2

Self-Reflective Consciousness and the Projectable Self

Janet Metcalfe & Hedy Kober

In 1619, while secluded in his poele, Descartes undertook to discover the indisputable source of all knowledge, the unshakable foundation upon which he could base his philosophy with certainty. Many years later, in his 1637 *Discourse on Method*, he reported that the one thing that he was unable to doubt, which became this foundation, is something that today we might dub self-reflective consciousness, metacognition, secondary representation, or autonoetic consciousness—the reflection of the self upon its own thoughts, memories, mental processes, and other possible worlds, including the ability to mentally project oneself outside the boundaries of one's immediate stimulus environment and thereby entertain counterfactuals. He affirmed that in order to have self-reflection of this sort, one must have a self (cf. Russell, 1945/1972). "But what then am I? A thing that thinks. What is that? A thing that doubts, understands, affirms, denies, wills, refuses, and that also imagines and senses" (p. 66). In this chapter, we refer to the kind of self-reflective consciousness so aptly captured in Descartes's meditations and discourse as the *projectable self*. It can, as Descartes noted, perceive the present and understand, affirm, and deny; but it can also be projected into the past to allow episodic memory (Tulving, 2002; Wheeler, Stuss, & Tulving, 1997; chapters 1 and 7, this volume). It can imagine the future and itself in it. It can imagine other possible perspectives, including, importantly, the points of view of other people, both physically and also, even more interestingly, mentally and emotionally. The potential to project oneself in outer and inner space and time allows capabilities familiar to people, but perhaps rare or unknown in other animals: reminiscence, planning and scheming, connection to other people through empathy and understanding, and manipulation of other people through lies and deception.

Descartes equated this self-reflective ability with the soul and argued, primarily on religious grounds (but with the uniqueness of language being his empirical evidence), that only humans have it. Without it, a creature is a "mere" mechanism, regardless of the complexity of the behavior. While taking no stand on the ineffability of the soul, we here argue that the projectable self is the singular evolutionary adaptation underpinning the most advanced achievements of humans including our culture—without doubt what Dennett (1991) would call a Good Trick.

Humphrey (2003) advanced that we humans have what he unashamedly called an inner eye. This inner eye could look down on more basic cognitions and interpret them in a user-friendly way, that is, as motivations, feelings, goals, hopes, intentions, fears, thoughts, memories, and so on (rather than as, say, p300s, serotonin imbalances, hippocampal activation, or reverberating Hebb nets). He was specific about noting that this self-reflective capability is an evolutionary advance (though he did not specify exactly at what point it came into being). He allayed concern about the problem of having to postulate a homunculus within a homunculus within a homunculus, saying: "The problem of self-observation producing an infinite regress is, I think, phony. No one would say that a person cannot use his own eyes to observe his own feet. No one would say, moreover, that he cannot use his own eyes, with the aid of a mirror, to observe his own eyes. Then why should anyone say a person cannot, at least in principle, use his own brain to observe his own brain?" (p. 11).

Humphrey (2003) noted that one could exhibit many skills without this kind of self-reflective consciousness, as in the top panel of figure 2.1. For example, one might well be able to drive, play the piano, protect one's young, or do many of the things that people have been shown to be able to do automatically without input from the inner eye. An issue that is investigated in much of this book is whether or not much highly complex behavior of nonhuman primates might be possible without this kind of consciousness and what behavior actually requires self-reflective consciousness. On the phenomenology—the transcendent or otherworldly qualities of this kind of consciousness—Humphrey demurs, alluding only to "one curious feature: the output of the inner eye is part of its own input. As I expect you know, a self-referential system of this sort may well have strange and paradoxical properties—not least that so-called 'truth functions' go awry" (p. 12).

So why would such an inner eye evolve? The answer given by Humphrey, much like that of Terrace, Nelson, Higgins, and others in this volume, is that people are intensely social, as a species. Having an inner eye—which gives a quick and accessible description of how one feels and thinks oneself, what one wants, plans, and fears—may

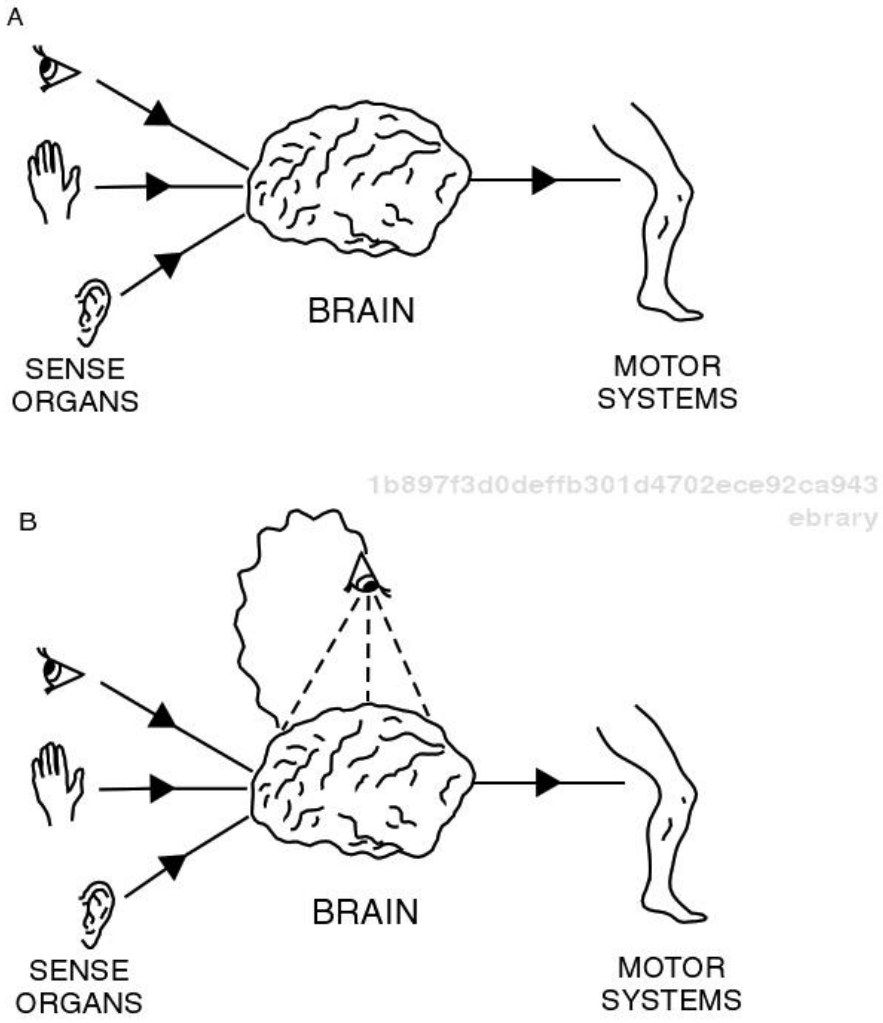

Figure 2.1 A. Diagrammatic representation of an entity lacking insight. B. Diagrammatic representation of an entity possessing insight. Illustrations courtesy of N. Humphrey.

allow one to make similar attributions to other people, and quickly. Facility at predicting the behavior of others, by getting inside their minds, provides an enormous social advantage. A person with this ability highly honed is attractive as a mate and formidable as a foe. This social aspect of the projectable self has come to be known as theory of mind, and some of the literature concerned with this specific issue is reviewed below.

A related just-so story about the evolution of the projectable self is that "one can let one's hypotheses die instead of oneself." The ability

to imagine possible future situations with oneself in them in considerable detail, and to play out possibilities in the mental rather than the physical world, allows a person to consider disastrous alternatives without physical consequences, and hence to find new solutions that might otherwise be impossible. Thinking, in this sense of simulating events—both social and physical—has clear advantages for a species that can do it, and some explorations of this are given in chapter 8. Impulsiveness due to the lack of projecting the self in this way (i.e., acting without first thinking) in humans is considered a mental disorder.

Finally, the ability to look back on the details of one's personal past, to be able to revisit what happened to oneself in the past via episodic memory (rather than merely having one's semantic memory or automatic responses changed, without the possibility of reflection on the earlier events that caused the changes), also may enhance survival. For example, one could update current judgments and courses of action in the light of new information about the causes of events in the past. Such malleability seems highly adaptive and also entertaining. In the print shown in figure 2.2, by the Japanese master Hiroshige, a powerful samurai is seen in three panels: the past, the present, and the future. He is portrayed as looking fiercely into his own personal past (but gesturing toward and guarding his much-hoped-for future protectively). He contemplates the ghosts of his now-vanquished enemies, which are multistable figures—at one moment innocuous snow-laden trees, at the next fearsome skulls that come back to haunt him. This time travel into his past, and the flexibility of the interpretation of that past from different vantage points, as is characteristic of episodic memory, enhances the beauty of his imagined future, portrayed as the beyond-lovely courtesan housed in the exquisite palace in the right panel. Episodic memory (see Nelson, 2000; chapter 4, this volume) constitutes the foundation for one's personal self, with all three panels, as well as the narrative that joins them. Such time travel into one's own personal past, also, according to Tulving (chapter 1, this volume), provides the basis for all future planning and is the mental foundation of human culture.

Formally, the idea that people have an inner eye that can look at other cognitive functions and content is the same as the description, agreed upon by the field, given by Nelson and Narens (1994) for metacognition (and see chapter 12, this volume). They propose that there are at least two levels, one of which is considered to be basic cognition, and the other "meta"—at a higher level looking down upon and making judgments about the events happening at the basic level. When they are judgments about events in the world, of brightness or numerosity, say, rather than about entities of the mind, such as memories or beliefs, then they are not properly called metacognitions. To be

Figure 2.2 *Taira Sees the Ghosts of His Enemies*, by Ando Hiroshige. From the collection of J. Metcalfe. See color insert.

in a position to have metacognitions, an animal must first have cognitions (i.e., mental representations). But the issue at stake here is not whether other primates have cognitions (which we assume they do) but rather whether they can reflect upon them. In Nelson and Narens's scheme, there is a feedback loop whereby the metacognitive level can change or modulate what goes on at the basic level. This feedback allows the person to have control over basic-level thoughts.

Some kinds of metacognition need not require a full-blown projectable self. For instance, metacognitive judgments might be made about semantic knowledge (bona fide representations, but not personal ones). The inner eye could stay stationary, in space and time, to make such judgments. Metacognitions might also be made about events involving the self over time past and future, however, or involving what it is like to be another person, and such judgments would involve projection. Insofar as having an inner eye that looks at one's own cognitions is true self-reflection, investigation of metacognition, even in the nonprojective form, can be informative concerning the evolution of the projectable self. Presumably the ability to make even simple metacognitive judgments is a precursor to full-blown projective consciousness. It may, indeed, be this particular capability in nonhuman primates that provides the most incisive clues to the evolution of our human consciousness, since it may have been the first to manifest itself (see also chapters 3 and 12, this volume).

In this chapter, we propose that the emergence of the projectable self is the central evolutionary advance culminating in human consciousness. For this conjecture, we owe much to Tulving (2002; Wheeler et al., 1997), to Suddendorf and Corballis (1997) and Suddendorf and Whiten (2001), to Humphrey (2003), and to Donald (1991). This projectable self, full blown, is the capability that makes us unique, but it is a very new adaptation. We seek to find evidence of where, in the evolutionary tree, this capability originated. The word "consciousness" is very slippery, insofar as there are undoubtedly many kinds and levels of consciousness. While the projectable self indisputably involves consciousness at the very highest level, there are many other lower forms of consciousness sometimes referred to as awareness, phenomenology, noetic and anoetic consciousness, sentience (see especially chapters 1, 4, and 5, this volume). Beings that do not have a fully developed sense of self-reflective projectable consciousness may nevertheless have these other kinds of consciousness, and furthermore, these are sufficient to underlie much clever behavior and learning (see chapter 12, this volume, for an argument that they may, indeed, underlie metacognition itself). The central tenet of this chapter is that while there is considerable evidence for these other kinds of consciousness in nonhuman primates, there are also some indications of at least some characteristics of the projectable self. For

example, there is now some suggestion that monkeys may have the beginnings of metacognition (see chapters 10, 11, and 12, this volume), though the flexibility of this capability is far from impressive. Chimps, bonobos, gorillas, and orangutans show greater metacognitive capabilities (see chapter 13, this volume). Glimmerings of episodic memory, it is asserted by some, may be within the range of chimps and gorillas (chapters 8 and 9, this volume;cf. chapter 1). At the human end of the scale, there is also an indication of the newness of this capability. Evidence from studies of autistic humans suggests that the fully articulated projectable self may have not yet fully saturated the human gene pool.

SELF-RECOGNITION

To have self-reflective consciousness, one must have a self. Thus, tasks that tap into a person's feeling of self, or self-recognition, provide an obvious empirical starting point. The most famous task of this ilk is the dye or rouge test, devised by Gallup (1970). Unbeknownst to the person or animal, a spot of dye or rouge is put on his or her face, in a location that cannot be seen without a mirror. Once the person or animal looks in a mirror, the question is, does he or she self-refer concerning the rouge? Gallup has argued that such self-reference indicates that the creature has a sense of self (though a number of other researchers, see Heyes 1994, 1998; Povinelli, 1993; Tomasello & Call, 1997, have suggested that self-awareness, in the Cartesian sense, might not be implied because the animal might simply be using the mirror to gain visual access to a part of the body previously perceived only tactually or proprioceptively, or it might be recognizing only its own behavior rather than its psychological or Cartesian self).

Humans

By the age of around 18 to 24 months (Amsterdam, 1972; Lewis & Brooks-Gunn, 1979), human babies pass the rouge test. Asendorpf, Warkentin, and Baudonniere (1996) note that at around the same time, and in synchrony, they show evidence of recognition of mind in others—showing empathy, communication through synchronic imitation, and cooperation. If this correspondence, and the implied cross-task dependence, turns out to be correct, then mirror self-recognition may indeed be a marker of a projectable self. Although young children without experience with mirrors showed some deficits in relating reflection to location in space, they showed no impairment in self-recognition (Priel & De Schonen, 1986).

Although most autistic children recognize themselves, some do not. Spiker and Ricks (1984) reported that only 36 of 52 autistics in their sample, between the ages of 4 and 12 years, showed evidence of self-

recognition. Failure to self-recognize was predicted by language impairment and level of functioning. Another study (Neuman & Hill, 1978) found that 1 out of 7 autistics between the ages of 5.5 and 11 years failed to recognize themselves. Last, Ferrari and Matthews (1983) found that those autistics who did not clearly recognize themselves (46.7% of their sample) had the lowest mental age in the sample (22 months) and were rated by their teachers as significantly lower-functioning on behavioral observations of affect, attentional skills, language, and interpersonal skills.

Nonhuman Primates

There is considerable evidence that chimps, so long as they are given experience with mirrors, and so long as they are socially raised, can pass the rouge test (Gallup, 1970; Gallup, McClure, Hill, & Bundy, 1971; Tomasello & Call, 1997). In Gallup's (1970) seminal studies with chimps and macaques, the chimps initially displayed aggressive behavior toward their own mirror image—as if the reflection were another, and threatening, chimp. Within a few days, however, they started using the mirror in a self-directed way—to pick food out of their teeth, and so on. When the chimps were anesthetized and marked with dye on a spot that they could not see directly without the use of the mirror, they would touch the spot on their own face—much as young children do when they are surreptitiously marked with a trace of rouge on their face (figure 2.3). Chimps, then, pass this

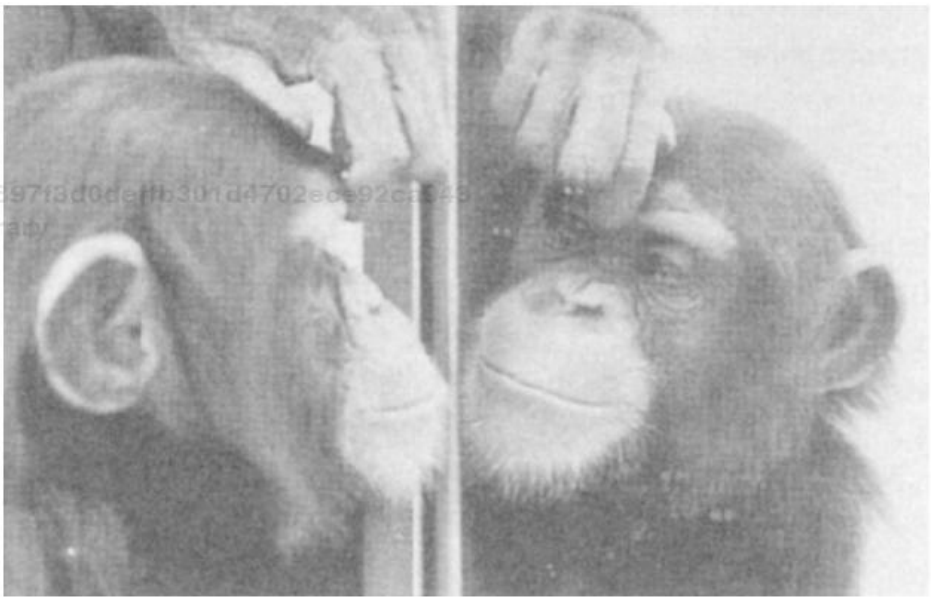

Figure 2.3 Megan looking in the mirror. Photo courtesy of the Cognitive Evolution Group, University of Louisiana at Lafayette.

test of self (see Tomasello & Call, 1997, for a review), as do orangutans (Suarez & Gallup, 1981), and some gorillas. Koko, for example, the famous human-reared gorilla (Patterson & Cohn, 1994), and King, the circus gorilla studied by Schwartz (chapter 9, this volume), both passed.

Despite extensive testing and extensive use of mirrors, there is still scant evidence for self-recognition in monkeys. An exception was reported for six cotton-topped tamarins (Hauser, Kralik, Botto, Garrett, & Oser, 1995), but the results did not replicate (Hauser, Miller, Liu, & Gupta, 2001). The fact that monkeys will look away rather than staring back at another monkey may qualify the results of mirror testing. Novak (1996, cited in Tomasello & Call, 1997) found that when monkeys were first trained to look at one another, they showed some evidence of mirror self-recognition. Overall, though, clear evidence for monkey self-recognition is lacking.

THEORY OF MIND

Gallup (1982) has argued that awareness of self implies awareness of other minds. The self, by this perspective, is an essentially social being who gains his or her definition from the social matrix (see also chapters 3, 4, and 6, this volume). However, one might imagine that a mind could exist in a solipsistic world, in which there is self-reflection, but no other-reflection. This may be an empirical question though, and, as mentioned above, self-recognition and empathy seem to develop at about the same time. Furthermore, Frith and Happe (1999) have shown an association between impairment on theory of mind tasks and an inability to introspect, reflect on one's own actions, or anticipate one's own actions. Even so, self-recognition typically develops earlier (it is usually in place by age 18 months) than does theory of mind (which is only entrenched by about 4 years). Insofar as an individual might have a mind, but not extend the attribution to other beings, a failure to show evidence of theory of mind does not necessarily rule out the possibility of self-reflective consciousness of some sort. But if a nonhuman primate showed evidence of having theory of mind, this in itself would be proof positive of self-reflective consciousness of a highly projectable form.

Humans

For most adult humans, theory of mind is a ubiquitous attribution to other people. It is often so overextended that we also attribute mind—intentions, goals, thoughts, memories, plans, emotions—to animals, indiscriminately, and even to inanimate objects. People will anthropomorphize even to moving abstract shapes, ascribing volitional mental states to them. For example, Adolphs (1999) reviewed a replication of

a classic study by Heider and Simmel (1944) in which people viewed a short film of an open square and some triangles and circles moving in various directions. A typical description was, "And then the big triangle chased the little triangle around. Finally, he went in, got inside the box to go after the circle, and the circle was scared of him . . . and they went off on their way, and the big triangle got upset and started breaking the box open" (p. 473). Contrast Klin's (2000, p. 840) transcription of an autistic's description of the same film: "The big triangle went into the rectangle. There were a small triangle and a circle. The big triangle went out. The shapes bounce off each other. The small circle went inside the rectangle. The big triangle was in the box with the circle." It is the rare person who can stick to the kind of strict operational description of the patterns of movement given by the autistic, stripped of intentional terminology. When they do so, they sound very odd indeed.

Wellman, Cross, and Watson (2001) concluded from a meta-analysis of experiments on theory of mind capabilities in human infants that the development of theory of mind begins at around 2 years of age, and continues until about 4 years of age (see also Nelson, 2000; Perner, 1991). Of course, insofar as social intelligence is highly related to theory of mind, the development of this capability may continue well into adolescence and beyond. There are some cases of patients, however, who apparently lack the ability.

Stuss, Gallup, and Alexander (2001) described a patient with lesions to the right frontal cortex (or with bifrontal lesions) who failed an inference-based theory of mind task. The patient was to indicate which of two graduate students (both of whom had previously demonstrated their trustworthiness) he wished to show where a prize, which was concealed from his own view by a barrier, was hidden. One student sat by the subject in front of the barrier as the experimenter hid the prize, while the other student sat next to the experimenter, behind the barrier, and could clearly see where the prize was hidden. When normal people, or patients with nonfrontal or left frontal lesions are asked, in this situation, they consistently choose the student who sat next to the experimenter, the one who could see where the prize was hidden. Presumably they choose this student because they imagine what they would see if they were him, and realize, effortlessly, that only the person behind the barrier could actually see where the prize was hidden. Without theory of mind, or the realization that another's experience is like their own, however, they might not make this inference. The right frontal patient chose at random between the two students.

In contrast to Stuss's finding, which implicates the right prefrontal cortex in theory of mind, Happe and colleagues (1996), in a positron emission tomography study using normals, found that a highly cir-

cumscribed region of the left medial prefrontal cortex was activated during story comprehension that involved taking another's perspective. This area was not activated in Asperger's autistics, however, who (as we shall review), have great difficulty with theory of mind tasks. Rowe, Bullock, Polkey, and Morris (2001) found theory of mind deficits in patient groups with either left or right unilateral frontal lesions, and Stone, Baron-Cohen, and Knight (1998) found deficits, similar to those of Asperger patients, in patients with bilateral orbitofrontal lesions.

It has been proposed (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, 1995) that the fundamental impairment in autism is a lack of theory of mind. Although numerous other idiosyncratic behaviors and perceptions are associated with autism (for example, stereotypy of movement, compulsions, extremely repetitious behavior, lack of appropriate voice modulation, heightened perceptions, pain tolerance, language deficits, etc.), the inability to understand another person's perspective is a theme that underlies much research (see, for example, Adolphs, Sears, & Piven, 2001; DeLong, 1992; Hobson, 1990; Hughes & Russell, 1993; Leslie, 1987; Ozonoff, Pennington, & Rogers, 1991, for other theories of the dysfunction). The lack of theory of mind would result in a social blindness that allows the highly developed Asperger's autistic to follow rules of behavior well, but without the sensitivity to nuance that we take for granted as part of a person's social skills. Such people often behave inappropriately without realizing it.

Two tasks, both with a number of variants, have come to be classic tests of theory of mind research. The first (Wimmer & Perner, 1983), often called the Sally-Anne task, is one in which a person (Sally), or sometimes a puppet, sees some kind of reward being hidden and then leaves the room. A second person (Anne) then changes the location of the reward. When Sally returns, the child being tested is asked where she will look. Normals typically say that she will search in the original location in which she left the reward, but autistic children (and normal children younger than age 4) typically say that she will look where the reward is actually currently located (Baron-Cohen et al., 1985). The second task is one in which a package that typically contains candy (i.e., M&M's) is opened and revealed to contain pencils. The child is asked what a new person will think is in the package. Normals say M&Ms, but autistic children frequently say pencils—revealing an inability to put themselves in the epistemic position of the other person (Perner, Frith, Leslie, & Leekam, 1989).

It would appear that, insofar as autism is a genetically predisposed condition, and insofar as the theory of mind hypothesis is correct, this very high level of self-reflective consciousness may not have yet fully saturated the human gene pool, or that it is, at the least, fragile.

Nonhuman Primates

The first exploration of theory of mind in nonhuman primates was done by Premack and Woodruff (1978), who gave their chimpanzee, Sarah, a series of problems to solve that they claimed involved inferences about the goals, wants, needs, desires, and knowledge of other people, that is, which implied theory of mind. For example, a videotape may have shown a human trying to get some inaccessible food, or trying to extricate himself from a locked cage, or shivering from a nonfunctioning, unplugged heater, or trying to play an unplugged record player. Sarah was then given a forced-choice test in which one of the alternatives was a solution to the human's problem. That she picked the correct choices most of the time suggested to the authors that she empathized with the human's problem, exhibiting theory of mind. Humans, in this situation, would solve the problem, essentially, by asking themselves what they would do if they were in the other person's shoes, and the claim was that Sarah did the same thing.

This interpretation, while intriguing, might be disputed. It is possible that the chimps simply gave the solution as a learned response, or remembered the correct sequence from having seen it before, rather than projecting themselves into the position of the other. The authors argued against mere familiarity as a possibility, since the chimps would have had the opportunity to view unplugged as well as plugged-in plugs, or burned wicks as well as new ones. They also argued for intentionality insofar as the responses varied depending upon whether Sarah liked the humans involved or not. Keith was Sarah's favorite keeper. When he was the actor in the problems, she solved them correctly eight out of eight times. However, when Bill, an acquaintance whom Sarah disliked, was the actor, she was right only two out of eight times.

The nature of the problems posed clouds a straightforward interpretation. These were not classic theory of mind problems, such as the Sally-Anne task or the M&M's task, but rather sequences that culminated in a solution. Notably, in the research on autism, three types of sequence problems are typically given: (1) mechanical, which are understood by all autistic participants; (2) behavioral, which can be done without reference to mental states; and (3) mentalistic, which require knowledge of the depicted person's state of belief, and which typically autistics cannot do. Only the third type provides information about theory of mind, though all three types seem to have been included in the Premack and Woodruff set. Thus, although Premack and Woodruff's contention is intriguing, their conclusions are not beyond dispute.

A second approach was taken by Povinelli and his group. A number of researchers have shown that chimps, like children, follow a

person's gaze. When people do this, presumably it is because they are looking to see what the other person sees, and it implies theory of mind. However, based on a series of clever experiments, Povinelli (2000; Povinelli & Eddy, 1996a, 1996b, 1996c) argued that this interpretation does not necessarily follow with chimps, insofar as they do not distinguish between people who can see or cannot (because they either have a mask on or a bucket over their heads, say). They seem to respond, instead, to body orientation, and conform to what Povinelli thinks is a lower order rule rather than to a notion about what the other person or animal is perceiving. He concludes that gaze following does not necessarily imply theory of mind.

Interestingly enough, Premack and Premack (1983), years earlier, noted that Jessie was the only one of the four non-language-trained juvenile chimps who, in trying to get a trainer to move across the room to help her with a task, without hesitation and on the first trial, removed a blindfold from the eyes of the trainer—apparently realizing that a blindfold on the eyes impaired vision. Jessie did not remove the blindfold when it was over the mouth. It may be the rare chimp who actually understands the relation between eyes and seeing, and Povinelli may not have had such a rare chimp in his cohort. It is possible, though, that despite this lack of realization, chimps might still understand that another creature has a mental life similar to their own. Even so, gaze-following by chimpanzees can no longer be taken as unequivocal evidence that chimps have theory of mind.

A third approach was taken by Call and Tomasello (1999), who devised a nonverbal false belief (Sally-Anne) task. One adult human (the hider) hid a reward in one of two identical containers. Another adult (the communicator) attempted to help the subject by placing a sticker on the container that she believed to hold the reward. Both 4- and 5-year-old children and the apes (chimps and orangutans) were able to use the stickers to locate the rewards in the control trials. In the critical trials, though, the communicator left the room, and the hider switched the location of the reward. When the communicator came back, she marked the location at which she had last seen the reward. The hider then gave the child or the ape the chance to look for the reward. The question was whether the subject went for the container that had the sticker. The children tended to discount the sticker—revealing that they knew that the communicator did not know; but the apes went for it—revealing what appears to be a lack of appreciation of the communicator's knowledge, and a lack of theory of mind.

Although these results mitigate against chimps having theory of mind, more recently Tomasello, Call, and Hare (2003) have revised their negative conclusions, and demonstrated what may be a genuine case of chimpanzees having theory of mind. The mind that they have

a theory of, though, has to be one they care about. Tomasello et al. used pairs of dominant and subordinate chimps, in competition for food. Some of the time, the subordinate chimp could see the food and could also see that the dominant could not see it, by virtue of a barrier. The subordinates took advantage of their knowledge of what the dominant did not know in a variety of flexible ways that suggested that they knew what he knew. In a second set of experiments, the subordinate watched a human hide the food and also observed whether or not the dominant also saw the hiding. The researchers found that when the subordinate had observed that the dominant had not observed the hiding, he behaved quite differently than when he knew that the dominant had seen the hiding. Povinelli and Vonk (2003) have criticized these experiments on the grounds that the chimps might have had past experience or innate wiring in such situations, and that, to test the notion of knowledge of mental states, one needs to use situations that could not possibly have innate or experiential precedents (which might allow "behavioral abstraction") for the subject chimps:

> Any experiments that rely upon a behavioral abstraction will be of little use, especially when this invariant is one the subject has previously witnessed, or that they are likely to have evolved to detect or exploit. Indeed, contrary to recent speculations, behavioral interactions that make the most ecological sense to the organism are precisely the ones that will be least diagnostic about whether the organism is reasoning about mental states and behavior or behavior alone. (p. 159)

But if, following Povinelli and Vonk, we grant that tests for theory of mind are valid only if the minds and situations have no importance or meaning (that could be generalized from any past experience or innate predispositions) or social importance to the person or animal being studied, it seems doubtful that anyone, including normal adult humans, would ever show positive results. Povinelli and Vonk (2003) concluded that "the idea that theory of mind is the 'holy grail' of comparative cognition needs to be abandoned" (p. 160). We are more inclined to agree with Tomasello et al.'s (2003) conclusion: "The stakes here are large. At issue is no less than the nature of human cognitive uniqueness" (p. 156).

In summary, then, there is no completely undisputed evidence that any nonhuman primate has theory of mind, though there is a suggestion that in some circumstances that are deeply social, chimps may have it. There is considerable evidence that some humans—those with autism—do not have theory of mind, or at least have deficits in this area. As mentioned above, however, this is an extreme form of projection. Not only does the individual have to have self-reflective consciousness, but he or she also has to be able to attribute it to others

and to correctly make inferences from that attribution. Some kinds of self-reflective consciousness might well exist prior to this highly articulated form.

DECEPTION

Humans

The ability to deceive is an interesting one, from the perspective of the projectable mind, because the tactical deceiver—to be effective—must be able, at least to some extent, to second-guess the target of his or her deception, understanding the thoughts, feelings, and inferences that will be made. It implies a glimmering of theory of mind—though perhaps short of that required for the tasks detailed in the previous section. Without an ability to, at least partially, project oneself into the other person's mind, deception would be ineffective. Furthermore, the deception has to be subtle enough to be undetected, since a detected deception can be disastrous for the perpetrator (see Cosmides & Tooby, 1992, for evidence and arguments concerning the evolution of so-called cheater-detector mechanisms in humans). People are extremely good at deception, as well as at the detection of potential deception. Novels and histories revel in it.

However, as with the other indications of social understanding, autistics appear to have particular problems both in deceiving and in detecting deception. Frith (1989) noted that the profiles given of many of the hermit saints suggest autism. These saints are often known for their simplicity and lack of deceit. Their truthfulness has historically been taken as virtue, but this characterization seems altogether undeserved since, if Frith is correct, it results not from self-restraint and goodness under pressure from the dark side, but rather from of a simple lack of understanding of the other's point of view.

Several studies have investigated deception in autistic populations, and all have shown them to be guileless. For example, Russell, Maunthner, Sharpe, and Tidswell (1991) compared autistic subjects, children with Down's syndrome, and normal 3- and 4-year-olds. The children learned that it was in their interest to tell the experimenter to look into an empty box for a chocolate, rather than into the box that actually contained the chocolate. Both the 4-year-olds and the children with Down's syndrome used the deceptive strategy, but the autistic children and the 3-year-olds consistently went with the box that actually contained the chocolate, failing to inhibit the knowledge of their own epistemic state. Similarly, Yirmiya, Solomonica-Levi, and Shulman (1996), using a procedure in which a doll creates a false trail of footprints, showed that while the autistic children could use the deceptive method as well as mentally retarded children (but not

nearly as well as normals), they did not realize that by so doing they manipulated the belief of the other person involved. Other researchers have shown deficits in deceptive tasks such as penny hiding (Baron-Cohen, 1992; Oswald & Ollendick, 1989) in autistic participants. Thus, deception, like other tasks requiring a projectable self, appears to show as a deficit in autism.

Nonhuman Primates

The literature on tactical deception in nonhuman primates is largely anecdotal. For example, de Waal (1992) cites a case in which a young female chimp was aggressively chased by an older female, but managed to escape. Ten minutes later, the older female made reconciliatory gestures, approaching the younger with an open hand and making soft panting noises (the usual chimp prelude to a kiss). However, when the younger chimp came close, the older one lunged and bit her fiercely, before she was able to free herself. Presumably, the sweet gestures were just a deceptive attempt to lure the foe in close enough for the bite. Many such tales of chimp subterfuge have been recorded by primatologists and tabulated, systematized, and categorized in an enormous study by Byrne and Whiten (1992; Whiten & Byrne, 1988) who requested all reports of observed tactical deception among primate researchers. The kinds of observations reported were like that of Coussi-Korbel (1994), in which a subordinate young male would move in an indirect route toward a food goal to mislead a dominant male (and thus get the food for himself), or in which monkeys would point to the wrong location for hidden food (Mitchell & Anderson, 1997), or in which females would use their charms to distract a male in order to get food. "One of the female baboons at Gilgil grew particularly fond of meat, although males do most hunting. A male, one who does not willingly share, caught an antelope. The female edged up to him and groomed him until he lolled back under her attentions. She then snatched the antelope carcass and ran" (Jolly, 1985, p. 412). A standardized request, which included a computation of hours of observation, species, and so on, resulted in 253 such records, which were then analyzed, categorized by deception type, and classified by species.

The most common reports of deception came from our nearest relatives, genus *Pan*—chimpanzees and bonobos, followed closely by baboons. Some primates—lemurs in particular—appear not to deceive at all, despite the fact that it would presumably be much to their advantage. Deception, among monkeys, was rare. Whiten and Byrne (1988) noted some caveats to their results. First, the species reported with such a high rate of deception is also one studied by researchers who have been vocally antibehaviorist. But even leaving aside the possibility that some researchers might see mind to a greater extent

than others, and while acknowledging that a more experimental approach would be desirable, it is difficult not to be convinced (and amused) by the many anecdotes provided in this massive work. While formal tests of theory of mind in nonhuman primates are equivocal, insofar as that capacity is reflected in deceptive behavior, we humans do not appear to be the only ones who have it.

METACOGNITION

Humans

A large and rapidly developing literature is investigating the metacognitive capabilities of adults (see Metcalfe, 1996, 2000, for fairly recent reviews of data and theory). Judgments about what one knows are used to guide problem-solving behavior in humans (Metcalfe & Wiebe, 1987; Simon, 1979; Simon & Reed, 1976) and to indicate how close to the solution to a problem one is, or how near one is to remembering a forgotten memory (Metcalfe, 1986a). Such judgments serve as controls of problem-solving search processes and also of memory retrieval efforts (Miner & Reder, 1994). Adults are capable of highly refined judgments of confidence, both prospectively and retrospectively (Morris, 1990). People are even able to predict, with high accuracy, what they will be able to remember later, even though they cannot remember the answer at the time they make such "feeling-of-knowing" judgments (Blake, 1973; Butterfield, Nelson, & Peck, 1988; Costermans, Lories, & Ansay, 1992; Cultice, Somerville, & Wellman, 1983; Gruneberg & Monks, 1974; Hart, 1965, 1967; Hertzog & Dixon, 1994; Lachman, Lachman, & Thronesbury, 1979; Leonesio & Nelson, 1990; Metcalfe, 1986a, 1986b; Nelson, Leonesio, Shimamura, Landwehr, & Narens, 1982; Schacter, 1983). Adults are not perfect, however, in their metacognitions, and some of their inaccuracies about their own cognitive processes and capabilities have garnered a great deal of attention (e.g., Bjork, 1994; Nelson & Dunlosky, 1991). These biases and inaccuracies notwithstanding, however, it is safe to say that normal adult humans are able to make remarkably accurate judgments about what they currently know and what they will know, and they put those judgments to use, either implicitly or explicitly (Koriat & Goldsmith, 1996), in their behavior.

There is consensus in the field that the judgments are made by monitoring mental contents, in a manner that is consistent with Humphrey's (2003) inner eye. For example, in making judgments of learning, people are thought to attempt to retrieve whatever they can, given the retrieval cues available to them, then to mentally look at the content of what they retrieved, as well as at the characteristics of the process in which they just engaged during the retrieval (was it easy,

fluent, and fast, or labored?). Then, based on the results of this inner looking, they give a judgment, numerical or otherwise, about how well they have learned the targeted item. In making feeling-of-knowing judgments, they look at the quantity and familiarity of all the information they have, including the cue and partial information about the target (Koriat, 1993; Metcalfe, 1993; Metcalfe, Schwartz, & Joaquim, 1993) and they assess this to make their judgment.

The monitoring and control involved in metacognition appears to be associated with the last-developed area of the brain, namely the frontal cortex. As with the frontal patient reported by Stuss et al. (2001) who showed deficits in theory of mind, deficits in metacognition are also associated with frontal damage (Janowsky, Shimamura, & Squire, 1989). Insofar as the frontal lobes mature late, it is not surprising that metacognitions also appear developmentally late.

There has been very little research on metacognition in autistics, so no firm conclusions can yet be reached. One study (Farrant, Boucher, & Blades, 1999) investigated whether children knew about strategies for doing things like enhancing their memory span, or whether or not retrieval cues or verbalization enhanced performance, and found no differences among autistic and other children. Another study (Farrant, Blades, & Boucher, 1999) showed some impairment in the correspondence between predictions and later performance between autistic and other children, but they may have stemmed from an underlying memory problem rather than from an inability of the autistics to predict what they should be able to remember. Clinical and self-reports of autistics, though, suggest that they may be able to make rather refined metacognitive judgments in some areas. They can, for example, understand others' behavior by constructing a theory based on their experience of contingencies (Frith, 1989; Sacks, 1995). Insofar as the same mechanisms might be used in some metacognitive judgments, and it is not necessary to mentally project to do the tasks, one might not expect deficits in autistics. Certain metacognitive tasks may thus be the simplest of tasks involving an inner eye, and may be possible where more complicated tasks that require not only an inner eye, but that the individual be able to project it in space and time, may be impossible.

Nonhuman Primates

There are now a number of indications that nonhuman primates, including even monkeys, may be capable of making some metacognitive judgments. Researchers (Shields, Smith, & Washburn, 1997; Smith, Shields, Allendoerfer, & Washburn, 1998; Smith, Shields, & Washburn, 2003; chapter 10, this volume) have shown that monkeys can make uncertainty judgments, although it is not generally agreed upon whether uncertainty judgments, in and of themselves, are metacogni-

tive (Metcalfe, 2003). Hampton (2001; chapter 11, this volume) has given an impressive demonstration of metacognition, with judgments about the contents of working memory, in monkeys. Finally, Son and Kornell's data (chapter 12, this volume) also indicate that even rhesus monkeys are able to perform metacognitive tasks. Insofar as metacognitions, however simple, may be the germ of full-blown self-reflection and may be the first indication of an inner eye, the discovery of these capabilities in species other than ourselves may be the strongest primordial indication of self-reflective capability in other primates.

Call (chapter 13, this volume) documents a particularly advanced metacognitive/control capability in chimps and other primates. Not only do they appear to know what they do not know (see also Son & Metcalfe, 2004, for a similar capability in humans), but they also seek to remedy their lack of knowledge. Call shows that several species of primates (unlike lower animals such as dogs, say), in the face of uncertainty, will actively seek information in a effort to remedy their ignorance.

EPISODIC MEMORY

Tulving and his colleagues (Tulving, 2002; Wheeler, Stuss, & Tulving, 1997; chapter 1, this volume) have made a compelling case concerning the close relation between episodic memory and autonoetic consciousness. We claim that both are closely related to what we, in this chapter, refer to as the projectable self. It follows that the pattern seen for other markers of the projectable self, outlined above, should also be manifested with episodic memory tasks. Nelson (2000; chapter 4, this volume) provides an authoritative description of the development of episodic memory in early childhood and illustrates how it relates to the development of self-reflection, self-concept, consciousness, and infantile amnesia. The timing of the onset of episodic memory seems to correspond well with that of other projectable-self capabilities. Whether or not nonhuman primates show any signs of episodic memory is a topic of focal concern in other chapters (chapters 8 and 9, this volume; and see Schwartz & Evans, 2001). Since these chapters deal with this issue specifically, and there is much discussion throughout the book, we defer to them.

The one domain in the literature that bears on the relation between the projectable self and episodic memory that is not covered in other chapters in this volume—the missing piece to the puzzle—is the memory capability of autistics. As reviewed above, autistics have demonstrated deficits in self-recognition, in theory of mind, and in deception. If autistics lack a projectable self, then this should also be manifested by impairments in episodic memory.

The few studies that we have been able to find on this topic suggest, though with some equivocation, that autistics do indeed have difficulties with just this kind of memory, though not with all kinds of memory. For example, Boucher (1981a) reported that recall of recent events in autistic children was inferior to that of both normal children and ability-matched retarded control children. In a different study, she (Boucher, 1981b) reported similar overall free recall of word lists, but the autistic children relied more heavily on recalling the most recently presented items and reported fewer of the earlier words—perhaps suggesting less "time travel" to the earlier parts of the list. Tager-Flusberg (1991) found that autistic children were not different from mentally retarded and normal children in free recall of an unrelated list of words, or in their use of semantic or rhyme cues to retrieve unrecalled words from memory. However, they were impaired in their free recall of the semantically related list.

Boucher and Warrington (1976) compared autistic children to controls in a variety of tasks. Tasks that we would classify as implicating episodic memory, namely, recall tasks (in this case of pictures, written words, and spoken words), revealed memory deficits. In tasks that involved less of an autonoetic component because they were cued, the autistic children were not impaired. They were not impaired on cued recall tasks or a test of unrelated paired associates. Oddly, though, the autistic children were impaired on a forced-choice recognition task, a task that might presumably be done on the basis of familiarity, and that need not entail autonoetic involvement. Overall—though the data are not entirely consistent—it would seem that the memory tasks on which the autistic children were impaired were those that may require autonoetic consciousness. The authors of these studies suggested a parallel between autism and amnesia.

Similarly, Bennetto, Pennington, and Rogers (1996) found that autistics were impaired on source memory, temporal order memory, free recall, and working memory (though the last need not involve a projective self, and its impairment is also not a consistent finding in the autism literature). They were unimpaired on short- and long-term recognition, cued recall, and new learning ability. Boucher and Lewis (1989) have pointed out that autistic children have difficulties answering questions about their own past activities.

Finally, Klein, Chan, and Loftus (1999) conducted an interesting case study investigating semantic and episodic self-knowledge in a high-functioning autistic individual. At the time of the study, R.J. was a 21-year-old whose autistic symptoms dated back to about 8 months of age. His immediate memory span was normal, as was his verbal fluency in generating category exemplars (measuring a kind of semantic memory). Like other autistics, R.J.'s free recall of unrelated nouns was impaired. Interestingly, R.J. was found to have rather accurate

knowledge of his own personality traits, as measured by both his own test-retest reliability and also by the concordance of his self-ratings with his mother's ratings of him. However, his accurate semantic assessment of himself was in stark contrast to his inability to retrieve autobiographical events from his own past, when given those traits as cues. Normal subjects were able to do this retrieval 10 out of 10 times, meeting all criteria of episodic scoring. R.J. attained a score of only 2 out of 10, and then only when the criterion for acceptance was extremely lenient. An example: "*Tester:* Can you remember a time when you acted friendly toward someone? *R.J.:* Mmm . . . when . . . when people were nice to me. *Tester:* Was this a particular person you remember being nice to you? *R.J.:* Anyone" (p. 422). As the authors put it, "Apparently, R.J. did not need to remember how he had behaved in the past to know what he was like" (p. 425).

In summary, then, autistic children appear to be selectively impaired on memory tasks that are episodic in nature and that would appear to involve autonoetic consciousness. These findings suggest that they may be unable to project themselves into their own pasts, as is required by such tasks. Both theory of mind tasks (requiring projection to another's point of view), and episodic memory tasks (requiring projection into one's own personal past) appear to be impaired in autism.

CONCLUSION

The conclusion that seems to be emerging, but which is debated and disputed more fully in the other chapters in this volume, is that there are some indications of self-reflective consciousness in some primates other than humans. There are indications of self-recognition in the great apes. There are suggestions of some fragmentary episodic (-like) memory in some of the great apes. Other primates, though hardly candidates for CEO of a major corporation, have the beginnings of an ability to deceive. The great apes, and even monkeys, appear to have some metacognitive capabilities. So far, the evidence that any primates other than humans have full-blown theory of mind is still under dispute, but then very little research has been conducted on this intriguing topic, and it may turn out that with further investigation, a consensus will emerge in favor of humans not being alone in this regard. Adult humans project their consciousness of themselves into their own past and future, and into the minds of others, with remarkable ease. Investigation of the emergence of these capabilities, in a primordial form in other primates, as well as the acknowledgment of their fragility, as shown in autistic and frontal lobe patients, provides insight into this most quintessentially human kind of knowledge—our consciousness of ourselves.

REFERENCES

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3,* 469–479.

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience, 13,* 232–240.

Amsterdam, B. K. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology, 5,* 297–305.

Asendorpf, J. B., Warkentin, V., & Baudonniere, P. (1996). Self-awareness and other-awareness II: Mirror self-recognition, social contingency awareness, and synchronic imitation. *Developmental Psychology, 32,* 313–321.

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. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA: MIT Press.

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

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development, 67,* 1816–1835.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–206). Cambridge, MA: MIT Press.

Blake, M. (1973). Prediction of recognition when recall fails: Exploring the feeling-of-knowing phenomenon. *Journal of Verbal Learning and Verbal Behavior, 12,* 311–319.

Boucher, J. (1981a). Memory for recent events in autistic children. *Journal of Autism and Developmental Disorders, 11,* 293–301.

Boucher, J. (1981b). Immediate free recall in early childhood autism: Another point of behavioural similarity with the amnesic syndrome. *British Journal of Psychology, 72,* 211–215.

Boucher, J., & Lewis, V. (1989). Memory impairments and communication in relatively able autistic children. *Journal of Child Psychology and Psychiatry, 30,* 99–122.

Boucher, J., & Warrington, E. K. (1976). Memory deficits in early infantile autism: Some similarities to the amnesic syndrome. *British Journal of Psychology, 67,* 73–87.

Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology, 24,* 654–663.

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: Evidence from tactical deception. *Man, 27,* 609–627.

Call, J., & Tomasello, T. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70,* 381–395.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptation for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 163–228). New York: Oxford University Press.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 142–150.

Coussi-Korbel, S. (1994). Learning to outwit a competitor in mangabeys. *Journal of Comparative Psychology, 108,* 164–171.

Cultice, J. C., Somerville, S. C., & Wellman, H. M. (1983). Preschoolers memory monitoring: Feeling of knowing judgments. *Child Development, 54,* 1480–1486.

DeLong, G. R. (1992). Autism, amnesia, hippocampus, and learning. *Neuroscience and Biobehavioral Reviews, 16,* 63–70.

Dennett, D. C. (1991). *Consciousness explained.* Toronto, Canada: Little, Brown.

Descartes, R. (1999). *Discourse on method.* London: Penguin Books. (Original work published 1637)

de Waal, F. B. M. (1992). Intentional deception in primates. *Evolutionary Anthropology, 1,* 86–92.

Donald, M. (1991). *Origins of the modern mind.* Cambridge, MA: Harvard University Press.

Farrant, A., Blades, M., & Boucher, J. (1999). Recall readiness in children with autism. *Journal of Autism and Developmental Disorders, 29,* 359–366.

Farrant, A., Boucher, J., & Blades, M. (1999). Metamemory in children with autism. *Child Development, 70,* 107–131.

Ferrari, M., & Matthews, W. S. (1983). Self-recognition deficits in autism: Syndrome-specific or general developmental delay? *Journal of Autism and Developmental Disorders, 13,* 317–324.

Frith, U. (1989). *Autism: Explaining the enigma.* Cambridge, MA: Basil Blackwell.

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

Gallup, G. G. (1970). Chimpanzees: Self-recognition. *Science, 167,* 86–87.

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

Gallup, G. G., McClure, M. K., Hill, S. D., & Bundy, R. A. (1971). Capacity for self-recognition in differentially reared chimpanzees. *Psychological Record, 21,* 69–74.

Gruneberg, M. M., & Monks, J. (1974). "Feeling of knowing" and cued recall. *Acta Psychologica, 38,* 257–265.

Hampton, R. R. (2001). Rhesus monkeys know when they remember. *Proceedings of the National Academy of Sciences U.S.A., 98,* 5359–5362.

Happe, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., Dolan, R., Frackowiak, R., & Frith, C. (1996). "Theory of mind" in the brain: Evidence from a PET scan study of Asperger syndrome. *NeuroReport, 8,* 197–201.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56,* 208–216.

Hart, J. T. (1967). Memory and the memory-monitoring process. *Journal of Verbal Learning and Verbal Behavior, 6,* 685–691.

Hauser, M. D., Kralik, J., Botto, C., Garrett, M., & Oser, J. (1995). Self recognition in primates: Phylogeny and the salience of species-typical traits. *Proceedings of the National Academy of Sciences U.S.A., 92,* 10811–10814.

Hauser, M. D., Miller, T. C., Liu, K., & Gupta, R. (2001). Cotton-top tamarins (*Saguinus oedipus*) fail to show mirror-guided self-exploration. *American Journal of Primatology, 53,* 131–137.

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

Hertzog, C., & Dixon, R. A. (1994). Metacognitive development in adulthood and old age. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 227–252). Cambridge, MA: MIT Press.

Heyes, C. M. (1994). Reflection on self-recognition in primates. *Animal Behavior, 47,* 909–919.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences, 21,* 101–134.

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

Hughes, C., & Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: Its implications for theories of autism. *Developmental Psychology, 29,* 498–510.

Humphrey, N. K. (2003). The uses of consciousness. In N. Humphrey (Ed.), *The mind made flesh: Essays from the frontiers of psychology and evolution* (pp. 65–85). New York: Oxford University Press.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Memory and metamemory: Comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology, 17,* 3–11.

Jolly, A. (1985). *The evolution of primate behavior* (2nd ed.). New York: Macmillan.

Klein, S. B., Chan, R. L., & Loftus, J. (1999). Independence of episodic and semantic self-knowledge: The case from autism. *Social Cognition, 17,* 413–436.

Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The social attribution task. *Journal of Child Psychology and Psychiatry, 7,* 831–846.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,* 490–517.

Lachman, J. L., Lachman, R., & Thronesbury, C. (1979). Metamemory through the adult life span. *Developmental Psychology, 15,* 543–551.

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 464–470.

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

Lewis, M., & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self.* New York: Plenum Press.

Metcalfe, J. (1986a). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 288–294.

Metcalfe, J. (1986b). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 623–634.

Metcalfe, J. (1993). Novelty monitoring, metacognition, and a control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Reviews, 100,* 3–22.

Metcalfe, J. (1996). Metacognitive processes. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of perception and cognition* (pp. 381–407). San Diego, CA: Academic Press.

Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 197–211). New York: Oxford University Press.

Metcalfe, J. (2003). Drawing the line on metacognition. *Behavioral and Brain Sciences, 26,* 350–351.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 851–864.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and non-insight problem solving. *Memory and Cognition, 15,* 238–246.

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: Its metacognitive role in regulating question answering. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 47–70). Cambridge, MA: MIT Press.

Mitchell, R. W., & Anderson, J. R. (1997). Pointing, withholding information, and deception in capucin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 111,* 351–361.

Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 223–232.

Nelson, K. (2000). Memory and belief in development. In D. Schacter & E. Scarry (Eds.), *Memory, brain, and belief* (pp. 259–289). Cambridge, MA: Harvard University Press.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL-effect." *Psychological Science, 2,* 267–270.

Nelson, T. O., Leonesio, R. J., Shimamura, A. P., Landwehr, R. S., & Narens, L. (1982). Overlearning and the feeling of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 279–288.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition?. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Neuman, C. J., & Hill, S. D. (1978). Self-recognition and stimulus preference in autistic children. *Developmental Psychobiology, 11,* 571–578.

Oswald, D. P., & Ollendick, T. H. (1989). Role taking and social competence in autism and mental retardation. *Journal of Autism and Developmental Disorders, 19,* 119–127.

Ozonoff, S., Pennington, B. F., & Rogers, S. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32,* 1081–1105.

Patterson, F. G. P., & Cohn, R. H. (1994). Self-recognition and self-awareness in lowland gorillas. In S. Taylor Parker, R. W. Mitchell, & M. L. Boccia (Eds.), *Self-awareness in animals and humans: Developmental perspectives* (pp. 273–290). Cambridge, MA: Cambridge University Press.

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

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

Povinelli, D. J. (1993). Reconstructing the evolution of mind. *American Psychologist, 48,* 493–509.

Povinelli, D. J. (2000). *Folk physics for apes.* New York: Oxford University Press.

Povinelli, D. J., & Eddy, T. J. (1996a). Chimpanzees: Joint visual attention. *Psychological Science, 7,* 129–135.

Povinelli, D. J., & Eddy, T. J. (1996b). Factors influencing young chimpanzees' recognition of "attention." *Journal of Comparative Psychology, 110,* 336–345.

Povinelli, D. J., & Eddy, T. J. (1996c). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development, 61,* record #247.

Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences, 7,* 157–160.

Premack, D., & Premack, A. J. (1983). *The mind of an ape.* New York: Norton.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences, 4,* 515–526.

Priel, B., & De Schonen, S. (1986). Self-recognition: A study of a population without mirrors. *Journal of Experimental Child Psychology, 41,* 237–250.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). "Theory of mind" impairments and their relationship to executive functioning following frontal lobe excisions. *Brain, 124,* 600–616.

Russell, B. (1972). *A history of western philosophy.* New York: Simon and Schuster. (Original work published 1945)

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The "window task" as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9,* 331–349.

Sacks, O. W. (1995). *An anthropologist on Mars: Seven paradoxical tales.* New York: Knopf.

Schacter, D. L. (1983). Feeling of knowing in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 39–54.

Schwartz, B. L., & Evans, S. (2001). Episodic memory in primates. *American Journal of Primatology, 55,* 71–85.

Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and rhesus monkeys (*Macaca mulatta*) in a psychophysical same-different task. *Journal of Experimental Psychology: General, 126,* 147–164.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology, 30,* 363–396.

Simon, H. A., & Reed, S. K. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology, 8,* 86–97.

Smith, J. D., Shields, W. E., Allendoerfer, K. R., & Washburn, D. A. (1998). Memory monitoring by animals and humans. *Journal of Experimental Psychology: General, 127,* 227–250.

Smith, D. J., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26,* 317–339.

Son, L. K., & Metcalfe, J. (2004). *Judgments of learning: Evidence for a two-stage process.* Manuscript submitted for publication.

Spiker, D., & Ricks, M. (1984). Visual self-recognition in autistic children: Developmental relationships. *Child Development, 55,* 214–225.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contribution to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640–656.

Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for theory of mind. *Brain, 124,* 276–286.

Suarez, S. D., & Gallup, G. G. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution, 10,* 175–188.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs, 123,* 133–167.

Suddendorf, T., & Whiten, A. (2001). Mental evolution and development: Evidence for secondary representation in children, great apes, and other animals. *Psychological Bulletin, 127,* 629–650.

Tager-Flusberg, H. (1991). Semantic processing in the free recall of autistic children: Further evidence for a cognitive deficit. *British Journal of Developmental Psychology, 9,* 417–430.

Tomasello, M., & Call, J. (1997). *Primate cognition.* New York: Oxford University Press.

Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends in Cognitive Sciences, 7,* 153–156.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53,* 1–25.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72,* 655–684.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory, the frontal lobes, and autonoetic consciousness. *Psychological Bulletin, 121,* 331–354.

Whiten, A., & Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences, 11,* 233–273.

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.

Yirmiya, N., Solomonica-Levi, D., & Shulman, C. (1996). The ability to manipulate behavior and to understand manipulation of beliefs: A comparison of individuals with autism, mental retardation, and normal development. *Developmental Psychology, 32,* 62–69.