Joint Attention

Its Origins and Role in Development

Edited by
Chris Moore
Philip J. Dunham
Dalhousie University

With a Foreword by Jerome Bruner

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1995 Hillsdale, New Jersey Hove, UK

Chapter 3

The Eye Direction Detector (EDD) and the Shared Attention Mechanism (SAM): Two Cases for Evolutionary Psychology

Simon Baron-Cohen
University of Cambridge

> *To be a viable hypothesis about human psychological architecture, the design proposed must be able to meet both* solvability *and* evolvability *criteria [emphasis added]: It must be able to solve the problems that we observe modern humans routinely solving and it must solve all the problems that were necessary for humans to survive and reproduce in ancestral environments.*
>
> —Tooby & Cosmides, 1992, p. 110

Imagine that you are walking onto a crowded train. You see a remaining empty seat, so you go across and sit down. You get out your book, and settle into it. During the journey, you become aware of a feeling that someone is looking at you. You glance along the carriage and, sure enough, someone is looking at you. As soon as you make eye contact with this stranger, he looks away. To my mind, this phenomenon is rather striking in that it is not immediately obvious how you would have known that someone was looking at you if you were already engaged in another activity. One possibility is that as the train was going along you were occasionally rapidly scanning the other passengers' faces. Given that there was quite a crowd, perhaps your perceptual system only superficially processed those faces whose eyes were directed away from you, because to have processed each and every face in any deeper way would have led to information overload, and would be of little adaptive value. But among that sea of faces there was one whose *eyes were directed at you*, which your perceptual system processed in more detail. In this chapter, I discuss the evidence that "eyes looking at me" are especially salient.

In doing this, I pursue the idea that evolution has produced a neurocognitive system for the rapid detection of the eyes of another organism, because of the considerable adaptive significance of such a mechanism. This system is called the *Eye Direction Detector*, or *EDD*. Its ontogenesis in humans is described more fully elsewhere (Baron-Cohen, 1994, 1995), as is its neurobiology (Baron-Cohen & Ring, 1994). In this chapter, the focus is on its evolution, and on its relationship to a second mechanism, the *Shared Attention Mechanism*, or *SAM*. This second mechanism functions to verify if you and another organism are attending to the same thing.[1] I argue that SAM is not only essential for joint attention, but also for the development of a "theory of mind" in the human child.

Back to EDD. To get more of a flavor of the evolution of EDD, consider its use in a second example. This time the scenario involves an encounter between Alex, who has just entered a new social group, and Thalia, with whom Alex is keen to become acquainted:

> Alex stared at Thalia until she turned and almost caught him looking at her. He glanced away immediately, and then she stared at him until his head began to turn toward her. She [quickly looked toward the ground], but as soon as Alex looked away, her gaze returned to him. They went on like this for more than 15 minutes, always with split-second timing. Finally, Alex managed to catch Thalia looking at him. He made the friendly eyes . . . [and then] approached [her]. (Leakey & Lewin, 1992, pp. 287–288)

You could be forgiven for assuming that this couple was human; in fact, Alex and Thalia are members of a troop of baboons that live near Eburru Cliffs, 100 miles northwest of Nairobi, on the floor of the Great Rift Valley. The aforementioned observation was made by Barbara Smuts, a primatologist at the University of Michigan, who studied the social life of this troop over several years. Discussing this particular social interaction, Smuts comments, "It was like watching two novices in a singles bar" (Leakey & Lewin, 1992, p. 288).

In this example we see the split-second timing of eye direction detection, as one animal checks to see if the other is looking at him or her, and as each animal tries to avoid the other animal's being aware that they are checking this. Like the baboon, humans are very accurate at identifying when someone is looking at them. Given that there may be only a small difference in geometric angle between *you* being looked at and *something next to you* being looked at, the psychophysics of EDD are likely to be impressive. (This remains to be studied in detail in different species). The example with Alex and Thalia also shows how in some species there is a risk that eye contact from a stranger can be

[1]This owes a lot to those who have drawn attention to the primitive nature of joint attention (e.g., Bruner, 1983; Butterworth, 1991; Hobson, 1993; Tantam, 1992; Tomasello, Kruger, & Ratner, 1994). As will become apparent, my proposal is somewhat different from these in suggesting the modularity of this mechanism, and in discussing the unique class of representations it processes.

misinterpreted as threatening, and how, if one's intentions are prosocial as Alex's were, eye contact must be offered in small doses at the start of such an encounter, if this risk is to be avoided.

EVOLUTIONARY PSYCHOLOGY

In my opening two examples, I jumped from humans to baboons. This was a deliberate strategy to help introduce the approach being adopted here, that of evolutionary psychology. This is defined as "psychology informed by the fact that the inherited architecture of the human mind is the product of the evolutionary process" (Cosmides, Tooby, & Barkow, 1992, p. 7). This phrase should of course strike a chord with the more well-known phrase "evolutionary biology," which has transformed the science of biology. As Cosmides et al. suggest, the time is ripe to integrate psychology with biology via evolutionary theory.

In this spirit, the adaptive problem I initially discuss is the rapid prediction of another organism's next action. By adaptive problem I mean a "problem whose solution can affect reproduction, however distally" (Cosmides et al., 1992, p. 8). It is evident, I think, that if another organism's immediate goal is that you should be its lunch, it would pay to know about it quickly! This is one key function of EDD because "eyes directed at me" is something it was evolved specifically to detect, and because eye direction tends to specify the object upon which an agent's next action is targeted. Of course, just because another organism is looking at you, it does not mean the goal of its next action is necessarily predatory or aggressive. It may be interested in you for some other reason (e.g., to signal that it wants to communicate with you, or that it is sexually attracted to you, etc.).[2] However, recognizing any of these motives in another organism is also likely to affect reproduction, however distally.

The Environment of Evolutionary Adaptedness (EEA): An Intensely Social Environment

I return to the specific evolution of EDD shortly. First, I want to say something about evolutionary psychology more broadly. Cosmides et al. (1992) remind us that modern human history is only about 10,000 years old (if one dates this from the advent of agriculture, for example), and that nothing about our biology is likely to be an adaptation to changes during this period. In contrast, the period relevant to human evolution can be seen as spanning two phases: the Pleistocene epoch (roughly the last 2 million years, when we lived as hunter–gatherers), and then several hundred million years before that (when we lived as one kind of

[2] I thank Alison Gopnik for suggesting that I invoke the old joke here: that EDD allows one to predict that another animal is interested in you for one of the three F's: fight, food, or sex.

forager or another). They point out that these time spans are important to keep in mind, because "they establish which set of environments and conditions defined the adaptive problems the mind was shaped to cope with: Pleistocene conditions, rather than modern conditions " (Cosmides et al., 1992, p. 5). That there was massive neurocognitive evolution during the Pleistocene epoch is beyond any doubt. Figure 3.1, for example, shows that a threefold increase in brain size occurred in the 3 million years since *Australopithecus afarensis* evolved, going from around 400 cubic cm to its current size of about 1,350 cubic cm.

The causes of the brain size increase are likely to reflect many factors, but one major factor upon which many theorists agree was the need for greater "social intelligence" (Brothers, 1990; Dunbar, 1993; Humphrey, 1984). This is because the vast majority of primate species are social animals, living in groups that range from as few as two individuals to as many as 200. If one looks at different kinds of social organization in existing primates, one is struck by the variety: from lifelong monogamy (e.g., the gibbon), to unimale polygyny (e.g., the gorilla, where a single male has control over a group of females and their offspring), to multimale polygyny (e.g., the chimpanzee, where several males cooperate to defend a group of widely distributed females and their offspring), to "exploded" unimale polygyny (e.g., the orangutan, where a single male defends

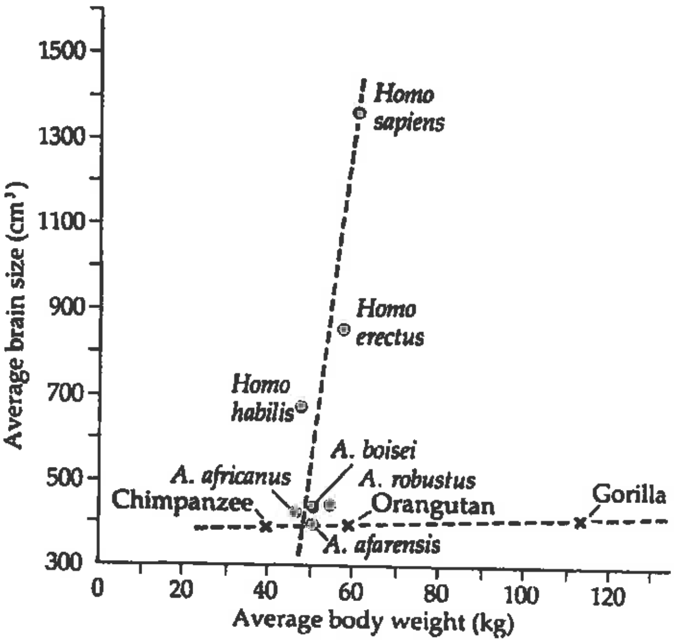

FIG. 3.1. Changes in brain size over the last 3 million years. From *Human Evolution*, p. 127, by R. Lewin (1989), Oxford, England, Blackwell Scientific Publications. Reprinted with permission.

a group of females and their offspring, but where the females do not live as a group but instead are distributed over a wide area).[3]

These different models hint at how primate social environments are enormously more complex than those of other social groups, even those that live in comparably sized groups, such as the antelope. The difference in complexity lies in the nature of the social interactions. "The group is . . . the center of intense social interaction that has little apparent direct bearing on the practicalities of life: In the human sphere we would call it socializing, the making and breaking of friendship and alliances" (Lewin, 1992, p. 46). The challenge for the primate was (and remains) to predict and manipulate the behavior of others in the group, what Humphrey (1984), and Byrne and Whiten (1988) depict as the Machiavellian nature of social interaction. It is this social intelligence that determines who wins higher status. Consider Lewin again on this point:

> When you observe other mammal species and see instances of conflict between two individuals, it is usually easy to predict which one will triumph: the larger one, or the one with the bigger canines or bigger antlers, or whatever is the appropriate weapon for combat. Not so in monkeys and apes. Individuals spend a lot of time establishing networks of "friendships", and observing the alliances of others. As a result, a physically inferior individual can triumph over a stronger individual, provided the challenge is timed so that friends are at hand to help the challenger and whilst the victim's allies are absent. (Lewin, 1992, p. 129)

Leakey and Lewin (1992) reached a similar conclusion: "In higher primates, the greatest reproductive success (in both males and females) is shaped much more by social skills than by physical displays, either of strength or appearance" (p. 293). So, the evolution of primates can be characterized by an increase in the complexity of social interaction, requiring on the cognitive level an increase in rapid and adaptive social intelligence, and on the biological level an increase in different brain mechanisms, to support this. I argue that EDD and SAM are two such brain mechanisms.[4]

In evolutionary psychology, then, a claim about an evolved neurocognitive system needs to be considered in the light of its adaptedness. Again, in standard biological terminology, a system with proposed adaptedness should have design features that would "exploit the enduring properties of the environment in which it evolved (termed its environment of evolutionary adaptedness, or EEA) and . . . solve the recurring problems posed by that environment" (Tooby & Cosmides, 1992, p. 69). Here, the recurring problem that the organism would need

---

[3]This survey comes from Lewin (1989).

[4]Another such brain mechanism that is postulated to have evolved for social intelligence is the *Theory of Mind Mechanism* (or *ToMM*; Leslie, 1991). This is touched upon at the end of this chapter, though because this book concentrates on joint attention, it is not discussed in depth here. For further details, see also Leslie (1994), [illegible]

to solve is predicting another organism's next move, and the enduring property of the environment that natural selection is postulated to have exploited to solve this problem is that an animal's eye direction reliably correlates with its next action. Animals, human beings included, tend to look at what they are about to act upon.[5]

EDD

From Reptiles to Humans. EDD can be likened to other visual mechanisms. For example, the retina of the rabbit has a "hawk detector" (Marr, 1982) and the retina of the frog has a "bug detector" (Barlow, 1972, cited in Tooby & Cosmides, 1992). I propose that EDD fires in response to eye-like stimuli, and fires most strongly if the direction of these eye-like stimuli are "looking at me." In this section I give a sketch of the range of species that show fine sensitivity to eyes and eye direction.

Ristau (1990, 1991) carried out some elegant experiments with plovers, to test whether 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 meters 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 toward the dunes than when the intruder was gazing toward 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 were capable of detecting whether an intruder was looking at them on their nest, and that the birds reacted 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 1 m 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 (Hennig, 1977), the blue

[5]If you need a more concrete image of this generalization, picture where your cat looks before she pounces on the ball of wool you teasingly dangle in front of her.

crab (O'Brien & Dunlap, 1975), and ducks (Sargent & 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 than by 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 behaviors (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 1.5 m, whether full face or half-profile), than to a human face tilted backward.

Mutual gaze, particularly in the form of a stare, is a well-documented component of threatening displays in many nonhuman primates, for example, adult male baboons (Hall & Devore, 1965), gorillas (Schaller, 1964), macaques (Altmann, 1967), and a number of other old world monkeys and apes (van Hooff, 1962). Chance (1967) described how struggles for dominance are often only ended with one animal averting its gaze, what he calls a "visual cutoff," possibly as a mechanism for reducing the physiological arousal produced by direct gaze (Nichols & Champness, 1971; Wada, 1961). This array of studies showing a fear response to eye direction is one source of evidence in support of the notion that there is an eye direction detector, one function of which is to detect threat. Note that recognizing that another animal's eyes are directed at you need not only be a form of threat detection. It may also signal to you that the other animal has you as its target for prosocial reasons. For example, eye contact occurs as part of grooming, greeting, and play facial expressions in old world monkeys and apes (van Hooff, 1962). However, Argyle and Cook (1976), in their important review of the literature, conclude that it is only in primates that gaze functions as an affiliative as well as an aggressive cue.

The Neurophysiology of EDD. Specific cells in the Superior Temporal Sulcus (STS) of the monkey brain appear to respond to the perspective view of the head (Perrett et al., 1985). Perrett's group found that different cell types in the STS of the macaque monkey brain respond selectively to the different views of the head, some to the left profile, others to the back of the head, and so forth. The STS is shown in Figure 3.2. Other studies have found that specific STS cells respond selectively to direction of gaze (Perrett et al., 1985; Perrett et al., 1990). For example, Perrett et al. (1985) found that 64% of the cells responsive to the face or profile views of the head were also selective for the direction of gaze. Other evidence for the neurophysiological basis of eye direction detection comes from neuropsychology. Lesions in the STS produce an impairment in the ability to discriminate gaze direction by monkeys (Campbell, Hey-

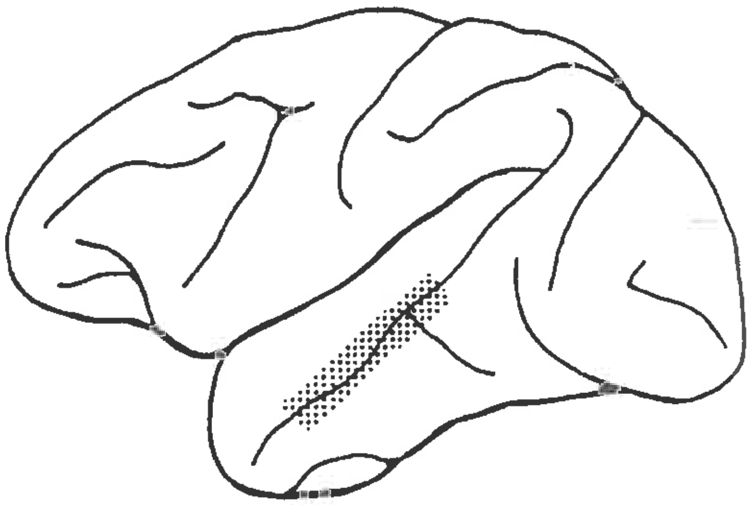

FIG. 3.2. The Superior Temporal Sulcus (STS: lateral view), in the macaque brain. From "The social brain: A project for integrating primate behaviour and neurophysiology in a new domain," by L. Brothers (1990), *Concepts in Neuroscience, 1*, 27–51. Reprinted with permission.

wood, Cowey, Regard, & Landis, 1990). Some patients with prosopagnosia are also impaired in this ability (Campbell et al., 1990; Heywood & Cowey, 1992). Perrett and his colleagues, in their most recent publications (e.g., Perrett, Hietanan, Oram, & Benson, 1992) refer to the cells in the STS that respond to gaze direction as cells responsive to the direction of attention of the other individual, and that have the primary function of detecting where another individual is looking. These studies suggest that EDD may be localized either regionally within the brain, or within specific neural circuits.

EDD's Representations: Dyadic. In the human case, I suggest that whenever EDD detects eyes, it gradually builds representations of eye behavior. These I call *dyadic representations.*[6] They are constructed by at least 4 months of age in humans, if not sooner (Johnson & Vicera, 1993). Dyadic representations specify the presence of two entities (Agent and Self; or Agent and Object; or Self and Object) standing in a relation to each other. Dyadic representations thus have one of four forms:

1. *[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-sees-me], or
[I-see-Mummy].

2. *[Agent–Relation–Proposition]*. Here, the relation term is unidirectional, because one of the elements is not an agent. An example of this form is:
   [Mummy-sees-the bus].
3. *[Agent$_1$–Relation–Agent$_2$]*. Here, again, the relation term is bidirectional, because both elements are agents. Examples of this are:
   [Mummy-sees-Daddy], or
   [Daddy-sees-Mummy].
4. *[Self–Relation–Proposition]*. Here, the relation term is unidirectional. So an example of this form is:
   [I-see-the house].

The form of these dyadic representations presumes that the organism has a concept of *seeing*, a concept of *self*, and even a concept of *itself seeing*. All of these concepts could be fairly simple. For example, seeing at the simplest level could be an awareness of the distinction between "eyes open" (light) and "eyes closed" (dark), which presumably even young infants have. Self could be quite low level too, namely, an awareness of the distinction between dual stimulation (when you touch yourself) and single stimulation (when you touch anything else; Gallup, 1970).

In summary, the first adaptive problem to which EDD is a solution is identifying whether another organism is looking at you, because of the obvious benefits of detecting this. There is, however, a second adaptive problem that would be worth solving: if the other animal is not looking at you, then what is he or she looking at? Might he or she have spotted something that would be worth knowing about? A food source, a rival, a mate, or a predator, for example? I have assumed that EDD can perform some geometric analysis to compute the direction that another animal is looking in, because this also only requires dyadic representation. Chimpanzees, for example, have been observed to use another animal's eye direction to search for a hidden object in a location being looked at (Menzel & Halperin, 1975). However, this is just about the limit of EDD's power. To see how EDD is used in more complex ways, I need to introduce the second mechanism, SAM.

SAM

Triadic Representations. The Shared Attention Mechanism (or SAM) functions to identify if you and another organism are both attending to the same thing. This is a further adaptive problem that it would be important to be able to solve, but EDD just cannot solve it. This is because EDD is limited to building dyadic representations, and according to my analysis, what is needed in order to verify if you and someone else are both attending to the same thing is what I call a *triadic* representation.[7] These are the special representations that SAM alone can build. Triadic representations differ in structure from 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:

1. *[Self–Relation–(Agent–Relation–Proposition)]*. 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)].
   Because this representation specifies that both I and Mummy are seeing the same bus at the same time, this fulfills the function of the triadic representation, namely, to identify shared attention.
2. *[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 formal descriptions 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 Figure 3.3, may be both more comprehensive, and simpler to "read." In the text, however, I use the aforementioned formal descriptions, because these are adequate for present purposes.

As is clear, the embedded term in triadic representations is the main advance over dyadic representations. It is this term that enables the animal to specify within the triadic representation that they are looking at the same object that the other animal is looking at. My theory proposes that the capacity to construct triadic representations is necessary for joint attention, and suggests that in the first instance, SAM builds these representations using dyadic representations it obtains from EDD's output. This is because triadic representations can be built more easily in the visual modality than can be done in other modalities. However, in children born blind, but in whom SAM is otherwise intact, shared attention can be achieved by building triadic representations via touch or audition, thus establishing joint tactile or joint auditory attention. This is not easy, but in principle possible.

The clearest pieces of evidence for SAM's functioning are in gaze monitoring, and in the "protodeclarative" pointing gesture (Bates, Benigni, Bretherton,

[7]This term is also 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

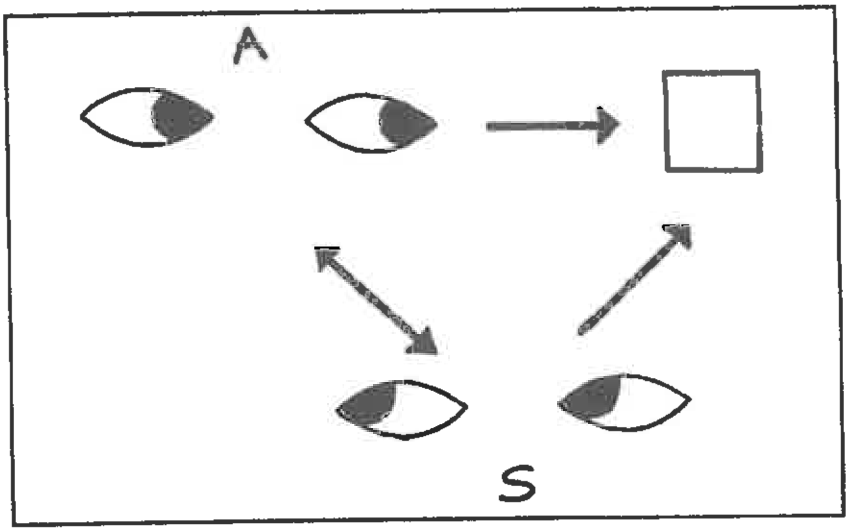

FIG. 3.3. A spatial description of the triadic representation [Agent–Relation–(Self–Relation–Proposition)]. Reproduced from "How to build a baby that can read minds: Cognitive mechanisms in mindreading," by S. Baron-Cohen (1994), *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *13*(5), 513–552. Reprinted with permission.

Camaioni, & Volterra, 1979), that is, the joint visual attention behaviors. 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 experimentally investigated (Cheyney & Seyfarth, 1990). Certainly, spontaneous gaze monitoring emerges clearly in human infants by the end of the first year, and is universally established between 9 and 14 months of age (Butterworth, 1991, this volume; Corkum & Moore, this volume; Desrochers, Morissette, & Ricard, this volume; Scaife & Bruner, 1975). Similarly, protodeclarative pointing emerges during the same period in normal development (Bates et al., 1979; Butterworth, 1991). Gaze monitoring continues to be present in human children (Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1993). The communicative function of gaze in human adults has also been amply documented (Argyle & Cook, 1976).

SAM's Role in the Development of a Theory of Mind. My suggestion is that in the human case, SAM has two other key functions. First, it imports volitional terms like "want" or "goal" into the relation slot of its triadic representations. These are some of the earliest mental state concepts that the human child possesses (Wellman, 1990). SAM can do this because volitional terms are themselves relation terms. It is adaptive to do this because it gives the child a way of inferring a person's desires and goals from the direction of their gaze. Secondly, SAM plays a necessary though not sufficient role in triggering a later mechanism, ToMM (or the *Theory of Mind Mechanism*; see Leslie, 1991).

Regarding evidence for SAM's function of inferring a person's goal from eye direction, Phillips, Baron-Cohen, and Rutter (1992) investigated this with normal infants ranging from 9 to 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), whereas 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 3–4-year-olds were asked "Which sweet will Charlie take?" after being shown a display of four sweets and with Charlie's face looking at one of these confections, they tended to pick the one he was looking at as the goal of his next action (see Fig. 3.4).

Regarding SAM's function of 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 four sweets, and placed the cartoon face of Charlie in the center of the display. Again, Charlie's eyes were depicted as pointing toward one of the four sweets, randomly selected (see Fig. 3.4). The subject was asked "Which one does Charlie want?". In another condition, 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, this volume) has also reported 18-month-olds' ability to use eye direction to infer a person's intended referent.

In summary, there is growing evidence that is consistent with the idea that in the human case SAM identifies when shared attention has been achieved, and infers a person's goal, desire, and intention to refer, from eye direction. Regarding SAM's function as a necessary (but not sufficient) precursor to a theory of mind, the evidence comes from autism.[8]

---

[8]In this chapter I simply mention one precursor to a theory of mind: joint attention. However, our earlier work (Baron-Cohen, Allen, & Gillberg, 1992) has shown that the combination of joint attention deficits and pretend deficits at 18 months predicts cases of autism. It is therefore likely that this same combination of deficits predicts impairments in theory of mind. If this is confirmed, then it suggests that both joint attention and pretend play may be precursors to theory of mind. Quite why these two behaviors should lie in this intimate relationship is not yet clear

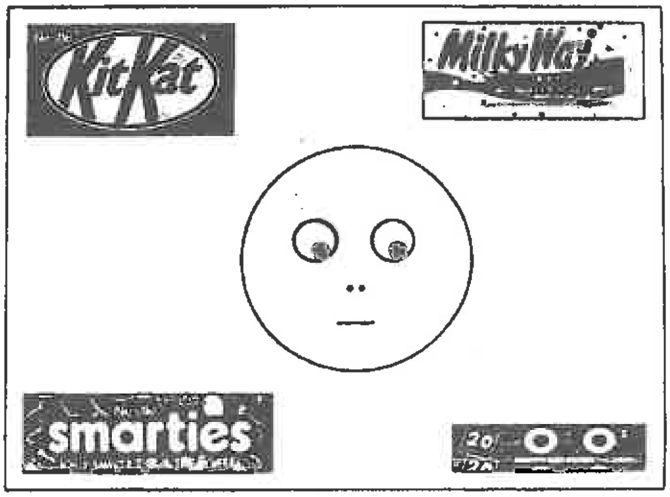

FIG. 3.4. The Four Sweets Display. From "Are children with autism blind to the mentalistic significance of the eyes?", by S. Baron-Cohen, R. Campbell, A. Karmiloff-Smith, J. Grant, & J. Walker (in press), *British Journal of Developmental Psychology*.

EDD, SAM, and Autism

EDD appears to be intact in children with autism, while SAM is in almost all cases impaired. Thus, children with autism are able to detect whether eyes are looking at them (Baron-Cohen et al., in press), but they show little if any joint-attention behaviors (Baron-Cohen, 1989a; Leekam et al., 1993; Phillips et al., 1992; Sigman, Mundy, Ungerer, & Sherman, 1986). This dissociation is consistent with the idea that children with autism can use EDD to build dyadic representations, but are impaired in using SAM to build triadic representations. If this is the case, then they should also fail to go on to import volitional mental state terms (*goal*, *desire*, and *intention-to-refer*) into triadic representations. There is some evidence consistent with these assumptions.

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 Four Sweets Task and found significant impairments in the use of eye direction in inferring want, goal, and refer. For the case of refer, this has also been replicated using Baldwin's (1991) paradigm by Baron-Cohen, Baldwin, and Crowson (1995).

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 (Baron-Cohen, 1989a, 1991b; Hobson, 1984; Tan & Harris, 1991)? One possibility is that they do this by employing dyadic representations of the form [Agent–Relation–Proposition]. As mentioned earlier, what is missing from these is the embedded term that is a necessary feature of triadic representations, and the possibility of employing other mental state terms in the relation slot. This would explain why they do not show spontaneous gaze monitoring, or attempt to direct another person's attention to an object, as an end in itself.

It follows from my earlier claim about the precursor relation between SAM and theory of mind that I see the joint-attention deficits in autism as developmentally related to their later theory of mind deficits.[9] The theory of mind deficits in autism is not reviewed here, but some references to this literature are Baron-Cohen, Leslie, and Frith (1985); Perner, Frith, Leslie, and Leekam (1989); Baron-Cohen (1989c, 1989d, 1990, 1993); Leslie and Roth (1993); and Frith (1989). Suffice it to say that children with autism appear severely and selectively impaired in many tasks that require judging another person's mental state. If this account is correct, it implies that the origins of theory of mind may lie in the apparently simple acts of joint attention in infancy, the origin of which may itself lie deep in the evolutionary history of the brain.

SUMMARY

In this chapter I described two neurocognitive mechanisms that have each evolved to solve a different, key adaptive problem. One is concerned with identifying if you are the target of another organism's attention, and the other is concerned with establishing a shared focus of attention with another organism. The first of these is achieved by building dyadic representations specifying whether the eyes of another organism are directed toward you or not. This can function both as an "early warning system" that a predator has you (literally) within its sights, or that it has other (more prosocial) designs on you. EDD has evolved to build such dyadic representations, and to build them rapidly. The second problem, establishing a shared focus of attention with another organism, requires the construction of triadic representations, which SAM has evolved specifically to build. Although these representations could in principle be built in a number of different modalities, it is far easier in the visual modality, hence SAM exploits EDD to perform this function.

However, EDD and SAM have evolved over vastly different time scales.[10] EDD is very old (indeed, work reviewed in this chapter identifies it in reptiles and birds, as well as primates, for example), whereas SAM appears only com-

---

[9]See Baron-Cohen (1989a, 1991a), Tantam (1992), and Sigman et al. (1986), for similar arguments.

[10]I am grateful to the editors of this volume for helping me clarify this.

paratively recently in evolution. It is best documented in humans, though it remains possible that it is present in some of the higher primates. These two mechanisms probably evolved in different EEA's: EDD is likely to have evolved in environments in which sighted organisms are preyed upon by other sighted organisms. Clearly this encompasses a large set of species, independent of how social such species are. SAM, on the other hand, is likely to have evolved only in highly social species because the benefits of SAM would be more obvious in such cases. Finally, in the human case, SAM appears to play a crucial role in the ontogenesis of a theory of mind.

ACKNOWLEDGMENTS

During the preparation of this work I was a member of the Departments of Child Psychiatry and Psychology, The Institute of Psychiatry, University of London. I am grateful for the support from the Mental Health Foundation and the Medical Research Council. Parts of this paper first appeared in Baron-Cohen (1994), and were presented at the Society for Research in Child Development Conference, New Orleans, March, 1993, the British Psychological Society, Welsh Branch, "Faces" Conference, University of Wales College of Cardiff, September, 1993, and the British Society for the Philosophy of Science "Roots of joint reference" Conference, University of Bristol, November, 1993. Chris Moore, Phil Dunham, and Alison Gopnik all gave me valuable feedback on the first draft of this chapter. I am also grateful to Digby Tantam, Leslie Brothers, Dave Perrett, Annette Karmiloff-Smith, Pierre Jacob, and Ruth Campbell for discussions of these ideas.

REFERENCES

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

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

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

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.

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

Baron-Cohen, S. (1989b). Joint attention deficits in autism: Towards a cognitive analysis. *Development and Psychopathology*, *1*, 185–189.

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. (1989d). 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. (1990). Autism: A specific cognitive disorder of "mind-blindness." *International Review of Psychiatry, 2*, 79–88.

Baron-Cohen, S. (1991a). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind* (pp. 233–252). Oxford, England: Basil Blackwell.

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. (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. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 59–82). Oxford, England: Oxford University Press.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive, 13*(5), 513–552.

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press/Bradford Books.

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., Baldwin, D., & Crowson, M. (1995). *Do children with autism use eye-direction to infer linguistic reference?* Unpublished manuscript, University of Cambridge, Department of Experimental Psychology.

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*, 173–186.

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., & 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* (pp. 183–210). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Bayliss, G., Rolls, E., & Leonard, C. (1985). Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Research, 342*, 91–102.

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

Bruce, C., Desimone, R., & Gross, C. (1981). Visual properties of neurones in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*, 369–384.

Bruner, J. (1983). *Child's talk*. Oxford, England: 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* (pp. 223–231). Oxford, England: Basil Blackwell.

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

Byrne, R., & Whiten, A. (1991). Computation and mindreading in primate tactical deception. In A. Whiten (Ed.), *Natural theories of mind* (pp. 127–142). Oxford, England: Basil Blackwell.

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.

Chance, M. (1956). Social structure of a colony of *Macaca mulatta*. *British Journal of Animal Behaviour, 4*, 1–13.

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.

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

Cosmides, L., Tooby, J., & Barkow, J. (1992). Introduction: Evolutionary psychology and conceptual integration. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 3–18). New York: Oxford University Press.

Desimone, R., Albright, T., Gross, C., & Bruce, C. (1984). Stimulus selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 8, 2051–2062.

Dunbar, R. (1993). Coevolution of neurocortical size, group size and language in humans. *Brain and Behavioral Sciences, 16*, 681–735.

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

Gallup, G. (1970). Chimpanzees: Self recognition. *Science, 167*, 341–343.

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.

Gomez, J. C. (1991). Visual behaviour as a window for reading the minds of others in primates. In A. Whiten (Ed.), *Natural theories of mind* (pp. 195–208). Oxford: Basil Blackwell.

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

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.

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* (pp. 1–128). Oxford, England: Oxford University Press.

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

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

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

Johnson, M., & Vicera, S. (1993). Cortical parcellation and the development of face processing. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life.* Dordrecht, Netherlands: Kluwer.

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

Kendrick, K., & Baldwin, B. (1987). Cells in the temporal cortex of conscious sheep can respond preferably to the sight of faces. *Science*, 236, 448–450.

Leakey, R., & Lewin, R. (1992). *Origins reconsidered.* London: Little, Brown, & Company

Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1993). *Eye-Direction Detection: A dissociation between geometric and joint-attention skills in autism.* Unpublished manuscript, University of Kent, Canterbury, UK.

Leslie, A. (1991). The theory of mind impairment in autism: Evidence for a modular mechanism of development? In A. Whiten (Ed.), *Natural theories of mind* (pp. 63–78). Oxford, England: Blackwell.

Leslie, A. (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, England: Cambridge University Press.

Leslie, A., & Roth, D. (1993). What autism teaches us about metarepresentation. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 83–111). Oxford, England: Oxford Medical Publications.

Lewin, R. (1989). *Human evolution.* Oxford, England: Blackwell Scientific Publications.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

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

Menzel, E., & Halperin, S. (1975). Purposive behavior as a basis for objective communication between chimpanzees. *Science*, *189*, 652–654.

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*, Rathbun): Its relation to threat of predation. *Journal of Comparative and Physiological Psychology*, 89, 86–94.

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., 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., 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, & D. Perrett (Eds.), *Processing the facial image: Philosophical Transactions of the Royal Society of London*, *B 335* (pp. 1–128). Oxford, England: Oxford University Press.

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: Wiley.

Perrett, D., Oram, M., Harries, M., Bevan, R., Hietanen, J., Benson, P., & Thomas, S. (1991). Viewer-centered and object-centered codings of heads in the macaque temporal cortex. *Experimental Brain Research*, 86, 159–173.

Perrett, D., Rolls, E., & Cann, W. (1982.) Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, *47*, 329–342.

Perrett, D., Smith, P., Potter, D., Mistlin, A., Head, A., Milner, A., & Jeeves, M. (1984). Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity, and relation to perception. *Human Neurobiology*, 3, 197–208.

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 the detection of goals: Evidence from normal toddlers, and children with autism or mental handicap. *Development and Psychopathology*, *4*, 375–383.

Ristau, C. (Ed.). (1990). *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* (pp. 209–223). Oxford, England: Basil Blackwell.

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., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265–266.

Schaller, G. (1964). *The mountain gorilla*. Chicago: 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.

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.

Tomasello, M., Kruger, A., & Ratner, H. (1994). Cultural learning. *Behaviour and Brain Sciences*.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). New York: Oxford University Press.

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

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

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

Wellman, H. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

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