10

How to build a baby that can read minds: Cognitive mechanisms in mindreading

Simon Baron-Cohen
University of Cambridge

The eight central claims of the paper are that: (1) Humans have evolved to be able to attribute mental states to interpret and predict action—that is, to "mindread" (Premack's claim). (2) The neurocognitive system dedicated to this function (the Mindreading System) has 4 modular components: (a) an Intentionality Detector (ID), akin to Premack's suggestion, whose function is to represent behaviour in terms of volitional states (desire and goal); (b) an Eye Direction Detector (EDD), whose function is initially to detect the presence of eye-like stimuli, and later to represent their direction as an Agent "seeing" the Self or something else; (c) a Shared Attention Mechanism (SAM), whose function is to represent if the Self and another Agent are attending to the same object or event; and (d) a Theory of Mind Mechanism (ToMM), as Leslie outlines, whose function is to represent the full range of mental states (including the epistemic ones), and to integrate mental state knowledge into a coherent and usable theory for interpreting action. (3) ID and EDD are held to process dyadic representations, SAM processes triadic representations, and ToMM processes what Leslie calls "M-Representations". (4) Whilst SAM can build triadic representations from the output of either ID or EDD, these are more easily built in the visual modality, hence the evolution of a special relationship between EDD and SAM. (5) SAM also functions to link ID with EDD, enabling eye-direction to be read in terms of volitional states. (6) When SAM outputs its triadic representations to ToMM, this triggers the latter to function. (7) A subgroup of children with autism are postulated to be impaired in SAM, whilst congenitally blind children are not. (8) In this subgroup of autism, ToMM is therefore not activated.

Correspondence should be sent to Simon Baron-Cohen, Departments of Experimental Psychology and Psychiatry, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K. (e-mail: spjtsbc@ucl.ac.uk).

In *Cahiers de Psychologie Cognitive* (Current Psychology of Cognition), 1994, 13 (5), 513–552.

> "One of the salient features of other persons in our social world is the action and focus of their eyes. The direction of another's gaze, which seems to indicate something about his or her orientation and focus of attention, may have significance for us and the organization of our own actions, particularly if we happen to be the target of these two distinctive orbs. But others are not invulnerable to the same scrutiny from us and likely will make similar adjustments depending on what they infer about our intentions, based in part on the presence or absence of our apparent attention to them."
> (Fehr & Exline, 1978, p. 225).

The idea that humans have evolved an ability to interpret and predict behaviour on the basis of mental status has a relatively recent history (Jolly, 1966; Premack & Woodruff, 1978; Humphrey, 1984; Byrne & Whiten, 1988; Cheney & Seyfarth, 1990). These authors have all argued in different ways that an ability to read behaviour in terms of mental states would confer selective advantages for survival and reproduction to those organisms who possessed this ability, in making them better able to make sense of another organism's action, and predict what it might do next.

In this article, I propose the existence of a *Mindreading System* that has evolved specifically to enable attribution of mental states to agents. In its most highly evolved form, the Mindreading System is postulated to have four modular components. Two of these have been described by others (ID: *pace* Premack, 1990; ToMM: Leslie, 1991). I will have relatively little to say about these two mechanisms, as readers can go back to the original papers by these authors for a fuller account of them. The other two mechanisms (EDD and SAM) are, I think, new, though of course they build on a large body of work which I review. My suggestion is that different neurological and phylogenetic populations may lack one or more of these four modules. In the normal human case, all four are present and functioning by 4 years of age. These four modules are "special purpose computational system(s)" (Fodor, 1983, p. 47) in what Brothers (1990) calls the "social brain".

In the main section of the paper, I review the four modules in the order in which I suggest they are active in human development. The proposed relationships between them are shown in Fig. 1, to which you can refer as I describe each mechanism. Before doing this, let me briefly defend my use of the term "module".

CRITERIA FOR MODULARITY

Fodor (1983) is the person who has given most impetus and serious consideration in modern psychology to the notion that the mind and the brain have modular organization. His modularity thesis is summarized in terms of nine tenets. Modules, Fodor argued, have:

1. Domain specificity
2. Informational encapsulation
3. Obligatory firing
4. Shallow outputs
5. Rapid speed
6. Inaccessibility to consciousness
7. A characteristic ontogenetic course
8. A dedicated neural architecture, and
9. A characteristic pattern of breakdown.

As Bates (1993) points out, the first 6 of these tenets also apply to "overlearning", whereby skills acquired through experience become automatized. The last 3 tenets are more applicable to "biological" modules. This clarifies that modules may result from both innate or acquired factors

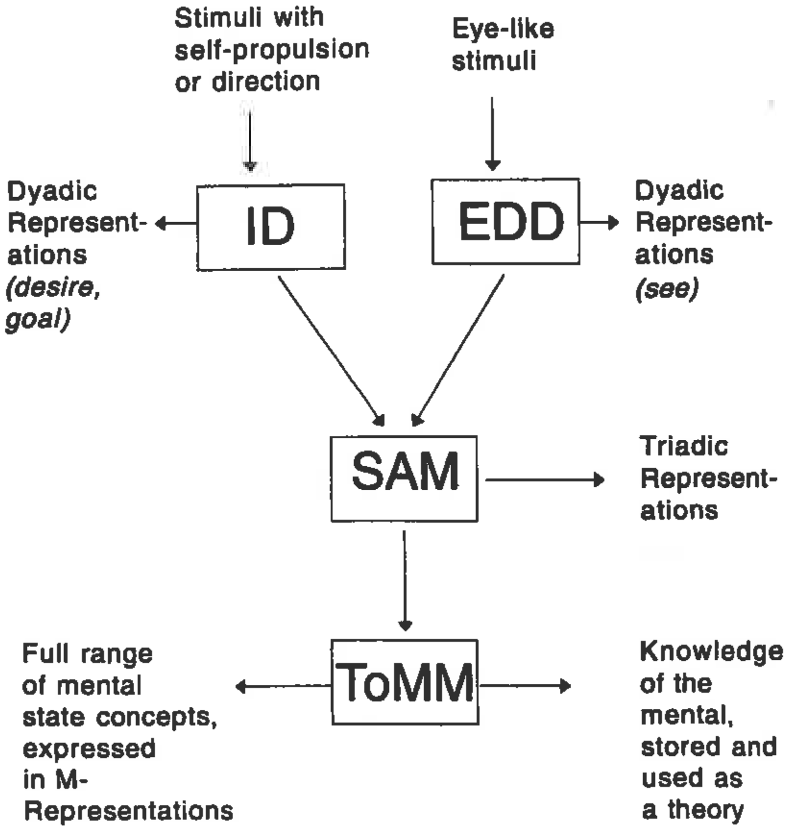


FIG. 1. The Mindreading System.

(Karmiloff-Smith, 1992). Of Fodor's 9 criteria, I reject the need for number 2 (informational encapsulation), since this seems to me to prevent the quite useful possibility that modules interact with one another in the way that I will suggest some of those in the Mindreading System do. Number 4 (shallow outputs) also seems to me to be both unnecessary and somewhat premature to include in the list, since it is an empirical question—in the case of each proposed module—as to what the output is. Number 6 (inaccessibility to consciousness) is also likely to be too strong a criterion, since different modules may vary in the extent to which their contents are accessible to introspection.

This therefore leaves 6 tenets (1, 3, 5, 7, 8, and 9). To each of the four components in the human Mindreading System, I will apply these 6 criteria, in order to see how each one stands up to this definition of a module.

FOUR COGNITIVE MECHANISMS

1. ID

ID stands for the *Intentionality Detector*. It is a primitive perceptual mechanism that I suggest is amodal, and that interprets stimuli in terms of that stimulus' goal or desire. That is, it reads stimuli as volitional. In this system, "goal" is defined as the target an action is directed towards, and "desire" is defined as a movement towards or away from a target. I intend this module to be very similar to—or coterminous with—Premack's (1990, 1991) notion of a module that is hard-wired into the human infant's visual system to detect intentionality. However, because I have modified slightly how I propose this system works, I have renamed it as ID. The suggestion is that human infants (by at least 6 months) read mental states of goal and desire into a wide range of stimuli with *direction* (e.g., a touch, a push, a jump, a shout, an arrow) or that manifest *self-propulsion* (i.e., an object that moves without an apparent external cause). This is basically what Premack also argues, although he implies a smaller set of stimuli triggers an interpretation of goal.

Applying the 6 criteria for modularity to ID, what do we find?

1.1. Domain specificity. Stimuli with direction, and self-propelling stimuli, appear to be a reasonably well-defined domain. Thus, a static circle or square has no direction, whilst a tree-branch does. Similarly, an object moving by apparent self-propulsion differs in key respects from an object that moves as a result of an external force (Mandler, 1992; Leslie, 1994).

A second approach to assessing domain specificity is to ask if the proposed system employs a unique class of representation. I will be using this approach for each of the four mechanisms I discuss in this paper. I suggest that ID builds *dyadic* representations. These specify the relation (in this case, a desire or a goal) between an Agent and something (or someone) else. They are dyadic because they code two entities in a volitional relationship. Dyadic representations can therefore take forms:

a. [Agent-Relation-Self]. Here, the relation term is bidirectional, because both elements are Agents and thus capable of an active relation with something. Examples of this form are:

[Mummy-wants-me], or
[I-want-Mummy].

b. [Agent-Relation-Object]. Here, the relation term is unidirectional, since one of the elements is not an Agent. An example of this form is:

[Mummy-wants-the cup].

c. [Agent$_1$-Relation-Agent$_2$]. Here, again, the relation term is bidirectional, since both elements are Agents. Examples of this are:

[Mummy-has goal-Daddy], or
[Daddy-wants-Mummy].

d. [Self-Relation-Object]. Here, the relation term is unidirectional. So an example of this form is:

[I-have goal-the toy].

Dyadic representations are not unique to ID, however, since I will argue that EDD (see below) also employs them. In this respect, the domain-specificity criterion is only partly fulfilled in the case of ID.

1.2. Obligatory firing. ID probably meets this criterion, since Heider and Simmel (1944) report that only one out of their whole sample of normal adult subjects did not attribute agency, goal and desire terms to moving geometric shapes in a silent film sequence, when asked to describe what they saw.

1.3. Rapid speed. No data related to speed are reported by Heider and Simmel, but the impression we have is that we do attribute desires and goals very fast. We don't need to ponder in any laboured fashion in order to compute that a stimulus might "want" something. We see the bee flying towards a colourful flower, and we 'instantly' compute that the bee *wants* to go to the flower.

1.4. Characteristic ontogenesis. The relevant human infancy studies have not progressed very far, with which to test if ID has a characteristic ontogenesis. However, these studies are underway in several centres (Gergely, Nádasdy, Csibra, & Biró, in press; Premack, 1993). Certainly older children (Dasser, Ullbaek, & Premack, 1989) readily interpret such minimal stimuli as moving geometric shapes as having goals and desires. And almost as soon as children start to speak, they refer to goals and desires (Wellman, 1990; Bartsch & Wellman, 1994).

1.5. Dedicated neural architecture. Perrett et al. (1992) report that specific cells in the Superior Temporal Sulcus of the monkey brain (the STS) fire in response to perceiving the *direction* of an Agent's head, or body-posture, or directed action (e.g., a hand opening a jar) suggesting that these cells may be detecting "goal" or "desire". However, since other cells in the STS respond to other features, the question of whether ID has a dedicated neural architecture needs more specific investigation.

1.6. Characteristic pattern of breakdown. There are reports that some forms of brain damage in human adult patients can lead to specific deficits in the recognition of agency or animacy (Farah, McMullen, & Meyer, 1991; Warrington & Shallice, 1984; Goodglass, Klein, Carey & Jones, 1966). This suggests that ID may have its own unique neurobiological basis that can be selectively damaged. Again, this needs more careful investigation, in order to establish if neurological patients have specific impairments in recognizing volitional states, versus agency, versus animacy.

In sum, ID fits some (but not all) of the criteria for modularity. I turn now to the other module that I propose is active in the human infant's Mindreading System.

2. EDD

EDD is the *Eye Direction Detector*. I suggest this system has two functions: first, to detect the presence of eyes (or eye-like stimuli). This function appears to be present in a large range of species (see below). Secondly, in the higher primates, EDD also represents eye-behaviour. Tantam (1992) hints that there must be such a mechanism, given the evidence for a primitive "social gaze response". Here I extend this important notion.

Applying the criteria for modularity to EDD, what do we find?

2.1. Domain specificity. Eye-like stimuli certainly constitute a well-defined domain. They typically occur in pairs, and often (though not always) comprise a dark, circular pattern. Human eyes also have contrast, as well as being able to move[1]. Their contrast is of both colour and luminescence, and is between the darker (iris/pupil—or *d*) region, and the white (sclera—or *w*) region. EDD in humans, I will argue, is sensitive to both the contrast and movement properties of eyes—it tracks and codes the spatial position of the *d* region relative to the *w* region. For convenience, I call these *d:w* representations, three examples of which are depicted in Fig. 2. Naturally, these

---

[1] Another observable property of the eyes is change in pupil size, independent of illumination. People clearly respond to this and are influenced by this in their judgement about attractiveness, etc. Indeed, the use of "belladonna" or atropine to dilate the pupils is correlated with judgements of beauty. See Hess and Petrovich (1978) for a review. I do not discuss pupillometry here.

illustrations only capture the static features of *d:w* representations, and not the dynamic, movement information that must also be specified in these. Whether such representations are unique to EDD, or simply make use of more general properties of the visual system, remains to be tested.

I suggest that, in the human infant, probably from about 4 months of age (Johnson and Vicera, 1993), and possibly in some of the other higher primates, EDD not only detects eye-like stimuli in an "obligatory" fashion, but goes on to build dyadic representations of their behaviour. These dyadic representations have an identical structure to those processed by ID, but whereas for ID the relation slot is filled with a term like "goal" or "desire", in the case of EDD the relation slot is filled with a term like "see". As in the case of ID, EDD's representations are dyadic because there are only two entities, connected by a relation term. And as before, EDD's dyadic representations must thus have one of four forms:

a. [Agent-Relation-Self]. Here, the relation term is bidirectional. Examples of this form are:
[Mummy-sees-me], or
[I-see-Mummy].

b. [Agent-Relation-Object]. Here, the relation term is unidirectional. An example of this form is:
[Mummy-sees-the bus].

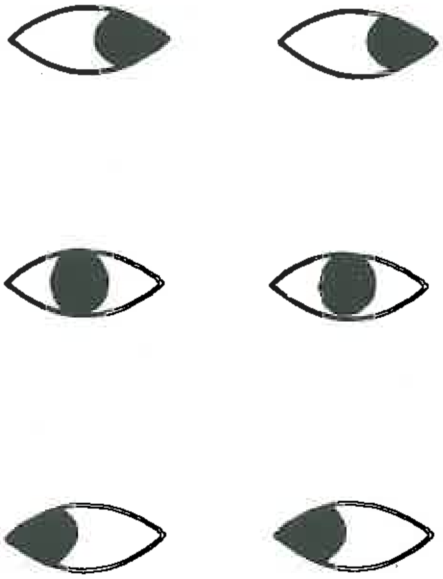


FIG. 2. Schematic illustration of some *d:w* representations.

c. [Agent$_1$-Relation-Agent$_2$]. Here, again, the relation term is bidirectional. Examples of this are:

[Mummy-sees-Daddy], or
[Daddy-sees-Mummy].

d. [Self-Relation-Object]. Here, the relation term is unidirectional. So an example of this form is:

[I-see-the house].

I have presented the dyadic representations in a 'sentence' form, but for clarity, some dyadic representations that EDD builds are also shown in Fig. 3.

In EDD, the relation term in dyadic representations is *the* crucial advance over *d:w* representations. It specifies that eyes belonging to Agent (or Self) stand in a relation to an object (or person). As with the relation term in ID's dyadic representations, I suggest that this allows the infant to represent some basic properties of intentionality (Brentano, 1984/1970), since intentionality (or "aboutness") is essentially relational[2].

2.2. Obligatory firing. It is hard not to notice people's eyes, especially when they are pointed at us (see Baron-Cohen, 1995). In this section, I consider the obligatory criterion by reviewing phylogenetic evidence for a special sensitivity to eyes and eye-direction.

Regarding eye-like stimuli, Blest (1957) reports that "eye-spots" are found on peacock tail feathers, on certain moth and butterfly wings (see Fig. 4), on some fish (e.g., Cichlid), on cobras, on some grasshoppers (e.g., *Ommatolampis perspicillata*), on some wasps (e.g., Mutillid) and beetles (e.g., Carabid), and on some pheasants, to name just a few. These eye-spots resemble vertebrate eyes. Evidence that such eye-spots are indeed readily detected comes, for example, from the finding that eye-spots have a deterrent effect on the predator. For example, certain species of moth flash the eye-spots on their wings, and this deters predators (Scaife, 1976). Similarly, Yellow Bunting birds retreat from butterflies with eye-spots, but eat the ones without eye-spots (Blest, 1957).

---

[2] The Self term in dyadic representations implies that the infant already has a localizable concept of self, distinct from another person. I assume infants already have a primitive self-concept, this having been derived from the perceptual distinction between single and dual sensations (Gallup, 1982; Perrett et al., 1990).

The relation slot in EDD's dyadic representations is filled by the term like *see*. I assume EDD can obtain knowledge about eyes seeing from the simple contingencies of closing and opening its own eyes (eyes closed produces the experience of darkness; eyes open produces the experience of light). Whilst this is initially based on its own experience, this knowledge could be generalized to an Agent by analogy with the Self.

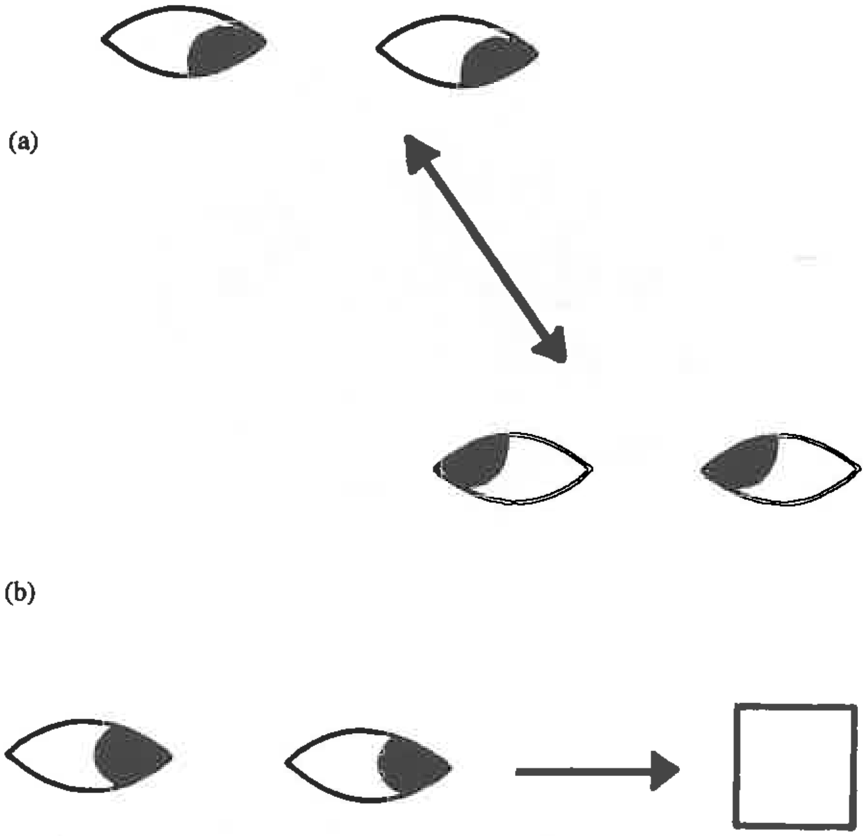


FIG. 3. Schematic illustration of some dyadic representations:
(a) [Agent-Relation-Self); (b) [Agent-Relation-Object].

All of this probably reflects natural selection taking advantage of the effect of "eyes". According to mimicry theory (Poulton, 1890), such signals may have evolved because eye-spots as stimuli cause the predator to hesitate, thus giving the moth valuable time to escape (Arduino & Gould, 1984). Whether the predator interprets the eye-spots as *eyes looking at them* is not clear: the predator's response could simply be a hard-wired response to a stimulus of this particular shape and colour. For example, Scaife (1976) found that 37 day old white leghorn chicks show most avoidance to two tracking, eye-like shapes[3].

---

[3] Note that eye-spots are not always responded to as dangerous. For example, in the male guppy fish, the eye-spot appears during *courtship*— the male turns his body sideways to present this display to the female (Argyle & Cook, 1916, p. 6).

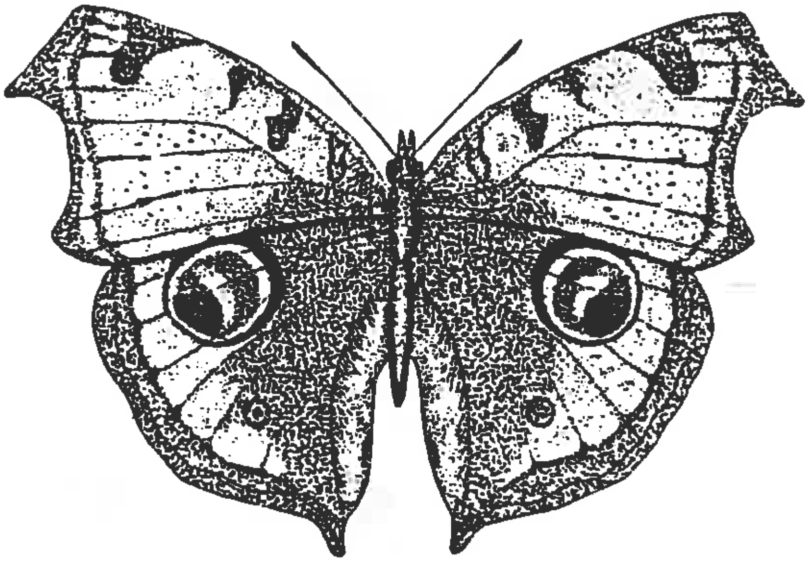


FIG. 4. Eye spots on a moth, adapted from Blest (1957).

As mentioned earlier, EDD is not solely for detecting the *presence* of eyes (though this is part of its job) but also for detecting the *direction* of eyes, since real eyes are directed at different targets in the environment. One likely primitive function of EDD is that it allows an organism to detect if another animal is aiming to attack it. Under this interpretation, eye-direction detection could be construed as a form of threat-detection.

Ristau (1990, 1991) carried out some elegant experiments with plovers, to test if these birds were sensitive to eye-direction, and whether they reacted to eyes directed at them as a threat. The birds were observed in the dunes on the beaches of Long Island, New York, where they nest. Ristau used two human intruders, one of whom looked towards the dunes, the other of whom looked towards the ocean. Each intruder walked up and down the same path, along the coastline, about 15–25 metres from the dunes. Trials began when an incubating parent plover was on her nest. Ristau found that the birds moved off and stayed off their nests for longer periods when the intruder was gazing towards the dunes than when the intruder was gazing towards the ocean. Moving away from the nest was interpreted as a sign of the parent-bird attempting to lead the intruder *away* from the nest. Ristau interpreted this as evidence that these birds are capable of detecting if an intruder is looking at their nest, and that the birds react to gaze so directed as a threat. One should note that in this study the birds had both eye-direction and head-direction available as cues.

Snakes have also been reported to be sensitive specifically to eye-direction as a cue to a potential threat (Burghardt, 1990). For example, if an intruder is about one metre from a hog-nosed snake, and looks directly at the snake, the snake will "feign death" for longer than if the intruder averts its eyes. The same is true of chickens, who also engage in *tonic immobility* for longer in the presence of a human who is staring at them than one who is not looking at them (Gallup, Cummings, & Nash, 1972). The phenomenon of tonic immobility has been documented in a range of other species, such as the lizard (*Anolis carolinensis*: Hennig, 1977), the blue crab (*Callinectes sapidus*: O'Brien & Dunlap, 1975), and ducks (Sargeant & Eberhardt, 1975). [See Arduino & Gould, 1984, for a review.]

Many animals do not react to the eyes with tonic immobility, but nevertheless react with avoidance and fear. For example, macaque monkeys look less at photographic slides of faces with eye-contact than with no eye-contact (Keating & Keating, 1982), and infant macaque monkeys show more emotional disturbance when confronted by a picture of a full face with eye-contact, compared with a picture of a face turned away to profile, with gaze averted (Mendelson, Haith, & Goldman-Rakic, 1982). Perrett and Mistlin (1990) further demonstrated that appeasement behaviours (lip-smacking and teeth-chattering) by macaque monkeys are controlled by gaze angle and head posture, in that they occur more often to a human face looking directly at the animal (from a distance of 1.5m, whether full-face or half-profile), than to a human face tilted backward (see Fig. 5).

Mutual gaze, particularly in the form of a stare, is a well-documented component of threatening displays in many non-human primates, e.g., adult male baboons (Hall & Devore, 1964), gorillas (Schaller, 1964), macaques (Altmann, 1967), and a number of other old-world monkeys and apes (van Hooff, 1962). Chance (1962) describes how struggles for dominance are often only ended with one animal averting its gaze—what he calls a "visual cut-off", possibly as a mechanism for reducing the physiological arousal produced by direct gaze (Wada, 1961; Nichols & Champness, 1971).

This array of studies might suggest the evolution of a "fear" response to eye-direction. However, in the higher primates gaze also occurs as part of grooming, greeting, and play facial expressions in old-world monkeys and apes (van Hooff, 1962). Argyle and Cook (1976), in their review of the literature, conclude that it is only in primates that gaze functions as an affiliative as well as an aggressive cue[4].

---

[4] It is of interest that in non-human primates, the sclera becomes pigmented by adulthood. Perrett and Mistlin (1990) have speculated that this itself may reflect an evolutionary adaptation. Their argument is that since (as we have seen) most primates respond to direct eye-contact as a threat, pigmentation of the sclera may allow one animal to watch another 'out of the corner of its eye" without risking overt staring eye-contact. This is because pigmentation of the sclera makes it more difficult for an animal to discern another animal's eye-direction, at least in profile direction. It may be however that in the human case it was evolutionarily adaptive to *maximize* the possibility of eye-direction being detected—in order to optimize the prosocial signal that eye-direction can convey. This would account for why in the human case, the sclera remains white throughout life.

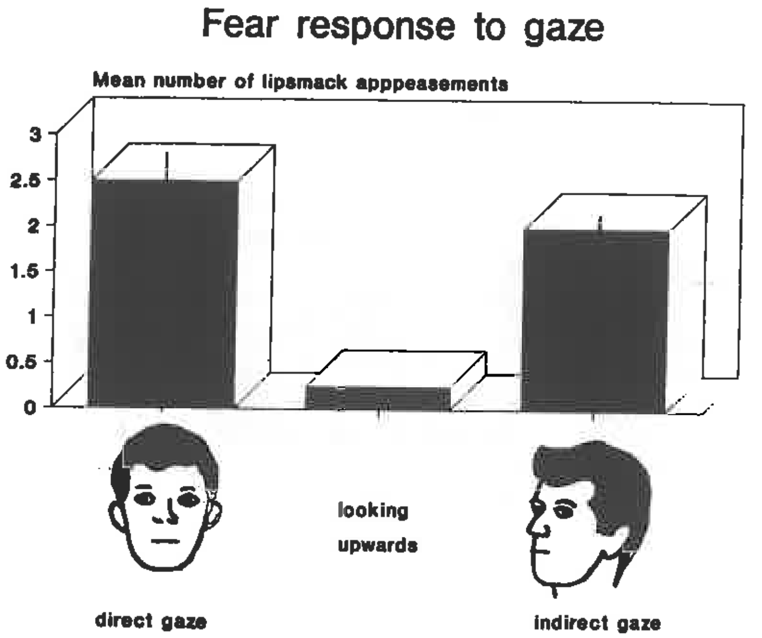


FIG. 5. Evidence of the fear response (measured in lip-smack appeasement) to eye-contact, by macaques (adapted from Perrett & Mistlin, 1990).

2.3. Rapid speed. As with ID, no specific reaction time data are available, but our impression is that "eyes looking at me" have a 'pop-out' effect—among a crowd of faces, we rapidly detect the face that is "looking at me".

2.4. Characteristic ontogenesis. If we focus on the human case, it is useful to think in terms of other people's eyes triggering EDD to fire, much like an Innate Releasing Mechanism (Tinbergen, 1951; Spitz, 1965). The idea is that when EDD detects eye-like stimuli, it fixates on these for relatively long bursts. We can infer the presence of EDD in the human neonate from Maurer's (1985) studies. For example, these report that 2 month old infants look almost as long at the eyes as at a whole face, whilst significantly less at other parts of the face[5]. These data are shown in Fig. 6. Haith, Bergman, and Moore (1977), and Hainline (1978) also found this, and Papousek and Papousek (1979) showed that 6 month olds look 2 to 3 times longer at a face looking at them than

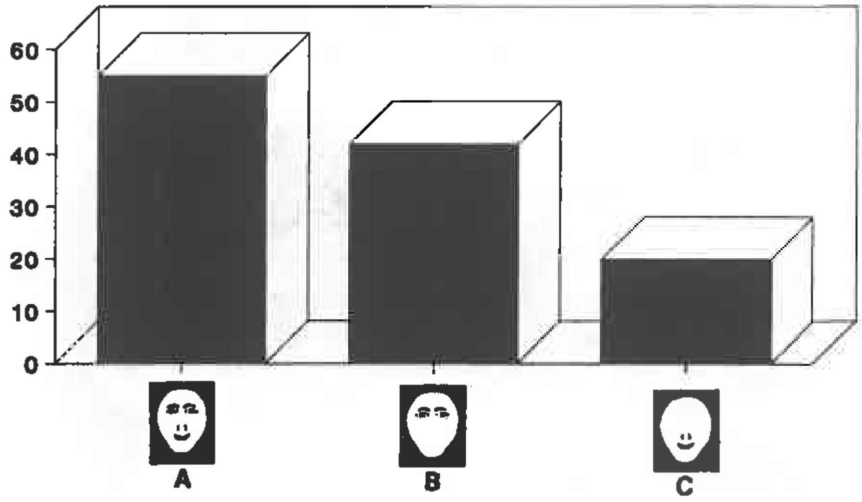
---

[5] Johnson and Morton (1991) suggest that a logically prior mechanism (CONSPEC) initially directs the newborn infant to look at faces. CONSPEC may thus locate the right "stimulus ball-park" for EDD to be triggered.

at a face looking away. At the very least, this shows a natural preference for looking at the eyes over and above other parts of the face.

Argyle and Cook (1976) note that during breast-feeding the infant is in an optimum position to see the mother's eyes, and that mothers use very long gaze durations (more than 30 secs) towards their infants, making them akin to "supernormal" stimuli (Stern, 1977). In addition, infant's control over their visual system appears precociously mature, enabling the infant to make or break eye-contact, and thus regulate the degree of eye-contact (Stern, 1977; Schaffer, 1977). Vine (1973) suggests it is the *contrast* of the mother's eyes which makes them salient within the face.

Developmental changes in the psychophysics of EDD merit further experimentation. Lasky and Klein (1979) found that 5 month old human infants look longer at a face showing direct eye-contact than one with averted gaze, and Johnson and Vecera (1993) report 4 month olds show a similar capacity to distinguish these. Butterworth's (1991) experiments reveal that 6 month olds can judge if an Agent's eyes are looking left or right, but not at which of two objects on one side the Agent is looking. By 12 months old, they can make these discriminations. Our own studies have shown that 3 year old normal children are perfect at distinguishing "Which one is looking at you?" (Baron-Cohen & Cross, 1992: see Fig. 7). Lord (1974) found 6 year olds' judgements



from Maurer, 1985

FIG. 6. Mean length of fixation by 2 month olds to face-like drawings (adapted from Maurer, 1985).

about gaze to different parts of the face were less accurate than adults', though Thayer (1977), using an on- versus off-face discrimination task, found that 6 year olds were as good as adults in detecting when another person was making eye-contact with them. However, they made more false positive errors by judging instances of off-face gazes as eye-contact.

Jaspars et al. (1973) found that human adults could discriminate eye-shifts of 1 cm at a distance of 100 cm, which corresponds to a change of 5 degrees, and Gibson and Pick (1963) also reported that human adults could detect an angular displacement of the eyeball of less than 3 degrees at 2 metres. Perrett

FIG. 7. "Which one is looking at you?": Photographic stimuli with (a) only eye-cues available, and (b) nose and eye-cues available (from Baron-Cohen & Coss, 1992).

and Milders (1992) found that adults were above chance in identifying a gaze shift of 2 degrees. Cline (1967) reports that human adults can detect even smaller deviations (1.4 degrees) of the eye, at a distance of 122 cm. It is not yet clear precisely how EDD distinguishes eye-contact from non-eye-contact—for example, is the relevant stimulus property the symmetry of the d region, as Anstis, Mayhew, and Morley (1969) suggest?[6]

2.5. Dedicated neural architecture. Some single-cell recording studies have found specific cells in the Superior Temporal Sulcus (STS) of the monkey brain respond selectively to direction of gaze (Perrett et al., 1985; Perrett et al., 1990; see Fig. 8). For example, Perrett et al. (1985) found that 64% of cells in the STS that are responsive to face or profile views of the head are also selective for the direction of gaze.

There is clear evidence of physiological arousal produced by mutual eye-contact. For example, galvanic skin responses increase with mutual eye-contact (Nichols & Champness, 1971). Wada (1961) also found brainstem activation in response to eye-stimuli in monkeys. These measures of arousal might of course be linked to either positive or negative emotions. In the case of human infancy, the evidence suggests it is linked with positive emotions[7], since eye-contact reliably triggers the infant to smile (Wolff, 1963; Stern, 1977; Schaffer, 1977). The neurobiology of EDD is considered in more detail elsewhere (see Baron-Cohen & Ring, 1994).

2.6. Characteristic pattern of breakdown. Lesions in the STS produce an impairment in the ability to discriminate gaze direction by monkeys (Campbell, Heywood, Cowey, Regard, & Landis, 1990). Some patients with prosopagnosia are also impaired in this ability (Campbell et al., 1990; Heywood & Cowey, 1992).

In sum, EDD fits most of the criteria for modularity.

3. SAM

SAM is the *Shared Attention Mechanism*. Its function is to identify if you and another organism are both attending to the same thing. This is a problem that it is important to be able to solve, but ID and EDD just cannot solve it. This is because the two earlier mechanisms are limited to building dyadic representations. SAM is held to be necessary for the development and production of

---

[6] I am grateful to Dave Perrett for directing me to this possibility.

[7] It is not only in infancy that eye-contact triggers pleasurable emotions. Rubin (1970) found that those couples rated as strongly in love gazed at each other more than those rated as weakly in love. Thayer and Schiff (1977) also found that judges rated reciprocated long gaze between mixed sex couples as a sign of greater sexual interest. See Kleinke (1986) for a review.

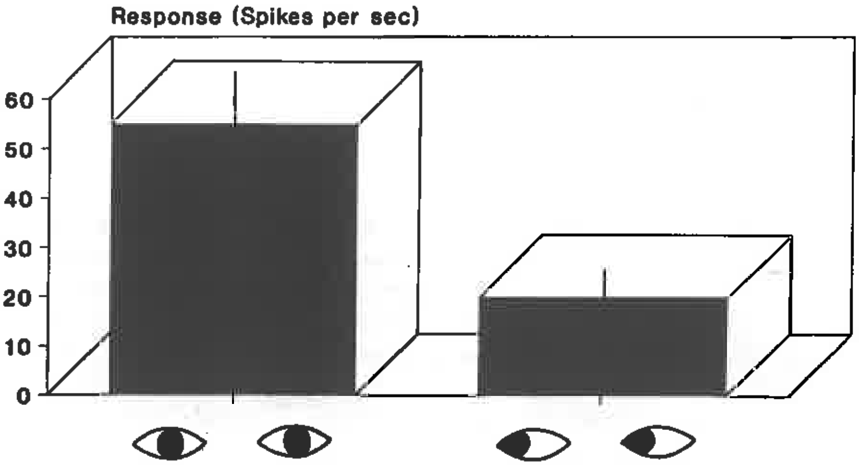


FIG. 8. Responsivity of cell M047 in the 575 to "pure" eye-direction (adapted from Perrett et al., 1990).

joint-attention behaviours[8]. In the human case, I suggest that SAM has two other key functions. First, to connect ID to EDD. Secondly, to trigger the last mechanism in the Mindreading System (ToMM).

Let us apply the 6 criteria for modularity to SAM.

3.1. Domain specificity. The conditions under which you and another person could be engaged in a mutually shared focus of attention are well-defined: there needs to be another Agent (let us call this Agent$_1$), and an additional object or Agent (if the latter, let us call this Agent$_2$). In addition, I suggest that, unlike the two earlier mechanisms, SAM codes these conditions using a unique representation, which I call a *triadic* representation[9]. Triadic representations differ in structure to dyadic representations in that they include an embedded element which specifies that Agent and Self are both attending to the *same* object. To capture this, they have one of two forms:

a. [Self-Relation-(Agent-Relation-Object)]. Here, the first relation term is bidirectional, so examples of this form are:

[I-see-(Mummy-sees-the bus)], and
[Mummy-sees-(I-see-the bus)].

---

[8] This mechanism could equally have been called a Joint Attention Mechanism, but the acronymn was obviously less desirable. Hence the name SAM.

[9] This term, like the term dyadic representation, is derived from Bakeman and Adamson (1984), and Trevarthen (1979). Hobson (1993) also refers to triadic relations. Note however that in my account, these are a class of *representation*.

Because this representation specifies that both I and Mummy are seeing the same bus at the same time, this fulfils the function of the triadic representation, namely, to identify shared attention.

b. [Self-Relation-(Agent$_1$-Relation-Agent$_2$)]. Here, both relation terms are bidirectional. So examples of this form are:

[I-see-(Mummy-sees-Daddy)], or
[I-see-(Daddy-sees-mummy)], or
[Mummy-sees-(I-see-Daddy)], or
[Mummy-sees-(Daddy-sees-me)], etc.

These sentence-like expressions are my attempt at specifying what triadic representations represent. However, it is questionable whether one can fully capture the complexity of the relations with such formal descriptions. The alternative spatial description, depicted in Fig. 9, may be both more comprehensive, and simpler to "read".

As mentioned earlier, the capacity to construct triadic representations is held to be necessary for joint-attention, and I suggest that in the first instance, SAM builds these representations using the dyadic representations that it obtains from EDD's output. This is because triadic representations can be built more easily in the visual modality than they can be in other modalities. However, in principle, SAM can build triadic representations from ID's output, in order to establish joint tactile or joint auditory attention.

3.2. Obligatory firing. Infants from the end of the first year of life begin to turn spontaneously to look in the same direction as another person, when that person's eye-direction suddenly changes (Scaife & Bruner, 1975). This is a response we continue to have as adults (Tantam, 1992). If we are told to suppress it, we can—but a short exercise in self-observation will prove to you that you monitor direction of gaze all the time. It is obligatory to that extent.

3.3. Rapid speed. Again, no reaction time data exist on the gaze-following response, but by any standards it would seem to be very rapid.

3.4. Characteristic ontogenesis. From about 9 months of age, normal infants begin to engage in a range of joint visual attention behaviours (Bruner, 1983). The clearest of these are gaze-monitoring (mentioned above), and the "protodeclarative" pointing gesture (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). In these behaviours, the infant alternates his or her own gaze between the adult's eyes and the object at which they are both attending, or at which the infant is directing the adult to look. Both of these behaviours are universally present by 9–14 months of age (Scaife & Bruner, 1975; Butterworth, 1991). Both continue to be present in older human children (Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1993) and adults (Argyle

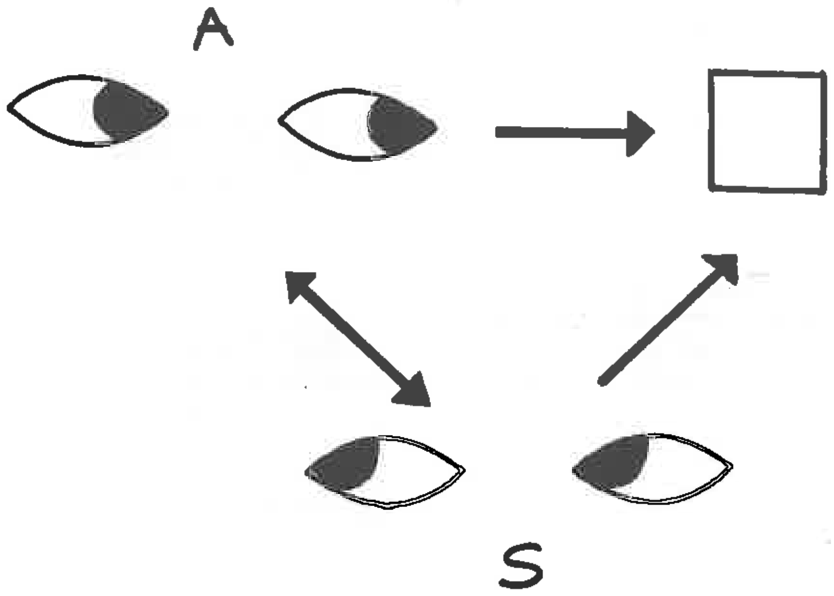


FIG. 9. Schematic illustration of a triadic representation: [Self-Relation-(Agent-Relation-Object)]. (S = Self; A = Agent).

& Cook, 1976), in whom they retain important communicative functions. There is some anecdotal evidence that chimpanzees and baboons may look in the same direction as another animal is looking, but this remains to be systematically investigated (Cheney & Seyfarth, 1990).

Regarding SAM's function of linking ID to EDD, this should result in volitional terms being imported into the relation slot of triadic representations, such that a person's *goal* or *desire* can be read from their eye-direction. Phillips, Baron-Cohen and Rutter (1992) investigated this with normal infants ranging from 9–18 months. The child was presented either with an ambiguous or an unambiguous action. One ambiguous action comprised *blocking* the child's hands during manual activity, by the adult cupping her hands over the child's. A second ambiguous action comprised offering an object to the child, but then at the last minute *teasingly* withdrawing it, just as the child began to reach for it. The unambiguous action simply comprised *giving* or presenting an object to the child.

This study found that, on at least half of the trials, 100% of the infants responded to the ambiguous actions by instantly looking at the adult's eyes (within the first 5 seconds after the tease or the block), whilst only 39% of them did so following the unambiguous action. This suggests that under conditions in which the goal of an action is uncertain, the first place young children (and indeed adults) look for information to disambiguate the goal is the eyes.

In a further study, we demonstrated that it is indeed eye-*direction* that children use to infer a person's goal (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, in press). Thus, when 3 year olds are asked "Which chocolate will Charlie take?", after being shown a display of 4 chocolates and Charlie's face looking at one of these, they tend to pick the one he is *looking at* as the goal of his next action (see Fig. 10).

Regarding inferring a person's *desire* from eye-direction, Baron-Cohen et al.'s (in press) study presented normal 3–4 year olds with the display of the 4 chocolates, and placed the cartoon face of Charlie in the centre of the display. Again, Charlie's eyes were depicted as pointing towards one of the 4 sweets, randomly selected (see Fig. 10). The subject was asked "Which one does Charlie *want?*". (In another, the subject was asked "Which one does Charlie *say* is the (x)?", in order to see if they used eye-direction to infer a person's intended referent.) Children of this age had no difficulty at all in inferring Charlie's desire (or his intended referent), from his eye-direction. Note that Baldwin (1991, 1995) has also reported 18 month olds' ability to use eye-direction to infer a person's intended referent.

3.5. Dedicated neural architecture. We currently have no clues at all as to any possible neural localization of SAM. One should note that because SAM is

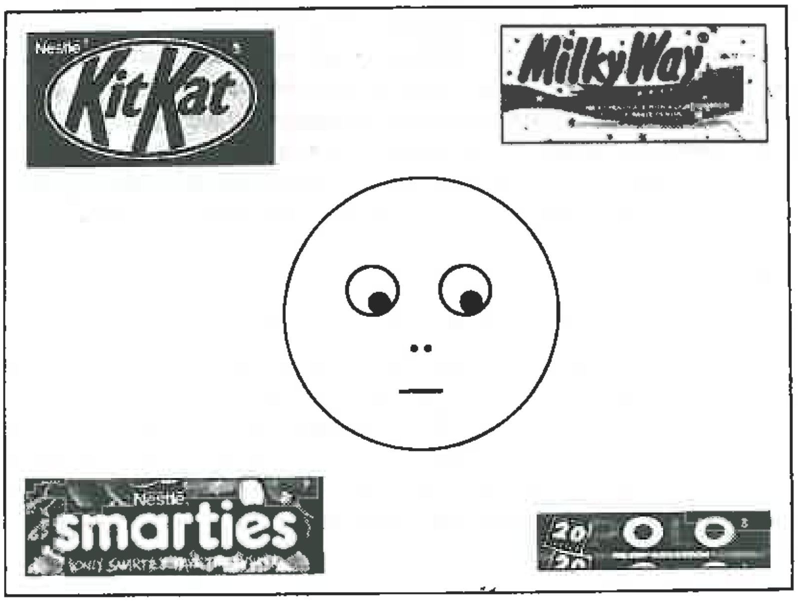


FIG. 10. The 4 Sweets display (From Baron-Cohen et al., in press).

amodal, the studies showing eye-direction sensitive cells in the STS can only be linked to EDD, and not necessarily to SAM.

3.6. Characteristic pattern of breakdown. SAM appears to be impaired in most cases of autism, whilst both ID and EDD appear to be intact. Evidence for ID being intact in autism includes the following: These children use words referring to goal-directed action and desire in their spontaneous speech (Baron-Cohen, Leslie, & Frith, 1986; Tager-Flusberg, 1992, 1993), and they can predict emotions on the basis of a person's desire (Baron-Cohen, 1991a). Evidence for EDD being intact is that they can detect when the eyes of another person are directed at them (Baron-Cohen et al., in press). Evidence that SAM is impaired in autism is that they show few if any joint-attention behaviours (Sigman, Mundy, Ungerer, & Sherman, 1986; Leekam et al., 1993; Baron-Cohen, 1989a; Phillips et al., 1992).

Phillips et al. (1992) tested very young children with autism for their ability to use SAM to detect a person's goals from their eye-direction, using the ambiguous and unambiguous actions described earlier. However, these children did not seem to use eye-contact to disambiguate the ambiguous actions, looking as little in both conditions (less than 11% looking, in each). Baron-Cohen et al. (in press) also tested children with autism on the 4 Sweets Task and found significant impairments in the use of eye-direction in inferring want, goal, or intended referent.

If children with autism are not capable of processing triadic representations, how are they able to pass visual perspective-taking tasks, which a number of experiments show that they do (Hobson, 1984; Baron-Cohen, 1989a, 1991b; Tan & Harris, 1991)? According to this theory, they must do this by employing dyadic representations, of the form [Agent-Relation-Object].

I will defer discussion of my earlier claim about the precursor relation between SAM and ToMM, until I have introduced this final mechanism. To summarize this section, SAM fits many of the criteria for modularity.

4. ToMM

ToMM is the *Theory of Mind Mechanism*. This is the name Leslie (1991) gives to the system underpinning our everyday ability to make sense of behaviour in terms of mental states, and predict an Agent's behaviour on the basis of such states[10]. I take the liberty of slightly elaborating Leslie's proposal, to clarify that ToMM has two principal functions: (1) To represent the full range of mental state concepts, including the epistemic ones; and (2) to integrate mental state knowledge into a coherent and usable "theory" for the human child and adult to employ.

---

[10] Unless I explicitly state otherwise, the reader can assume that I am using the term ToMM to refer to what Leslie (1994) also calls ToMM System$_2$.

Applying the 6 criteria for modularity to ToMM, what do we find?

4.1. Domain specificity. Certainly, the domain of application for ToMM appears highly specific: something either is a mental state, or it is not (Brentano, 1874/1970). Understanding the full range of mental states also requires the processing of a unique kind of representation. Leslie and Roth (1993) call these *M-Representations*[11], and suggest they have the following structure:

[Agent-Attitude-"Proposition"].

An example of what could fill these three slots is shown here:

[John-thinks-"The money is in the biscuit tin"]

In the above example, the whole M-Representation can be true even if the proposition is false. For example, the M-Representation can be true if John indeed *thinks* the money is in the biscuit tin, even if the money is in fact in the teapot. The usefulness of M-Representations is that they allow sensitive prediction of an Agent's future action. For example, they lead us to predict that John will go to the biscuit tin if he wants the money, despite the money in reality being in the teapot. M-Representations also allow one to make sense of an Agent's behaviour. For example, they help make sense of why John might look disappointed when he opens the biscuit tin[12].

4.2. Obligatory firing. ToMM also seems to meet this criterion. As adults, it takes enormous effort for us *not* to interpret behaviour in mentalistic ways. Students of behaviourism will recall having to be specially taught to suppress the urge to refer to mental states in their descriptions—it does not come naturally to us to refer to behaviour exclusively. Rather, reference to mental states seems to be our involuntary and spontaneous way of "reading" behaviour (Dennett, 1978).

4.3. Rapid speed. Whilst no one has attempted to time our ability to compute actions in terms of mental states, our impression is that this is not only effortless (to most of us), but indeed rapid.

4.4. Characteristic ontogenesis. ToMM appears to have a very characteristic timetable (Astington, Harris, & Olson, 1988). Let me sketch this here: According to Leslie (1987, 1991), ToMM comes on line in the middle of the second year of life, and its arrival is marked by the production and

---

[11] They adopt this term in order to avoid the confusions arising from the earlier term "metarepresentations" (Perner, 1993).

[12] In recent writings, Leslie suggests M-Representations have a 4th term, to express an anchor in reality (e.g., see Leslie, German, & Happé, 1993).

comprehension of pretence, around 18–24 months of age (Harris & Kavanaugh, 1993). By 3 years of age, preschoolers are able to understand aspects of what people know (Pratt & Bryant, 1990), and by 4 years of age they can distinguish true and false beliefs (Wimmer & Perner, 1983). Their mental-state knowledge is also highly organised into a coherent "theory" which the child uses for both explanation and prediction of action (Wellman, 1990). Preliminary cross-cultural studies support the universality of this developmental timetable (Avis & Harris, 1990; Riviere, personal communication; Jolly, 1966).

4.5. Dedicated neural architecture. So far, there is little evidence regarding the specific neurobiological basis of ToMM. However, since non-human primates do not show any convincing evidence of possessing ToMM, if one judges this by whether they understand epistemic mental states (Premack, 1988; Whiten, 1993; Hayes, in press), this suggests it may have appeared only with the evolution of the *Homo sapiens* brain. A recent functional imaging study that we have carried out using SPECT (Single Photon Emission Computed Tomography) suggests that the right orbito-frontal cortex may play a specific role in supporting ToMM in normal human adults (Baron-Cohen, Ring, Moriarty, Schmidt, Costa, & Ell, 1994). This would be consistent with the neurological evidence that damage to this area can produce "loss of social judgement" (Eslinger & Damasio, 1985) and a breakdown in the pragmatics of language (Kaczmarek, 1984). (See Baron-Cohen & Ring [1994] for a review of the possible neurobiology of ToMM.)

4.6. Characteristic pattern of breakdown. Children with autism are impaired in their understanding of epistemic mental states, such as knowing and believing (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, 1989b & c, 1991a & b, 1992; Leslie & Frith, 1989; Perner, Frith, Leslie, & Leekam, 1989; Sodian & Frith 1992; Goodhart & Baron-Cohen, 1994). However they are unimpaired in their understanding of non-mental representations (such as photographs, drawings, maps, and models: Leslie & Thaiss, 1992; Leekam & Perner, 1991; Charman & Baron-Cohen, 1992, 1993)[13]. They also lack spontaneous pretend play (Wing & Gould, 1979; Baron-Cohen, 1987). Finally, their spontaneous speech also lacks terms referring to epistemic mental states (Baron-Cohen et al., 1986; Tager-Flusberg, 1993). This specific deficit in ToMM is thought to relate to the abnormalities in social and communicative behaviour that are key symptoms of autism (Baron-Cohen, 1988, 1990; Frith, 1989).

In sum, ToMM fits the criteria for modularity fairly well.

---

[13] See Baron-Cohen, Tager-Flusberg, and Cohen (1993), where the relevant evidence is brought together and debated.

5. The Relationship Between the Four Mechanisms

There is a big difference between the first three mechanisms and ToMM, in that the small set of mental states that the first three are able to represent possess only two of the properties of Intentionality: aboutness, and aspectuality (Dennett, 1978; Perner, 1991). By contrast, the attitude concepts that ToMM can represent progressively include *pretend*, *know*, *think*, and *believe*, amongst others. These possess a third property of Intentionality: the possibility for misrepresentation (Leslie, 1987; Perner, 1991). ToMM is therefore both more versatile than the other three, and has a larger set of mental state terms that it can represent.

A second important relationship to bring out is the earlier claim that SAM stands in a causal relationship to ToMM: that the key way ToMM is activated is by taking as input SAM's triadic representations. Elsewhere, I have referred to this special relationship in terms of SAM *facilitating* the development of ToMM (Baron-Cohen and Cross, 1992; Baron-Cohen, 1995). Focusing on the disorders of autism and congenital blindness, as Hobson (1990, 1993) recommends we do, brings out the reasons for suggesting this causal relationship.

6. Autism and Congenital Blindness

The claim is that SAM is a necessary (though not sufficient) condition for the development of ToMM. This is a testable claim about a certain kind of precursor relationship between the two systems in development. Note that whilst SAM usually builds triadic representations using EDD's dyadic representations, this need not be the case. For example, since children with congenital blindness lack EDD, SAM must be restricted to building triadic representations specifying joint-attention via touch or audition (e.g., [I-touch-(Mummy-touches-the cup)]). That is, SAM, in their case, must build triadic representations using ID's dyadic representations. These kinds of triadic representations are likely to be considerably more difficult to build than those derived from EDD (see Baron-Cohen & Cross, 1992).

Given the evidence presented earlier that most children with autism fail to develop a fully functioning SAM, it follows from the theory outlined here that in these children this would have the knock-on effect of not activating ToMM. The theory thus predicts two subgroups of autism:

Subgroup A: both SAM and ToMM are impaired, as explained by the knock-on hypothesis. There is considerable evidence suggesting that many subjects with autism fall into this group (see Baron-Cohen, Tager-Flusberg, & Cohen, 1993; Baron-Cohen, 1991d and 1993, for reviews). In addition, absence of SAM is one predictor of autism at 18 months of age (Baron-Cohen, Allen, & Gillberg, 1992).

Subgroup B: SAM is intact, whilst ToMM is impaired in its own right. These children might correspond to those children with autism who are reported to have a period of normal development up to the age of 18 months (Volkmar & Cohen, 1989; Derek Ricks, personal communication, 1987), and then show clear signs of autism. This group remains to be fully investigated.

Finally, it follows that in children with congenital blindness, since SAM is intact, ToMM should develop, although a slight delay in this would not be surprising given the need for SAM to use ID instead of EDD.

ALTERNATIVE ACCOUNTS

Alan Leslie's Account

Leslie (1994) suggests two further mechanisms: ToBy (the *Theory of Bodies Mechanism*) detects if an object moved as a result of external causes, or is an Agent capable of self-propulsion, whilst ToMM System$_1$ detects an Agent's action as goal-directed. In effect, Leslie proposes two mechanisms to cover the functions that Premack (1990) and I suggest can be covered by one (ID). It remains to be shown that Leslie's division is justified by neuropsychological or neurological dissociation. An additional difference between ToBy and ID is that the former is a "mechanics module" (Leslie, 1994)—it processes information about physical causality as well as Agency. Patients who 'dissociate' in these two skills would pose a problem for Leslie's account of ToBy, whilst they would not for Premack or my account, since ID is not a mechanics module[14].

Furthermore, neither ToBy nor ToMM System$_1$ would be sufficient to account for why in the normal case we attend disproportionately to an Agent's eyes, nor why children with autism may attend to other aspects of Agents (such as their hands), but not especially to their eyes (Phillips, Laa, Gomez, Baron-Cohen, & Riviere, 1994). For this, I think you need EDD and SAM. Leslie's two systems would predict that the aspect of an Agent to which we might attend could be entirely arbitrary—it could be their shoulders, their hands, or even their feet! In my account, it is no coincidence that these aspects of people are passed over relatively quickly when we engage in "person perception": It is because we have two specific mechanisms (EDD and SAM) which drive us to attend to the eye-region, first and foremost.

Johnson and Morton's Account

Johnson and Morton (1991) suggest that CONLERN is a mechanism that comes on line in human development around 2 months of age, and that is sufficient for

---

[14] Indeed, I would prefer to see a separation between a mechanism (like ToBy) that processes the mechanics of objects, and a mechanism (like ID) that identifies Agency and interprets this in terms of volition.

learning about all aspects of faces and conspecifics (facial identity, expression, etc.). On their account, EDD and SAM would not be necessary. In my view, the data that I have reviewed suggest that we orient disproportionately to the eye-region, over and above other regions of a conspecific. EDD (and later, SAM) are my suggestions to account for this bias in information processing. Moreover, it is not an arbitrary bias. Rather, I suggest that this bias is highly adaptive, in that eye-direction is a reliable source of information about an Agent's goals, desires, and focus of attention. A further problem for Johnson and Morton's account is to explain why many aspects of CONLERN-processing—such as identity recognition, relationship perception, and perceptual role-taking—are intact in autism (Baron-Cohen, 1991c), whilst joint-attention is impaired. SAM is my way of overcoming this difficulty.

Digby Tantam's Account

In an important recent article, Tantam (1992) argues that (1) there is "an innate and involuntary orientation to social stimuli" (p.84); (2) "this *social attention response is* . . . absent or impaired in autism and . . . is the primary social abnormality" (p.84); (3) In the normal child there is a "second gaze response, in which gaze is drawn from a person's eyes to the object at which they are looking" (p.85); (4) "This second gaze response presupposes the first" (p.85); (5) The second gaze response is therefore also impaired in autism; and (6) "Weakness or absence of the social gaze response is enough . . . to account for many of the typical symptoms of autism, including the failure to acquire a theory of mind" (p.83).

Tantam's thesis is interesting. Effectively, he has renamed eye-contact as the "first gaze response" and joint-attention as the "second gaze response". If we equate what he calls the "first gaze response" with EDD, and the second gaze response with SAM, then his claim would presumably be that the primary deficit in autism is in EDD. However, my earlier review of the data suggests that EDD is largely intact in autism. It may be that Tantam is right, but as yet the evidence for this is insufficient.

The theory I have presented here differs from Tantam's thesis in describing the two mechanisms driving eye-contact and joint-attention (EDD and SAM); and in suggesting that whilst SAM is necessary for the development of ToMM, it is *not sufficient* for this—since ToMM is held to be independent of SAM. I would therefore take issue with Tantam's suggestion that absence of the "first social gaze response" (here taken to be equivalent to EDD) is enough to account for the failure to acquire a theory of mind. Indeed, the prediction from Tantam's theory is that the blind should also be impaired in ToMM, which, as far as we know, they are not.

Other Accounts

Some authors have argued that the origins of ToMM lie in (a) the capacity for neonatal imitation (Meltzoff & Gopnik, 1993), or (b) the capacity for social-mirroring (Gergely et al., in press), or (c) the capacity for affective responsivity (Hobson, 1990). Part of my reason for highlighting SAM in the precursor role is that the evidence for autism-specific deficits in these other domains is less robust than it is in the domain of joint-(visual)-attention. However, a plausible case has been made for all of these as candidate precursors to ToMM, and the real test will be to examine which of these (either singly, or in combination) is predictive of ToMM. Such longitudinal studies are time-consuming, but ultimately will be the only way to choose between these hypotheses.

CONCLUSIONS

I close this paper by drawing the reader's attention to the 8 claims in the Abstract, in order to summarize the arguments. In addition to these, I hope it is clear that as well as suggesting a model of the human Mindreading System, each component mechanism in this system has been evaluated in terms of the degree to which it meets criteria for modularity. In Table 1 below, I summarize the neuropsychological dissociations among the four mechanisms both in normally developing populations, and in some developmentally abnormal populations (all human). Future work will need to address the extent to which the four proposed mechanisms are present in different species. In terms of normal development, this model implies that there are at least two distinct

TABLE 1
Summary of the neuropsychological dissociations between the 4 mechanisms.

	ID	*EDD*	*SAM*	*ToMM*
Normal children				
6 month olds	+	+	–	–
14 month olds	+	+	+	–
48 month olds	+	+	+	+
Abnormal children*				
Autism (A)	+	+	–	–
Autism (B)	+	+	+	–
Blind#	+	–	+	+
AB type prosopagnosia@	+	–	+	+

* = given a mental age above 48 months approximately;
\+ = mechanism is functioning;
– = mechanism is not functioning;
= congenitally blind;
@ = see Campbell et al. (1990).

theories of mind that exist at different ages: 14 month olds have what I have elsewhere called an "attention-goal" psychology (Baron-Cohen, 1993), whilst 4 year olds (and older people) have what is usually referred to as a "belief-desire" psychology (Wellman, 1990).

ACKNOWLEDGEMENTS

During the preparation of this work, I was at the Institute of Psychiatry, London, and received support from the Mental Health Foundation and the Medical Research Council. This paper evolved through presentations at the Society for Philosophy and Psychology, Montreal, June 1992, the European Conference of Developmental Psychology, Seville, September 1992, the Society for Research in Child Development, New Orleans, March 1993, and the European Society for Philosophy and Psychology, Sheffield, June 1993.

I am grateful to Annette Karmiloff-Smith, Ruth Campbell, Dave Perrett, Sue Leekam, Helen Tager-Flusberg, Pierre Jacob, and Donald Cohen for their comments on earlier drafts of this paper, and to Leslie Brothers, Juan Carlos Gómez, Digby Tantam, John Swettenham, and Ami Klin for discussions. Finally, the title of this paper owes a lot to Mandler (1992).

REFERENCES

Altmann, S. (Ed.) (1967). *Social communication among primates*. Chicago, IL: University of Chicago Press.

Anstis, S., Mayhew, J., & Morley, T. (1969). The perception of where a face or a television 'portrait' is looking. *American Journal of Psychology, 82*, 474–489.

Arduino, P., & Gould, J. (1984). Is tonic immobility adaptive? *Animal Behaviour, 32*, 921–922.

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge, MA: Cambridge University Press.

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

Avis, J., & Harris, P. (1991). Belief-desire reasoning among Baka children: evidence for a universal conception of mind. *Child Development, 62*, 460–467.

Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interaction. *Child Development, 55*, 1278–1289.

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*, 875–890.

Baldwin, D. (1995). Understanding the link between joint attention and language acquisition. In C. Moore & P. Dunham (Eds.), *Joint attention: its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Baron-Cohen, S. (1987). Autism and symbolic play. *British Journal of Developmental Psychology, 5*, 139–148.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders, 18*, 379–402.

Baron-Cohen, S. (1989a). Perceptual role-taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology, 7*, 113–127.

Baron-Cohen, S. (1989b). Are autistic children behaviourists? An examination of their mental–physical and appearance–reality distinctions. *Journal of Autism and Developmental Disorders, 19*, 579–600.

Baron-Cohen, S. (1989c). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30,* 285–298.

Baron-Cohen, S. (1990). Autism: a specific cognitive disorder of "mindblindness". *International Review of Psychiatry, 2,* 19–88.

Baron-Cohen, S. (1991a). Do people with autism understand what causes emotion? *Child Development, 62,* 385–395.

Baron-Cohen, S. (1991b). The development of a theory of mind in autism: deviance and delay? *Psychiatric Clinics of North America, 14,* 33–51.

Baron-Cohen, S. (1991c). The theory of mind deficit in autism: how specific is it? *British Journal of Developmental Psychology, 9,* 301–314.

Baron-Cohen, S. (1991d). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind.* Oxford: Basil Blackwell.

Baron-Cohen, S. (1992). Out of sight or out of mind: another look at deception in autism. *Journal of Child Psychology and Psychiatry, 33,* 1141–1155.

Baron-Cohen, S. (1993). From attention-goal psychology to belief-desire psychology: the development of a theory of mind, and its dysfunction. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism.* Oxford: Oxford University Press.

Baron-Cohen, S. (1995). The Eye-Direction Detector (EDD) and the Shared Attention Mechanism (SAM): two cases for evolutionary psychology. In C. Moore & P. Dunham (Eds.), *The role of joint attention in development.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry, 161,* 839–843.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (in press). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology.*

Baron-Cohen, S., & Cross, P. (1992). Reading the eyes: evidence for the role of perception in the development of a theory of mind. *Mind and Language, 6,* 166–180.

Baron-Cohen, S., & Goodhart, F. (1994). The "seeing leads to knowing" deficit in autism: the Pratt and Bryant probe. *British Journal of Development Psychology* (in press).

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

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology, 4,* 113–125.

Baron-Cohen, S., & Ring, H. (1994). A model of the Mindreading System: neuropsychological and neurobiological perspectives. In P. Mitchell & C. Lewis (Eds.), *Origins of an understanding of mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmidt, P., Costa, D., & Ell, P. (1994). Recognition of mental state terms: a clinical study of children with autism, and a functional neuroimaging study of normal adults. *British Journal of Psychiatry* (in press).

Baron-Cohen, S., Tager-Flusberg, H., and Cohen, D. I. (Eds.) (1993). *Understanding other minds: perspectives from autism.* Oxford: Oxford University Press.

Bartsch, K., & Wellman, H. (1994). *Talk about the mind.* Oxford: Oxford University Press.

Bates, E. (1993). Invited lecture to the MRC Cognitive Development Unit, London.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). Cognition and communication from 9–13 months: correlational findings. In E. Bates (Ed.) *The emergence of symbols: cognition and communication in infancy.* New York: Academic Press.

Blest, A. (1957). The function of eyespot patterns in the Lepidoptera. *Behaviour, 11,* 209–256.

Brentano, F., von (1874/1970). *Psychology from an empirical standpoint.* [O. Kraus (Ed.), translated by L. MacAllister]. London: Routledge and Kegan Paul.

Brothers, L. (1990). The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. *Concepts in Neuroscience, 1*, 27–51.

Bruner, J. (1983). *Child's talk: learning to use language*. Oxford: Oxford University Press.

Burghardt, G. (1990). Cognitive ethology and critical anthropomorphism: a snake with two heads and hog-nosed snakes that play dead. In C. Ristau (Ed.), *Cognitive ethology: the minds of other animals*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of mind*. Oxford: Basil Blackwell.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9*, 55–72.

Byrne, R., & Whiten, A. (1988). *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Oxford University Press.

Campbell, R., Heywood, C., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia, 28*, 1123–1142.

Campbell, R., Walker, J., & Baron-Cohen, S. (1993). Aspects of part-of face processing in children with autism: A discriminant function analysis. Unpublished manuscript, Goldsmiths College, University of London.

Chance, M. (1967). The interpretation of some agonistic postures: the role of "cut-off" acts and postures. *Symposium of the Zoological Society of London, 8*, 71–89.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world*. Chicago, IL: University of Chicago Press.

Charman, T., and Baron-Cohen, S. (1992). Understanding beliefs and drawings: a further test of the metarepresentation theory of autism. *Journal of Child Psychology and Psychiatry, 33*, 1105–1112.

Charman, T., & Baron-Cohen, S. (1993). Understanding models, photos, and beliefs: a further test of the modularity thesis of misrepresentation. Unpublished manuscript, University College London, Psychology Department.

Cline, M. (1967). Unpublished manuscript, Cornell University.

Dasser, V., Ulbaek, I., & Premack, D. (1989). The perception of intention. *Science, 243*, 365–367.

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

Eslinger, P., & Damasio, A. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology, 35*, 1731–41.

Farah, M., McMullen, P., & Meyer, M. (1991). Can recognition of living things be selectively impaired? *Neuropsychologia, 29*, 185–193.

Fehr, B., & Exline, R. (1978). Social visual interaction: a conceptual and literature review. In A. Siegman & S. Feldstein (Eds.), *Nonverbal behaviour and communication* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fodor, J, (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: MIT Press.

Frith, U. (1989). *Autism: explaining the enigma*. Oxford: Basil Blackwell.

Gallup, G. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2*, 237–248.

Gallup, G., Cummings, W., & Nash, R. (1972). The experimenter as an independent variable in studies of animal hypnosis in chickens (*Gallus gallus*). *Animal Behaviour, 20*, 166–169.

Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (in press). Taking the intentional stance at 12 months of age. *Cognition*.

Gibson, J. & Pick, A. (1962). Perception of another person's looking behaviour. *American Journal of Psychology, 76*, 386–394.

Goodglass, H., Klein, H., Carey, P., & Jones, K. (1966). Specific semantic word categories in aphasia. *Cortex, 2,* 74–89.

Hainline, L. (1978). Developmental changes in visual scanning of face and non-face patterns by infants. *Journal of Experimental Child Psychology, 25,* 90–115.

Haith, M., Bergman, T., & Moore, M. (1977). Eye contact and face scanning in early infancy. *Science, 198,* 865–855.

Hall, K., & Devore, I. (1965). Baboon social behaviour. In I. Devore (Ed.), *Primate behaviour.* New York: Holt, Rinehart and Winston.

Harris, P., & Kavanaugh, R. (1993). Young children's comprehension of pretence. *Monographs of the Society for Research in Child Development, 58* (1).

Heyes, C., (in press). Social cognition in primates. In N. Mackintosh (Ed.), *Handbook of perception and cognition: Volume 9: Animal learning and cognition.* New York: Academic Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behaviour. *American Journal of Psychology, 57,* 243–259.

Heit, G., Smith, M., & Halgren, E. (1988). Neural encoding of individual words and faces by the human hippocampus and amygdala. *Nature, 333,* 773– 175.

Hennig, C. (1977). Effects of simulated predation on tonic immobility in *Anolis carolinensis*: The role of eye contact. *Bulletin of the Psychonomic Society, 9,* 239–242.

Hess, E., & Petrovich, S. (1978). Pupillary behaviour in communication. In A. Siegman & S. Feldstein (Eds.), *Nonverbal behaviour and communication* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heywood, C., & Cowey, A. (1992). The role of the 'face cell' area in the discrimination and recognition of faces in monkeys. In V. Bruce, A. Cowey, A. Ellis, & D. Perrett (Eds.), Processing the facial image. *Philosophical Transactions of the Royal Society of London, B 335,* 1–128.

Hobson, R. P. (1984). Early childhood autism and the question of egocentrism. *Journal of Autism and Developmental Disorders, 14,* 85–104.

Hobson, R. P. (1990). On acquiring knowledge about people and the capacity to pretend: response to Leslie (1987). *Psychological Review, 97,* 114–121.

Hobson, R. P. (1993). *Autism and the development of mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Humphrey, N. (1984). *Consciousness regained.* Oxford: Oxford University Press.

Jaspars, J., et al. (1973). Het observen van ogencontact. *Nederlands Tijdschrijt voor de Psychologie, 28,* 67–81. [Cited in Argyle and Cook (1976)]

Johnson, M., & Morton, J. (1991). *Biology and cognitive development.* Oxford: Basil Blackwell.

Johnson, M., & Vicera, S. (1993). Cortical parcellation and the development of face processing. In M. de Boysson-Hardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: speech and face-processing in the first year of life.* Dordrecht: Kluwer Academic Press, NATO ASI Series.

Jolly, A. (1966). Lemur social behaviour and primate intelligence. *Science, 153,* 501–506.

Kaczmarek, B. (1984). Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes. *Brain and Language, 21,* 52–58.

Kanner, L. (1943). Autistic disturbance of affective contact. *Nervous Child, 2,* 217–250. [Reprinted in L. Kanner (1973), *Childhood psychosis: Initial studies and new insights.* New York: John Wiley and Sons]

Karmiloff-Smith, A. (1992). *Beyond modularity.* MTT Press/Bradford Books.

Keating, C., & Keating, E. (1982). Visual scan patterns of rhesus monkeys viewing faces. *Perception, 11,* 211–219.

Kendrick, K., & Baldwin, B. (1981). Cells in the temporal cortex of conscious sheep can respond preferably to the sight of faces. *Science, 236,* 448–450.

Kleinke, C. (1986). Gaze and eye-contact: a research review. *Psychological Bulletin, 100,* 78–100.

Langdell, T. (1978). Recognition of faces: an approach to the study of autism. *Journal of Child Psychology and Psychiatry, 19*, 225–238.

Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1993). Eye direction detection: a dissociation of geometric and joint-attention skills in autism. Unpublished manuscript, Institute of Social and Applied Psychology, University of Kent, Canterbury.

Leekam, S., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*, 203–218.

Lempers, J., Flavell, E., & Flavell, J. (1977). The development in very young children of tacit knowledge concerning visual perception. *Genetic Psychology Monographs, 95*, 3–53.

Leslie, A. M. (1987). Pretence and representation: the origins of "theory of mind". *Psychological Review, 94*, 412–426.

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*. Oxford: Basil Blackwell.

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. Cambridge, MA: Cambridge University Press.

Leslie, A. M., German, T. P., & Happé, F. G. (1993). Even a theory-theory needs information processing: ToMM, an alternative theory-theory of the child's theory of mind. *Behavioral and Brain Sciences, 16*, 56–57.

Leslie, A. M., & Thaiss, L. (1992). Domain specifcity in conceptual development: Neuropsychological evidence from autism. *Cognition, 43*, 225– 251.

Leslie, A. M., & Roth, D. (1993). What autism teaches us about metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Ed.), *Understanding other minds: perspectives from autism*. Oxford: Oxford University Press.

Lord, C. (1974). The perception of eye contact in children and adults. *Child Development, 45*, 1113–1117.

Mandler, J. (1992). How to build a baby, II: Prelinguistic primitives. *Psychological Review, 99*, 587–604.

Maurer, D. (1985). Infants' perception of facedness. In T. Field & M. Fox (Eds.), *Social perception in infants*. Norwood, NJ: Ablex.

Meltzoff, A., & Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism*. Oxford: Oxford University Press.

Mendelsohn, M., Haith, M., & Goldman-Rakic, P. (1982). Face scanning and responsiveness to social cues in infant monkeys. *Developmental Psychology, 18*, 222–228.

Nichols, K., & Champness, B. (1971). Eye gaze and the GSR. *Journal of Experimental Social Psychology, 7*, 623–626.

O'Brien, T., & Dunlap, W. (1975). Tonic immobility in the blue crab (*Callinectes sapidus*): Its relation to threat of predation. *Journal of Comparative and Physiological Psychology, 89*, 86–94.

Papousek, H., & Papousek, M. (1979). Early ontogeny of human social interaction: Its biological roots and social dimensions. In M. von Cranach, K. Foppa, W. Lepenies, & D. Ploog (Eds.), *Human ethology: claims and limits of a new discipline*. Cambridge, MA: Cambridge University Press.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: Bradford Books, MIT Press.

Perner, J. (1993). The theory of mind deficit in autism: rethinking the metarepresentation theory. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism*. Oxford: Oxford University Press.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. (1989). Exploration of the autistic child's theory of mind: knowledge, belief, and communication. *Child Development, 60*, 689–700.

Perrett, D., Hietanen, M., Oram, W., & Benson, P. (1992). Organization and function of cells responsive to faces in the temporal cortex. In V. Bruce, A. Cowey, A. Ellis, & Perrett, D. (Eds.), Processing the facial image. *Philosophical Transactions of the Royal Society of London, B 335*, 1–128.

Perrett, D., Harries, M., Mistlin, A., Hietanen, J., Benson, P., Bevan, R., Thomas, S., Oram, M., Ortega, J., & Brierley, K. (1990). Social signals analyzed at the single cell level: someone is looking at me, something touched me, something moved! *International Journal of Comparative Psychology, 4*, 25–55.

Perrett, D., & Milders, M. (1992). Unpublished data, Department of Psychology, University of St Andrews, Scotland.

Perrett, D., & Mistlin, A. (1990). Perception of facial characteristics by monkeys. In W. Stebbins & M. Berkley (Eds.), *Comparative perception, Vol II: Complex signals*. New York: John Wiley and Son.

Perrett, D., Smith, P., Potter, D., Mistlin, A., Head, A., Milner, A., & Jeeves, M. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, B223*, 293–317.

Phillips, W., Baron-Cohen, S., & Rutter, M. (1992). The role of eye-contact in goal detection: evidence from normal infants, and children with mental handicap or autism. *Development and Psychopathology, 4*, 375–383.

Phillips, W., Laa, V., Gomez, J.-C., Baron-Cohen, S., & Riviere, A. (1994). Treating people as objects, agents, or subjects. Unpublished manuscript, Institute of Psychiatry, London.

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973–983.

Premack, D. (1988). 'Does the chimpanzee have a theory of mind?' revisited. In R. Byrne & A. Whiten (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford: Oxford University Press.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1–16.

Premack, D. (1991). The infant's theory of self-propelled objects. In D. Frye & C. Moore (Eds.), *Children's theories of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Premack, D. (1993). Invited address, MRC Cognitive Development Unit, University College London, November.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a 'theory of mind'? *Behavioral and Brain Sciences, 1*, 515–526.

Poulton, E. (1890). *The colours of animals*. London. [Cited in Blest (1957)]

Ristau, C. (1990). Aspects of the cognitive ethology of an injury feigning plover. In C. Ristau (Ed.), *Cognitive ethology: the minds of other animals*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ristau, C. (1991). Attention, purposes, and deception in birds. In A. Whiten (Ed.), *Natural theories of mind*. Oxford: Basil Blackwell.

Rubin, A. (1970). Measurement of romantic love. *Journal of Personal and Social Psychology, 16*, 265–273.

Sargent, A., & Eberhardt, L. (1975). Death feigning by ducks in response to predation by red foxes (*Vulpes fulva*). *American Midland Naturalist, 94*, 108–119.

Scaife, M. (1976). The response to eye-like shapes by birds. II: The importance of staring, pairedness, and shape. *Animal Behaviour, 24*, 200–206.

Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the human infant. *Nature, 253*, 265.

Schaffer, H. (1977). Early interactive development. In H. Schaffer (Ed.), *Studies in mother–infant interaction*. New York: Academic Press.

Schaffer, G. (1964). *The mountain gorilla*. Chicago, IL: Chicago University Press.

Sigman, M., Mundy, P., Ungerer, J., & Sherman, T. (1986). Social interactions of autistic, mentally retarded, and normal children and their caregivers. *Journal of Child Psychology and Psychiatry, 27*, 647–656.

Spitz, R. (1946). The smiling response: a contribution to the ontogenesis of social relations. *Genetic Psychology Monographs, 34*, 57–125.

Stern, D. (1977). *The first relationship: infant and mother.* Cambridge, MA: Harvard University Press.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: deficits in the early acquisition of a theory of mind. *Child Development, 63*, 161–172.

Tager-Flusberg, H. (1993). What language reveals about the understanding of minds in children with autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism.* Oxford: Oxford University Press.

Tan, J. & Harris, P. (1991). Autistic children understand seeing and wanting. *Development and Psychopathology, 3*, 163–174.

Tantam, D. (1992). Characterizing the fundamental social handicap in autism. *Acta Paedopsychiatrica, 55*, 88–91.

Thayer, S. (1971). Children's detection of on-face and off-face gazes. *Developmental Psychology, 13*, 673–674.

Thayer, S., & Schiff, W. (1977). Gazing patterns and attribution of sexual involvement. *The Journal of Social Psychology, 101*, 235–246.

Tinbergen, N. (1951). *The study of instinct.* Oxford: Oxford University Press.

Trevarthen, C. (1979). Communication and cooperation in early infancy: a description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech.* Cambridge: Cambridge University Press.

Van Hooff, J. (1962). Facial expressions in higher primates. *Symposium of the Zoological Society of London, 8*, 97–125.

Vine, I. (1973). The role of facial signalling in early social development. In M. von Cranach & I. Vine (Eds.), *Social communication and movement: Studies of men and chimpanzees.* London: Academic Press.

Volkmar, F., & Cohen, D. (1989). Disintegrative disorder or 'late onset' autism? *Journal of Child Psychology and Psychiatry, 30*, 717–724.

Wada, J. (1961). Modification of cortically induced responses in brainstem by shift of attention in monkeys. *Science, 133*, 40–42.

Warrington, E., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*, 829–854.

Wellman, H. (1990). *Children's theories of mind.* Bradford, MIT Press.

Whiten, A. (1993). Evolving a theory of mind: the nature of non-verbal mentalism in other primates. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: perspectives from autism.* Oxford: Oxford University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Wolff, P. (1963). Observations on the early development of smiling. In B. Foss (Ed.), *Determinants of infant behaviour* (Vol 2). New York: Wiley.

Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferior temporal cortex? *Experimental Brain Research, 73*, 209–214.

Submitted May 11, 1993
Accepted November 10, 1993