin*, 123, 3–32.

Marr D. 1982: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W. H. Freeman.

Maurer, D., Lewis, T. and Brent, H. 1989: The Effects of Deprivation on Human Visual Development: Studies of Children Treated for Cataracts. In F. J. Morrison, C. E. Lord and D. P. Keating (eds), *Applied Developmental Psychology*. vol. 3. New York: Academic Press, 139–227.

Nakayama, K., He, Z. and Shimojo, S. 1995: Visual Surface Representation: A Critical Link Between Lower-Level and Higher-Level Vision. In S. M. Kosslyn and D. Osherson (eds), *Visual Cognition*, vol 2, *An Invitation to Cognitive Science*, 2nd edn. Cambridge, MA: MIT Press, 1–70.

Newport, E. 1990: Maturational Constraints on Language Learning. *Cognitive Science*, 14, 11–28.

Nichols, S., Stich, S., Leslie, A. and Klein, D. 1996: The Varieties of Off-Line Simulation. In P. Carruthers and P. Smith (eds), *Theories of Theories of Mind*. Cambridge University Press, 39–70.

Perner, J. 1991: *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Perner, J., Baker, S. and Hutton, D. 1994: Prelief: The Conceptual Origins of Belief and Pretense. In C. Lewis and P. Mitchell (eds), *Children's Early Understanding of Mind: Origins and Development*. Hillsdale, NJ: Erlbaum.

Perner, J., Leekam, S. and Wimmer, H. 1987: Three-year-olds' Difficulty with False Belief. *British Journal of Developmental Psychology*, 5, 125–37.

Perner, J., Ruffman, 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 and Psychiatry*, 36, 459–74.

Pinker, S. 1997: *How the Mind Works*. New York: Norton.

Pylyshyn, Z. 1978: When Is Attribution of Beliefs Justified? *Behavioral and Brain Sciences*, 1, 592–3.

Pylyshyn, Z. 1985: Plasticity and Invariance in Cognitive Development. In J. Mehler and R. Fox (eds), *Neonate Cognition: Beyond the Blooming Buzzing Confusion*. Hillsdale, NJ: Erlbaum, 403–15.

Pylyshyn, Z. in press: Is Vision Continuous with Cognition? The Case for Impenetrability of Visual Perception. *Behavioral and Brain Sciences*.

Roth, D. and Leslie, A. M. 1998: Solving Belief Problems: Toward a Task Analysis. *Cognition*, 66, 1–31.

Saltmarsh, R., Mitchell, P. and Robinson, E. 1995: Realism and Children's Early Grasp of Mental Representation: Belief-based Judgments in the State Change Task. *Cognition*, 57, 297–325.

Scholl, B. J. 1997: Neural Constraints on Cognitive Modularity? *Behavioral and Brain Sciences*, 20(4), 575–6.

Segal, G. 1996: The Modularity of Theory of Mind. In P. Carruthers and P. Smith (eds), *Theories of Theories of Mind*. Cambridge University Press, 141–57.

Slaughter, V. and Gopnik, A. 1996: Conceptual Coherence in the Child's Theory of Mind: Training Children to Understand Belief. *Child Development*, 67, 2967–88.

Stich, S. and Nichols, S. 1998: Theory Theory to the Max: A Critical Notice of Gopnik and Meltzoff's *Words, Thoughts and Theories*. *Mind and Language*, 13, 421–49.

Tager-Flusberg, H., Boshart, J. and Baron-Cohen, S. in press: Reading the Windows to the Soul: Evidence of Domain-Specific Sparing in Williams Syndrome. *Journal of Cognitive Neuroscience*.

Tager-Flusberg, H., Sullivan, K. and Boshart, J. 1997: Theory of Mind Abilities in Williams Syndrome: Is There Evidence of a Spared Cognitive Domain? Paper read at the Society for Research in Child Development, Washington DC, April 1997.

Vosniadou, S. 1994: Universal and Culture-Specific Properties of Children's Mental Models of the Earth. In L. Hirschfield & S. Gelman (eds), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 412–430). Cambridge University Press.

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

Wellman, H. 1998: Culture, Variation, and Levels of Analysis in Folk Psychologies: Comment on Lillard (1998). *Psychological Bulletin*, 123, 33–36.

Wiesel, T. and Hubel, D. 1963: Single-cell Responses in Striate Cortex of Kittens Deprived of Vision in One Eye. *Journal of Neurophysiology*, 26, 1003–1017.

Wimmer, H. and Perner, J. 1983: Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. *Cognition*, 13, 103–128.