

## II How Does Joint Attention Work? Mechanisms and Processes



## 9 Sociality, Attention, and the Mind's Eyes

Stephen Shepherd and Massimiliano Cappuccio

They sit together, one with his arm slung over the other's neck. They cluck over an infant—a boy—and they raise and inspect him, chattering over his sex. As they stare intently, one adult kneads the other's shoulder. Except for one detail, we could be forgiven for assuming the two *jointly attended* their newly arrived companion. However, this detail is key: the individuals in question are not humans, but Barbary macaques.

Many nonhuman animals seem sensitive to the attentional states of others, but joint attention is thought to be uniquely human, intimately tied to our ability to represent the minds of others. While the precise definition of joint attention continues to be debated, it involves (at minimum) attention to another individual and to the subject of his or her attention. In this chapter, we will explore the behavioral and neural evidence regarding animals' understanding of others' attention, in particular that implied by gaze following. We will conclude by describing a crucial aspect of human sociality—the use of pointing gestures to manipulate states of joint attention—and discuss what this gesture implies about differences between human and nonhuman social cognition. We provisionally distinguish four levels of socially cued attention:

1. *Gaze following, or the tendency to look where we see someone else looking* For ourselves—or, indeed, any animal—to know the subject of another's attention, we must evaluate how he or she directs his or her sense organs. In the visual domain, for example, this involves perceiving his or her gaze, translating between his or her reference frame and our own by replicating or simulating the other's viewpoint, and extending our attention to include the other's putative visual focus. The ability to follow attention is prerequisite to the ability to jointly attend, or moreover, to infer another's intentions and goals from his or her bodily behavior. Nevertheless, this level of socially cued attention is merely a reorienting of attention in the direction attended

by another—and can therefore occur with neither awareness nor intervening conceptual understanding of the other's mental or intentional state.

2. *Seeing gaze as intention* At the next level of sophistication, an individual A observing an individual B focusing its attention on C recognizes B and C to be linked. This linkage may be achieved through simultaneous, split attention toward both B and C or through some direct perception of an intentional relationship B→C. No intrinsic content need be ascribed to this linkage, which is not representational in character. Nevertheless, linking the individual with that individual's attention focus makes possible the ascription of enactive intentions such as *seeing*, *wanting*, and *responding*. An individual attending B *looking toward* C might become aware from this that B *sees* C, that his or her expressed affective state is *about* C, and that these things predict certain types of *interaction* with C.

3. *Nonrepresentational joint attention: perception of collective attentional states* The perception that another's attention refers to the same focus as our own affords certain uniquely social behaviors. Without another individual's attending to the same object as ourselves, it is impossible to compete or cooperate. Perhaps more compellingly, the perception that another individual attends to the same focus as ourselves makes possible social affordances such as *outnumbering our foe*. Many animals work in groups to defend against and drive off threats—predators, rival social groups—and individuals may regulate their participation in these behaviors through nonrepresentational recognition of collective action (see also Gallagher, this volume, and Hutto, this volume). However, instances of joint attention recognized in this way would be constrained by each animal's repertoire of species-typical social behaviors.

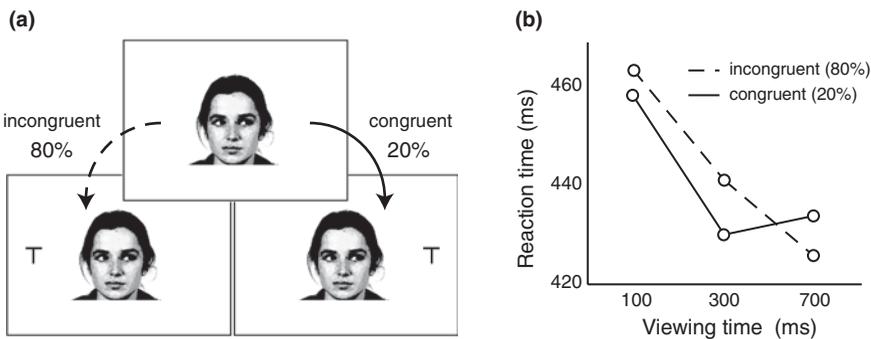
4. *"Full-blooded" joint attention: participation in a shared mental state* In conventional usage, joint attention requires mutuality. This is more than mere *cooperation* or *competition*, for example: I can compete with an individual who is unaware of me, as long as I'm aware our actions conflict. Moreover, I can cooperate, taking that individual's side in a conflict while he or she is unaware that I'm behind him or her, threatening his or her foe. Humans, however, can go farther, treating joint attention as a state in which each party participates in the representation of some "common ground" (Tomasello, 1995, 1999, 2008, pp. 94–95). Humans learn to actively manipulate these mutual representational states, shaping their content through the use of declarative pointing gestures.

It is broadly agreed (but not universally agreed; see Leavens, this volume) that only humans are capable of full-blooded joint attention, but what of

these proposed subordinate components? If animals cannot fully jointly attend, can they recognize communality, mutuality, or similar social states? What do other animals understand of the dyadic relations between animals and their distal behavioral goals? Which species follow their companion's gaze, and is their gaze following behavior reflexive and unconscious or flexible and goal directed? In the following discussion, we will first survey animal gaze following behaviors, then discuss possible neural correlates of these behaviors, and finally return to the question of how human joint attention—and in particular our manipulation of joint attention through gesture—compares to those forms of socially cued attention which are widespread among nonhuman species.

**Following Gaze**

To respond to gaze, it must first be perceived—but at least theoretically, this perception need not be precise, conscious, or intentional. Using a simple computerized cuing task, Friesen and Kingstone (1998) demonstrated that eye gaze cues influenced subjects' attention within a tenth of a second; moreover, attention was allocated in the direction of gaze despite the fact that the gaze cues had no predictive value and were thus irrelevant. Langton and Bruce (1999) found similar results using observed head orientations, and Driver and colleagues (1999) strengthened the case for automaticity, reporting subjects follow gaze even when explicitly informed that cues are *counter*predictive of target location (see figure 9.1). These



**Figure 9.1** Reflexive gaze following. Humans respond faster to targets which appear congruent to previously observed gaze, even when targets consistently appear at the incongruent location (from Shepherd, 2010, based on Driver et al., 1999).

results indicate that not all gaze following responses are susceptible to willful modulation. Perhaps more importantly, these findings show that gaze following requires only those cognitive resources that arise within a few hundred milliseconds of stimulus onset.

The simplest act of gaze following would be to identify a directional visual pattern—an averted nose, or the offset concentric circles of averted eyes—and to shift attention in that direction until a thing of interest is seen. It could be fast, simple, and purely reflexive: one needn't even know that one had done it. Moreover, it might be automatic: one might be unable to stop it, even if one knew the cue would be misleading. Gaze following thus might be a fully modular behavior: reflexive, and once triggered, unalterable.

Joint attention requires more than merely following gaze or otherwise replicating another's attention: one must, at the very least, simultaneously attend to the cuing individual. This suggests we should dissociate reflexive gaze following (sometimes termed "gaze cuing") from more sophisticated understanding of another's eye movements. Further, it prompts the question of how social perception and representation relate: how "looking" grades into "seeing," or how "attending" grades into "knowing," "wanting," or "believing"—or, saliently, "sharing." Nonetheless, a middle ground intervenes between these two phenomena.

It is our impression that we do not follow gaze *solely* as an unconscious reflex. Upon seeing a group of people excitedly craning their necks, we often strain to determine what they're looking at: our attempt to follow gaze is *goal directed*, and informed by clear *expectations* as to what might and might not be the focus of their excitement—and, moreover, as to whether the focus of their excitement is something about which we truly care. To accurately follow gaze, we interpret another's bodily orientation in spatial relation to our own. This rich understanding of others' bodies and of three-dimensional visual space permits us to take their perspective, following their gaze *geometrically* to objects outside our immediate visual field.

Gaze-following behavior (level 1) in many animals likewise involves these features and so appears to transcend simple reflexes, suggesting a more flexible understanding of other's attentional states. The ability to follow gaze *geometrically*, and to modify gaze following behavior based on *expectations* and *goals*, implies some sort of understanding that others are linked to the subjects of their attention (level 2) and appears to build toward those social abilities implicit in joint attention (levels 3–4). What does the gaze following behavior of animals tell us about their understand-

ing of attention? What, if anything, can we infer about the neural substrates of this understanding?

### Gaze Following in Nonhuman Animals

Social attention abilities have been differently explored in humans, primates, and nonprimate animals, respectively emphasizing connections to language, theory of mind, and learning signals. Nonetheless, responses to gaze can, to some extent, be unified in a common framework (Shepherd, 2010). Overwhelming evidence suggests that sensitivity to direct gaze—to being watched—is early developing and phylogenetically widespread, appearing soon after birth in humans (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000; Farroni, Csibra, Simion, & Johnson, 2002) and exhibited by diverse fish, reptiles, birds, and mammals (Senju & Johnson, 2009; Sowards & Sowards, 2002). However, while sensitivity to being watched has clear and obvious relevance to survival, the benefit of sensitivity to averted gaze is more subtle. Four factors seem paramount: first, animals must generally coordinate their behavior with that of their predators, prey, or mates; second, social animals must coordinate their movements with one another to maintain group coherence; third, conspecific behaviors may suggest adaptive responses to yet unnoticed environmental features (cf. Shepherd, Steckenfinger, Hasson, & Ghazanfar, 2010); finally, social primates communicate and negotiate alliances through facial expression and gaze behavior and may eavesdrop on the signal exchanges of others to better monitor their social environment.

Evidence for gaze following by animals comes from the study of an individual's orienting responses to head or eye postures of an experimenter or of a covertly manipulated conspecific. Ideally, these orienting movements are spontaneous rather than reinforced since in principle any sensible stimulus can be used for conditioning (albeit more easily when stimuli are salient to natural behavior). To date, evidence has strongly supported gaze following abilities in most primate species, as well as some domesticated species, trained marine mammals, and birds.

### Following Gaze

For example, Michael Tomasello and colleagues captured the attention of a cue individual by discretely showing it food, then observed the responses of an otherwise oblivious conspecific over the next ten seconds. For each of the five primate species tested (chimpanzees, mangabeys, and three types of macaque), individuals were significantly more likely to look

toward the experimenter after seeing a companion looking in the same direction (Tomasello, Call, & Hare, 1998). Two studies extend this data to other primate groups, providing evidence consistent with gaze following: one found that pair-tested cotton-top tamarins (a new world monkey) gazed in similar directions while viewing experimenter cues and that the pairs' gaze was most strongly correlated among individuals who typically resided together (Neiworth, Burman, Basile, & Lickteig, 2002); another found that ring-tailed lemurs (a prosimian or "primitive" primate) look in similar directions during natural behavior and are especially likely to co-orient with individuals toward whom they've just attended (Shepherd & Platt, 2008). Additionally, several studies have examined the tendency to follow conspecifics' gaze among nonprimate species. Using Tomasello's paradigm, Kaminski and colleagues showed that goats follow other goats' gaze (Kaminski, Riedel, Call, & Tomasello, 2005). Moreover, two types of birds—ravens (Schloegl, Kotrschal, & Bugnyar, 2007) and ibises (Loretto, Schloegl, & Bugnyar, 2010)—appear to follow one another's skyward glances, and, similarly, dogs appear able to use other dogs' attention when searching for hidden food (Hare & Tomasello, 1999).

Most commonly, however, gaze following has been studied using human-produced directional cues such as eye movement, head movement, pointing gestures, tapping, and so on. Animal subjects have sometimes been tested simply for spontaneous orienting behavior but more often have been challenged with an opportunity to seek or avoid food at cued locations. These studies suggest that apes and monkeys (see Rosati & Hare, 2009) and human-reared ravens (Bugnyar, Stöwe, & Heinrich, 2004) spontaneously follow human gaze while dogs (Hare & Tomasello, 1999) and dolphins (Pack & Herman, 2004) have both been reported to follow human gaze when rewarded. Further work has suggested that while many domesticated animals (including cats, dogs, goats, horses, seals, and dolphins; see Miklósi & Soproni, 2006) follow human cues to get food, wild animals may more typically *avoid* food that has been targeted by another (Flombaum & Santos, 2005; Hare & Tomasello, 2004). Together, these findings suggest that sensitivity to gaze direction may be widespread but that animals may learn to use this information in very different ways.

### **Understanding Gaze: Geometry, Expectation, and Goal Directedness**

The question remains, however, what animals truly *understand* about gaze. Animals might co-orient because of an innate and modular gaze following reflex, because they have been conditioned that gaze predicts the location of rewards or salient features, or because they use others' gaze to infer

something about their shared environment or about the others' behavioral or mental states. Psychophysical comparison suggests that monkeys and humans follow head and eye gaze quickly, independently of task relevance, using neural pathways that operate within a few hundred milliseconds (Deaner & Platt, 2003). Such gaze cuing is consistent with a behavioral module, operating with neither volition nor awareness. However, like those of humans, animal gaze responses sometimes betray surprising sophistication.

First, most animals appear to understand others' gaze as a vector within a rich, three-dimensional environment. In response to humans staring behind a barrier, or to a distant object beyond a closer distractor, most tested species have been shown to look toward the true target of the humans' gaze (great apes: Bräuer, Call, & Tomasello, 2005; new world monkeys: Amici, Aureli, Visalberghi, & Call, 2009; Burkart & Heschl, 2006; even ravens: Bugnyar, Stöwe, & Heinrich, 2004). Recent investigation of great apes suggest spatially sophisticated gaze comprehension, with ability increasing from orangutans to gorillas, chimpanzees, and finally bonobos (Okamoto-Barth, Call, & Tomasello, 2007). In only one case—that of a bird, the ibis—have adults of a species been shown to follow gaze but to be confused by barriers (Loretto, Schloegl, & Bugnyar, 2010). On the other hand, developmental trajectories suggest that gaze following arises before a rich spatial representation is present. Human infants begin to co-orient with adults within the first year but only later gain the ability to follow gaze toward objects outside their immediate field of view or hidden behind barriers (Butterworth, 1991; Moll & Tomasello, 2004). Moreover, gaze processing accuracy continues to improve throughout human children's first six years (Doherty, Anderson, & Howieson, 2009). While comparative data are sparse, a report in ravens suggests nonhuman animals follow a similar trajectory (Bugnyar, Stöwe, & Heinrich, 2004).

Second, like humans, many animals appear to have expectations about what they should see when they follow another's gaze: if they find nothing, they do a double take. These expectations appear to be shaped by experience. For example, old-world monkeys and apes following gaze cues toward nothing, check back toward the cue, and then follow gaze again (Bräuer, Call, & Tomasello, 2005; Call, Hare, & Tomasello, 1998; Goossens, Dekleva, Reader, Sterck, & Bolhuis, 2008; see also Horton & Caldwell, 2006); moreover, adult apes have been shown to do so more than infants even after controlling for baseline gaze following frequency (Bräuer, Call, & Tomasello, 2005). These experiments have been performed

in few, if any, nonprimates, and even among primates the data are somewhat equivocal: while apes and old-world monkeys are reported to “double look,” negative results have been reported in new-world monkeys (Amici, Aureli, Visalberghi, & Call, 2009), and the apes vary in the sophistication of their expectations regarding gaze (Okamoto-Barth, Call, & Tomasello, 2007).

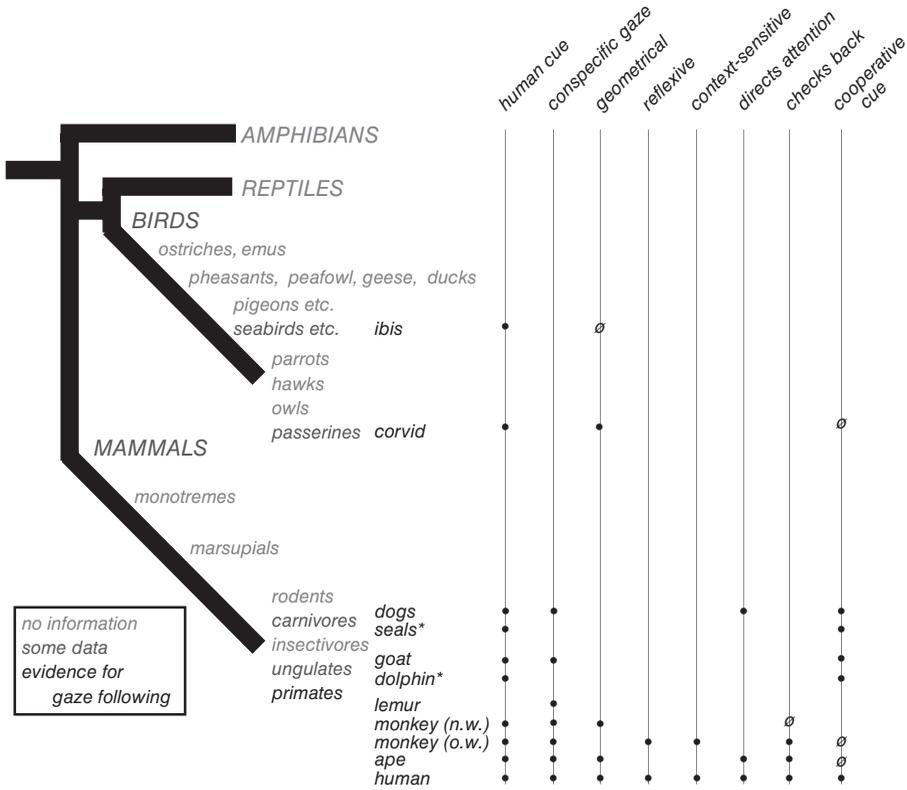
Third, despite evidence for fast, reflexive gaze cuing in primates (Deaner & Platt, 2003), it appears that both human and nonhuman animals modulate their gaze following behavior based on context: gaze is followed more readily when coming from informed, familiar, emotionally expressive, high-status, or otherwise exceptional individuals. In reality, the issue of context-sensitivity folds together two distinct questions. First, are fast, reflexive gaze responses the result of a dedicated, encapsulated module for gaze following? Second, can animals flexibly integrate observed gaze with other clues to an animal’s mental state or social significance? The first question is addressed by psychophysical findings that the strength of gaze cuing is modulated by the social status of the observed individual (monkeys: Shepherd, Deaner, & Platt, 2006; humans: Jones et al., 2010), though these effects may spare the very fastest (less than 200 milliseconds) responses (Shepherd & Platt, 2009). These findings limit the potential modularity of any gaze following mechanism (see Shepherd, 2010, for discussion) and suggest that within a split second, gaze following behaviors are modulated by the social relevance of the cue.

Moreover, co-orienting in monkeys appears to be enhanced by specific facial expressions (Goossens, Dekleva, Reader, Sterck, & Bolhuis, 2008; Teufel, Gutmann, Pirow, & Fischer, 2010; cf. Paukner, Anderson, Fogassi, & Ferrari, 2007), suggesting that gaze following in natural behavior is well integrated with other forms of social response. These findings suggest that animals may follow gaze differently depending on a cue’s social context and perceived relevance. In particular, they are reminiscent of results from human infants, in which gaze following behaviors are potentiated by “ostensive” cues suggesting specific communicative intent: infants more strongly follow gaze when associated with direct gaze or infant-directed speech (Farroni, Mansfield, Lai, & Johnson, 2003; Senju & Csibra, 2008). Gaze following behaviors may thus involve components with distinct psychophysical and developmental time courses, and while the fastest responses may be reflexive and modular, later gaze following behaviors (still arising within a second) are sensitive to contextual factors including those which suggest communicative significance.

### Gaze Following: Learning and Development

These “advanced” gaze-following behaviors appear to have extended developmental time courses, both in human and nonhuman animals. This suggests that the mechanisms underlying our ability to understand others’ viewpoints, like most cognitive abilities, develop through the interaction of genetic predispositions and learning mechanisms with our experienced social environment. Complicating things further, different components may follow distinct ontogenetic trajectories. For example, monkey gaze following becomes increasingly prominent between five months and one year of age (Teufel, Gutmann, Pirow, & Fischer, 2010). While adult monkeys and chimpanzees habituate to gaze cues directed toward empty space, infants do not (Tomasello, Hare, & Fogleman, 2001), suggesting development increases the discretion and selectivity of gaze following behavior. However, from the start, gaze following is facilitated when cuing individuals issue an expressive “comment” on observed interactions—and this facilitation is strongest for monkey observers around one year of age as they begin to navigate society independently of their mothers (Teufel, Gutmann, Pirow, & Fischer, 2010). Thus, while habituation to useless gaze cues increased over development, modulation by vocal expression peaked at a crucial developmental landmark. Similar complexity has been reported in the development of gaze following by ravens: juvenile ravens habituate to irrelevant cues during simple gaze following, but older ravens following gaze geometrically do not (Schloegl, Kotrschal, & Bugnyar, 2007). In this case, the authors speculate, early-developing gaze following is specialized for predator detection while later-developing geometrical gaze following is specialized for another function: pilfering the hidden food caches of other ravens.

In summary, many primates and some birds and domestic animals appear to be capable of interpreting and responding to gaze—and many other species remain untested (see figure 9.2). Fast gaze cuing may be quick and reflexive, consistent with mediation by an encapsulated neural module shared across (at least) higher primates. However, the spatial and social sophistication of gaze following behavior suggests it is goal directed, flexible, and generally well integrated with other cognitive processes. In all species studied, gaze following behaviors appear to be constructed during early life. However, gaze cues are not used in all the contexts in which we imagine they might be helpful, and the use of gaze as a collaborative signal seems uncommon in the animal world—seemingly restricted to coordinating aggression or antipredator behavior, for example, rather than



\* trained/enculturated individuals only

**Figure 9.2**

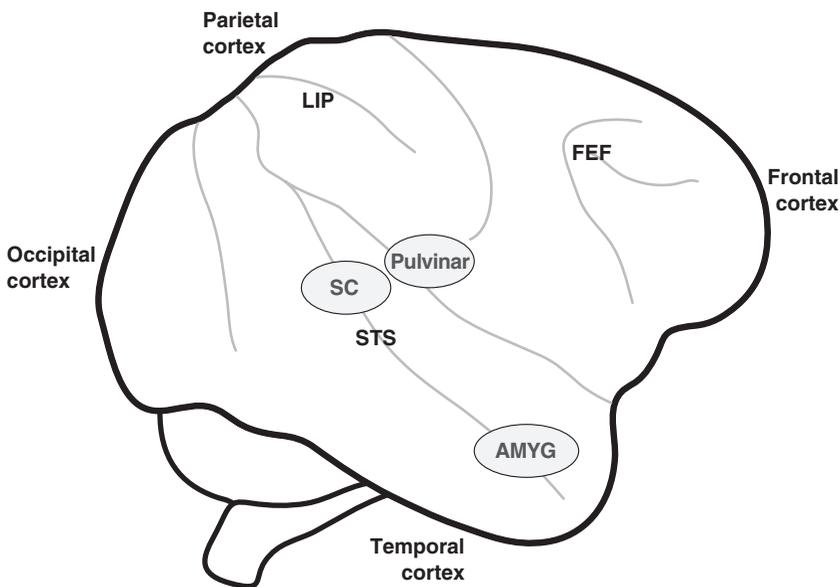
Family tree of gaze following abilities. Positive evidence for gaze following abilities is illustrated with a small black dot, negative evidence with a slashed circle. Many species remain untested (adapted from Shepherd, 2010). n.w., new world; o.w., old world.

to acquiring food, resources, or mates. Moreover, animals—and especially adult animals—lack the extraordinary and pervasive interest in others’ activities displayed by human children (Tomasello, Carpenter, Call, Behne, & Moll, 2005; but see Pika & Zuberbühler, 2008; Tanner & Byrne, 2010). This suggests that while many wild-reared animals appear to perceive the relational and referential aspects of attention, they have limited ability or opportunity to learn emergent affordances of socially cued attention, with these abilities generally restricted to group defense and coalitional conflict. We shall explore these social affordances in more detail in our consider-

ation of pointing gestures. First, however, let us consider possible neural substrates for forms of socially cued attention which are widespread among nonhuman animals.

### Gaze Following by Neural Circuits

When most people think of a brain, they think of the convoluted mass that comprises the human cortex (see figure 9.3). Those of us who study brains may be inclined to overemphasize the uniqueness of the cerebral cortex, a superficial and conspicuous layer of the brain useless without the important and complex subcortical structures beneath it. Moreover, while primates and especially humans have elaborately developed cortices, homologs of mammalian cortex are found within other vertebrate groups (Allman, 1999; Wang, Brzozowska-Prechtel, & Karten, 2010). It is thus



**Figure 9.3**

Neuroanatomical regions of interest. The pathways used in gaze following remain unknown, but several candidate areas may play a role. In this sketch, lobes are identified in plain text, cortical areas in bold (LIP, lateral intraparietal area; FEF, frontal eye fields; STS, superior temporal sulcus), and subcortical nuclei in outline (SC, superior colliculus, homolog of bird's optic tectum; Pulvinar, pulvinar nucleus of the thalamus; AMYG, amygdala).

possible that the neural mechanisms of gaze following are derived from some gaze following ancestor and are homologous across all occurrences. It appears more likely, however, that circuits for gaze following behavior have developed independently (i.e., are analogous) between divergent species such as birds and primates.

### The Subcortical Pathway

The fast time course of reflexive gaze cuing suggests that a fairly short and direct circuit translates gaze observation to orienting movements. It has been theorized that the fastest reactions to basic, biologically crucial stimuli are through an innate subcortical system that is broadly conserved across vertebrates (Johnson, 2005; Sowards & Sowards, 2002). This subcortical system is thought to be interconnected with slower and more nuanced cortical networks, suggesting that while reflexive gaze cuing may be mediated by the former, the latter may mediate the more recently evolved, late developing, or computationally challenging features which characterize naturalistic gaze following.

The mammalian subcortical visual pathway is believed to pass from the retina to the superior colliculus, the pulvinar, and then the amygdala (Jiang & He, 2006; Johnson, 2005; Morris, Ohman, & Dolan, 1999). The amygdala is a sensory convergence zone associated with emotional processing and can influence attention both directly, through modulation of early sensory processing, or indirectly, via projections to social and reward cortices in the temporal and frontal lobes (Amaral, Behnia, & Kelly, 2003; Ghashghaei & Barbas, 2002; Vuilleumier, 2002a). Moreover, the superior colliculus and the pulvinar nucleus of the thalamus influence attention in their own right, serving, respectively, as a final common output pathway for gaze orienting (Platt, Lau, Glimcher, Hall, & Moschovakis, 2004) and as a partner to ongoing cortical processing (Saalmann & Kastner, 2009). While the amygdala is known to be sensitive to observed gaze (monkey electrophysiology: Gothard, Battaglia, Erickson, Spitler, & Amaral, 2007; monkey imaging: Hoffman, Gothard, Schmid, & Logothetis, 2007; human imaging: Kawashima et al., 1999), it is not known whether this information derives from cortical or subcortical signals, nor is it clear whether the amygdala distinguishes between directions of averted gaze or merely between averted gaze and direct staring. While lesions of the amygdala have been reported to disrupt human gaze responses (Akiyama et al., 2007; Okada et al., 2008), these disruptions may be secondary to damaged white matter or decreased social attention.

A homolog of the superior colliculus—the optic tectum—is the principle visual processing center of fish, amphibians, reptiles, and birds (Allman, 1999; Sowards & Sowards, 2002), but in mammals, these subcortical structures are believed to access only the coarsest and least-digested ascending visual input. This makes it unlikely that mammalian subcortical systems are sufficient to mediate spatially sophisticated behaviors such as geometrical gaze following. In contrast, two meta-analyses have identified regions sensitive to gaze direction in human cortex, most notably near the superior temporal sulcus (STS) and in the dorsal and ventral frontoparietal attention networks (Grosbras et al., 2005; Nummenmaa & Calder, 2009).

### Social Processing in Mammalian Cortex

While it is possible that social processing centers include cortices homologous across all primates (Rosa & Tweedale, 2005; Tootell, Tsao, & Vanduffel, 2003) and perhaps other mammals (Kendrick et al., 2001), it is clear that great diversification has occurred across species during mammalian evolution. Two of the best studied mammalian species, macaques and humans, are believed to share similar visual systems; even so, the layout of their socially responsive cortices has proved difficult to reconcile (Op De Beeck et al., 2008; Pinsk et al., 2009; Tsao, Schweers, Moeller, & Freiwald, 2008). Broadly speaking, different social processing regions are believed to specialize in features that are relatively static (such as identity) or dynamic (such as expression, gaze, and biological motion; Hoffman & Haxby, 2000); in humans, the former is associated with the fusiform gyrus (Kanwisher & Yovel, 2006) and the latter with the posterior portion of the STS (Allison, Puce, & McCarthy, 2000). In monkeys, most socially responsive cortices appear to lie along the STS, starting deep in the sulcus posteriorly and dipping slightly ventral anteriorly, with several additional patches in the inferior prefrontal lobes (Pinsk et al., 2009; Tsao, Schweers, Moeller, & Freiwald, 2008). The functional specialization of these patches is currently undergoing intensive study and may be mapped, at least in outline, by 2015.

Already, however, anatomical and electrophysiological research in the temporal lobes provides rich information about gaze processing. STS comprises several cortical regions with distinct cellular structure and connectivity, and while posterior regions communicate both with posterior parietal and frontal areas, anterior regions communicate preferentially with frontal and visual cortices (Seltzer & Pandya, 2004; Seltzer & Pandya, 1991). The first neurons sensitive to observed gaze were reported in macaques near the STS (Hasselmo, Rolls, & Baylis, 1989; Perrett et al., 1985; Perrett, Rolls,

& Caan, 1982); moreover, Perrett and colleagues argued that the patchy distribution of gaze-responsive cells corresponded to the patchy distribution of STS neurons communicating with parietal orienting areas (Perrett, Hietanen, Oram, & Benson, 1992). Importantly, neurons in the middle anterior upper bank of the STS represent gaze direction independently of whether it arises through head or eye posture (Perrett, Hietanen, Oram, & Benson, 1992), and while more caudal neurons respond symmetrically to gaze averted to either the right or left, anterior neurons differentiate specific gaze directions (De Souza, Eifuku, Tamura, Nishijo, & Ono, 2005; see also Jellema, Baker, Wicker, & Perrett, 2000).

Human imaging studies support these findings. Though imaging studies have shown peak gaze sensitivity in the posterior STS (Allison, Puce, & McCarthy, 2000; cf. Grosbras et al., 2005; Nummenmaa & Calder, 2009), a recent adaptation study (Calder et al., 2007) showed direction-differentiating neurons concentrate anteriorly. Intriguingly, the STS is especially active in humans when we observe surprising or incongruous gaze behavior (Pelphrey, Singerman, Allison, & McCarthy, 2003): this covert activation may be analogous to overtly observed “check backs” or “double looks.” Human neural pathology provides further support for the role of the temporal lobes in perceiving gaze direction in space: two split-brain patients reflexively followed gaze in only one visual hemifield, consistent with cortical mediation by a single hemisphere (usually the right) specialized for face processing (Kingstone, Friesen, & Gazzaniga, 2000). Moreover, in a rare patient with a lesion affecting the entire right superior temporal gyrus, both gaze perception (Akiyama, Kato, Muramatsu, Saito, Nakachi, et al., 2006) and reflexively gaze following (Akiyama, Kato, Muramatsu, Saito, Umeda, et al., 2006) were absent.

In addition to temporal cortices strongly and selectively activated by faces, macaques appear to have three face-selective regions of frontal cortex not evident in humans (Tsao, Schweers, Moeller, & Freiwald, 2008). Interestingly, however, at least one human subject with lesions in orbitofrontal cortex lacked both gaze cuing and symbolic cuing (Vecera & Rizzo, 2004). While it is difficult to draw conclusions from a single subject, especially given interindividual variability in cuing behaviors, these reports may suggest that those frontal regions which process facial signals in monkeys have been generalized to process more diverse signals in humans. This possibility is especially intriguing given recent evidence that nonpredictive (but socially conventionalized) symbols, such as arrows, evoke reflexive cuing responses in adult humans (Kuhn & Kingstone, 2009; Tipples, 2002; Borjon et al., forthcoming).

It is crucial to recognize that perception of gaze is only a first step in gaze following: any extracted signal must next be converted either directly into an orienting response or indirectly into some sort of flexible knowledge. Moreover, just as attention may arise in other modalities than vision, joint attention might arise through other means than the sight of another's gaze. Our own routine gaze following responses transcend modality (Borjon et al., forthcoming), and blind individuals are capable of sharing attention, interpreting action, and attributing intent using similar brain networks to those of the sighted (Bedny, Pascual-Leone, & Saxe, 2009; Ricciardi et al., 2009; albeit with some developmental delay; see Hobson & Bishop, 2003). Both arguments suggest that gaze following must impact supramodal brains systems which coordinate attention and movement in space: most notably, frontoparietal action networks. These neural systems both direct attention and organize behavior—and have been implicated in understanding the attention and behavior of others.

### The Frontoparietal Orienting System

The muscles orienting the eyes are controlled from centers in the brainstem, which receive projections from the superior colliculus and frontal eye fields. These centers, in turn, are interconnected with the lateral intraparietal area (LIP) and supplementary eye fields, and together they comprise the dorsal attention network (Platt, Lau, Glimcher, Hall, & Moschovakis, 2004). This system is likely the final step in converting gaze perception into an orienting response: that is, making the focus of another's orienting system the focus of our own.

This step could occur through a great number of pathways. In macaques, projections from gaze-sensitive regions of the STS to cortical areas 7A and LIP comprise one likely route through which observed gaze could influence orienting (Perrett, Hietanen, Oram, & Benson, 1992; see also Seltzer & Pandya, 2004; Seltzer & Pandya, 1991). LIP is particularly interesting because its neurons respond to both direct stimulation of, or orienting movements toward, their spatial response field (RF; Colby & Goldberg, 1999). Intriguingly, neurons in LIP are further modulated by the behavioral relevance of their RF, reflecting both the juice reward value of trained eye movements (Platt & Glimcher, 1999) and the intrinsic reward of acquiring social information (Klein, Deaner, & Platt, 2008). LIP constitutes a unified spatial saliency map, and neuronal activity is associated both with covert attention and overt motor intention toward a region (Colby & Goldberg, 1999).

Shepherd and colleagues (Shepherd, Klein, Deaner, & Platt, 2009) recorded the activity of LIP neurons while monkeys performed a simple test of gaze cuing. Monkeys fixated a central location, a face appeared gazing one direction or another, and then a target appeared unpredictably in either the cued (“congruent”) or uncued (“incongruent”) direction. When trials were congruent, monkeys were only slightly faster than when trials were incongruent—perhaps because the cuing effect is smaller in monkeys than in humans or perhaps because these monkeys had performed many thousands of trials and were practiced at suppressing the irrelevant gaze-cuing response. Nonetheless, gaze-cuing effects were significant. Classically, LIP neurons would be expected to ignore the gaze cue and instead fire only when targets appeared in their RF: one great burst bridging target onset and saccade initiation. Instead, we found that a large subpopulation of neurons had a small but significant response to gaze cues presented outside, but directed toward, their RF. Generally, these responses began within 100 to 200 milliseconds of the cue onset and persisted through the cue period. We saw some neurons that were enhanced and some neurons that were suppressed by gaze. The population that was enhanced responded with a time course quite similar to the behaviorally observed gaze following response, peaking between 200 and 400 milliseconds and decaying by 800 milliseconds. The inhibited population was tonically active throughout the cue period.

How are we to interpret these findings? The enhanced neurons would be expected to contribute to gaze following, the suppressed neurons would be expected to dampen it—and both effects were plausible, given that the monkeys had been extensively trained to ignore the irrelevant gaze cue. Close examination of the modulation dynamics suggested that the enhanced neurons may have been secondary players, however. If LIP were mediating the fastest and most reflexive gaze-cuing responses, enhanced activity would have led the behavioral effect by thirty milliseconds: it did not. In summary, these findings support the claim that gaze cues modulate LIP neurons from outside their classical RFs and suggest that observed gaze changes the value of orienting toward the cued location. Thus, even if LIP is not involved in the fastest gaze cuing effects, it is almost certain to contribute to the influence of social context on gaze following (e.g., Goossens, Dekleva, Reader, Sterck, & Bolhuis, 2008; Shepherd, Deaner, & Platt, 2006; Teufel, Gutmann, Pirow, & Fischer, 2010).

Consistent with these findings, human neurons that track observed gaze directions are clustered in two places: STS and posterior parietal cortex (Calder et al., 2007). However, homology between macaque and human

brains remains uncertain, with the relevant cortical areas greatly expanded in humans relative to macaques (Orban, Van Essen, & Vanduffel, 2004). Complicating things further, there is not just one but two frontoparietal attention systems: dorsal and ventral (Corbetta & Shulman, 2002). While monkey electrophysiology implicates the dorsal orienting system, human joint attention tasks often activate a right-lateralized ventral network that strongly resembles the ventral attention network identified by Corbetta and Shulman—including the right temporoparietal junction and right lateral prefrontal cortex—along with medial frontal cortices (Calder et al., 2002; Grosbras, Laird, & Paus, 2005; Nummenmaa & Calder, 2009; Saxe, 2006). Monkey homologs of these areas remain uncertain. Finally, a study of human patients suffering visuospatial neglect suggests that these orienting networks—or at least, the parietal components thereof—are not essential to reflexive gaze cuing. In fact, not only did these patients exhibit normal gaze-cuing effects, they could be induced to follow gaze from their preserved hemifield into that which they more typically neglected (Vuilleumier, 2002b).

### *Reflecting Gaze*

To follow someone's gaze is to replicate their direction of attention. Such a shift in our attentional state can be achieved overtly and observably (by changing our bodily orientation) or covertly and privately (by changing solely our neural activity). Changes in overt orienting alter the physical coupling between our bodies, brains, and environment and, therefore, necessarily alter the contents of our sensory processing. Likewise, our covert attentional deployments alter our cognition, affecting future orienting decisions and subtly influencing our postural stance (e.g., microsaccadic drift: Engbert & Kliegl, 2003; Hafed & Clark, 2002). Moreover, it now appears that we use overlapping neural tissues to physically orient our eyes and to covertly shift our attention (e.g., Moore & Fallah, 2004). Because attention systems receive inputs from many cortical areas, there are many pathways through which social cues could exert their influence on gaze, and with no clear rubric for estimating which route might be faster or more flexible than another.

Importantly, however, the orienting system is embedded in a larger cortical network for interacting with the environment. A broad family of frontoparietal circuits link the dorsal (spatial or action-oriented) visual system and exterior body-sense areas, both associated with the parietal lobes, to frontal systems that evaluate, organize, and execute body movements. Seen through this lens, gaze following is but one of many

mechanisms which couple individuals' behavioral states. Contagion between observed and enacted behavior has been described for skeleto-motor action (Blakemore & Frith, 2005), facial expression (Dimberg & Thunberg, 1998), affective state (Hatfield, Cacioppo, & Rapson, 1993), sensory orienting and attention (Shepherd, 2010), and even perceptual states (Borjon et al., forthcoming). It is important to note that while this coupling is mediated through observable body states, it operates at a certain level of abstraction: our goals and emotions are more faithfully replicated than the postures or communication signals which conveyed them (Tamietto et al., 2009).

Some of the most interesting neural devices which may play a role in relating observed and executed behavioral states are the "mirror" neurons found in the prefrontal (area F5; Rizzolatti et al., 1996; Gallese et al., 1996) and parietal (inferior parietal lobule, mainly area PF; Fogassi et al., 2005; Gallese et al., 2002) cortices of macaque monkeys. These neurons respond to the execution of hand and mouth actions at various levels of abstraction, often sensitive to action goals rather than mere postural shifts (Umiltà et al., 2001, 2008). Moreover, these neurons are accompanied by "canonical neurons," which fire both during actions and when action-relevant objects are presented: they appear to categorize affordances, or potential actions, rather than actions currently being performed (Murata et al., 1997). Since they conflate self-produced and witnessed behaviors, mirror neurons are sometimes claimed as a crucial mechanism through which we interpret the behavior of others (e.g., by using our own motor capabilities to support a first-person subjective understanding of another's action goals; Di Pellegrino et al., 1992; Rizzolatti et al., 2001; Gallese, 2005a). This hypothesis is hotly debated. While there are various lines of evidence that that one's own motor expertise supports the understanding of others' actions (Beets et al., 2010) and that mirror neurons are both selectively and flexibly recruited during the recognition of goal-oriented actions (Gazzola et al., 2007; Cattaneo et al., 2007), objections have been raised on both philosophical and empirical grounds. These objections cite a lack of evidence that mirror neuron play a causal role in action recognition (Hickok, 2009) and claim no existing model can adequately explain how neurons mirroring another's action can support understanding of its intent (Borg, 2007; Csibra, 2007). Alternative explanations have also been suggested for mirror neurons: they may be produced by Pavlovian association, for example (Mahon & Caramazza, 2008; Catmur, Gillmeister, Bird, Liepelt, Brass, & Heyes, 2008) or may predict the next move an agent will perform to achieve a previously represented

goal (Jacob, 2008). (A rebuttal to these objections can be found in Rizzolatti & Sinigaglia, 2010.)

Shepherd et al.'s 2009 report of gaze "mirror" neurons makes this debate immediately relevant to the phenomenon of joint attention. Shepherd and colleagues identified neurons which were active both when a monkey looked a certain direction and when it saw a photographed monkey looking the same direction. If these neurons are, in fact, a kind of mirror neuron, then mirror processes may operate both for directly manipulative acts and for those (e.g., gaze) whose physical enaction is subtle and distal and whose consequences are covert. In our discussion of comparative gaze following abilities, we identified three markers of sophistication in gaze following: *geometry*, *expectation*, and *goal directedness*. Gaze mirror neurons could, in principle, accommodate each of these features. Like frontoparietal motor mirror neurons, gaze mirror neurons would be expected to operate at a certain level of abstraction: they should be sensitive to the environmental locus of attention rather than to the specific physical posture implementing a direction of gaze (e.g., eyes alone vs. head and eyes together). Similarly, because mirror neurons inhabit motor cortices focused on sensorimotor transformation across diverse bodily reference frames, it would be surprising if they were incapable of mediating geometrical gaze following. If mirror neurons, like canonical neurons, identify environmental affordances, then we should expect that they are associated with strong expectations. Finally, because the frontoparietal orienting networks intervene between perceptual and motoric cortices, we would expect minimal encapsulation, suggesting mirror-mediated gaze following would be highly context-sensitive and minimally reflexive.

To review, many animals have a rich capacity to respond to others' attention, as evident in their tendency to follow gaze (level 1) with their geometrical sophistication, expectations regarding where other animals will look, and context-dependent and goal-directed utilization of gaze cues (level 2). These abilities arise early in development, are sensitive to experience, and may be mediated by several neural pathways. Monkeys, like humans, exhibit both fast and reflexive gaze cuing and slower and more nuanced gaze following behavior. In monkeys, neurons in the LIP appear to differentiate gaze directions, with some acting to "mirror" gaze and some acting in opposition to gaze following. These data suggest that gaze is represented in the frontoparietal orienting system, which seems well-suited to mediate geometrical gaze following in a flexible and goal-directed manner (level 2). It is worth noting that if theorists are

correct in claiming mirror neurons permit a preconceptual and nondeclarative attribution of motor intentions (Gallese, 2005b), then they could support an understanding of nonrepresentational gaze as a form of motor intention. Gaze toward an object significantly increases the likelihood it will be acted upon, and while attention is not a strict precondition for transitive action, it is strongly associated with various forms of motor intent. Furthermore, gaze plays a significant role in the dyadic signaling of nonhuman primates and as such plays an *instrumental* role in signaling intention to interact with a target, typically in the context of aggression (gaze enacts a *threat*) or mating (gaze enacts *courtship*). For these reasons, it is not surprising that systems processing motor intention could be generalized to gaze signals (level 2). Crucially, by tracking the social contingencies necessary for species-typical triadic behaviors, these neurons could provide a nonrepresentational pathway for joint attention (level 3): *challenge* individual B for C, *team up with* individual B *against* individual C, *share* object C with individual B. The first two situations arise frequently (and the latter, rarely) in the social lives of group-living animals and comprise an important part of their natural behavioral repertoire.

If uniquely social affordances suggest a nonrepresentational mechanism for joint attention in animals, then how does this form of joint attention compare with that found in humans? A key feature of joint attention, as expressed by humans, is the ability to manipulate the focus of joint attention through gesture. However, this ability is not completely unique: animals, too, direct the attention of others through communicative signals. This aspect of socially cued attention—its active *exploitation* to meet behavioral goals—will be the subject of our final section.

### Sharing Attention through Gesture

Generally, some form of joint attention is said to arise whenever two observers are simultaneously and symmetrically aware of a common stimulus and of the other's attention toward that stimulus (e.g., Moore & Dunham, 1995). Because gaze following can synchronize attentional shifts toward a common stimulus, it is likely to play a pivotal role in achieving joint attention, provided two additional conditions can be met: first, the observed attention must be recognized as intentional in character; second, mutuality must somehow be recognized, for example, through unique social affordances or mentalistic representations permitting an inference of symmetry.

Developmental studies suggest that a human infant's ability to follow and subsequently to direct an adult's gaze predicts the infant's acquisition of full-blooded joint attention abilities over the following months (Carpenter, Nagell, & Tomasello, 1998). Joint attention among humans—even among young children—appears to involve rich representational systems and cooperative motivations that are rarely claimed of animals. While many animals appear to follow gaze, and while many appear to do so in a flexible and sophisticated manner suggesting they recognize gaze as intentional, they seem to lack the pervasive interest in others' attention which characterizes humans, and they appear to recognize social affordances only in limited and stereotyped contexts. Similarly, human infants younger than nine to twelve months of age seem incapable of jointly attending despite possessing the relevant perceptual and motor competencies from about the fifth month (Tomasello, 1999, p. 66). Something dramatic seems to change early in typical human development—and only atypically (if ever) in animal development—to promote the sharing of attention. How close do animals come, then, to achieving “full-blooded” joint attention? What additional features are necessary to construct human-like abilities from those building blocks we share with other species? Though the ancestral origin of our joint attention is highly uncertain, our ritualization of reaching gestures to manipulate others' attention (along with increasing reliance on imitative learning) seems to have played a crucial role (Tomasello, 1999, pp. 87–88).

### Joint Attention and Deixis in Humans

The human experience of joint attention, after all, entails much more than merely *following* it. The key difference is that the focal stimulus becomes “mutually manifest” to both coattenders (Tomasello, 1995; Clark, 1996). This mutuality implies that the attentional focus is openly accessible to each coattender and becomes salient by reason of a “common ground”: a shared context defined by the personal and cultural background of the individuals involved. In Tomasello's view (1995, 1999, 2008), both recognition of mutuality and conceptualization of the common ground are mentalistic in character, requiring perspective-taking and simulation skills believed to be unavailable to nonhuman primates (see also Meltzoff & Gopnik, 1993): this account builds on the intuition that while apes might only understand the meaning of a deictic gesture from the point of view of the signaler (“what the communicator wants for herself”), humans also understand the gesture from a self-reflexive point of view (“what the communicator thinks the recipients want for themselves”). Such an

understanding necessitates a mutually known framework of goals and interests: the common ground. This notion of joint attention is recursive in nature (Tomasello, 2008, pp. 94–95, regarding Grice, 1957, and Sperber & Wilson, 1986) and has been challenged as requiring infinite regress (Schiffer, 1972, and Campbell, 2005).

Tomasello argues (1999, 2008) that the crucial behavioral signature of joint attention, as experienced by humans, is the use of indicative gesture for the sole purpose of sharing attention. These gestures are arbitrary and produced merely to inform or declare: pointing something out or holding an object just to “show” it. The prototypical gesture in this family—index-finger pointing—is present in all known human cultures (Kita, 2003) and emerges in infancy at twelve months (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004). The extension of a finger toward a distant object is purely deictic and expressly communicative: it suggests no goal but communication, achieves no independent aim, and conveys no intrinsic content or affective valence. In contrast, both deictic gaze behaviors and proximal reaches accomplish personal objectives, and gaze, in particular, has affective gloss on account of its instrumental role in dyadic interaction and its integration into facial expressions. Human pointing is thus ostensive and arbitrary, given significance only by virtue of situational features and the presence of a common ground.

In this manner, human pointing can convey just about anything; it can even be performed to refer symbolically to an absent object or, when targeting a present object, to indicate an incidental feature or abstract association. In fact, human pointing is so conventionalized that it can be fully abstracted from the details of the immediate environment and given meaning solely through manipulation of the common ground: pointing can direct the imagination, designating and referencing locations in an arbitrary symbolic space (Kendon, 2004). Such gestural communication has clear ties to language (Corballis, 2003; Deacon, 1997), and indeed, early fluency with pointing and showing gestures predicts language competency later in infancy (Mundy et al., 2007); conversely, early disinterest in pointing and showing is a crucial diagnostic feature of childhood autism (American Psychiatric Association, 1994, p. 66).

In summary, human pointing gestures can be purely deictic, expressly communicative, arbitrarily conventionalized, altruistically/communally motivated, informative/declarative in character, and given meaning by dint of a representational common ground requiring full-blooded joint attention (level 4). Such manipulation of joint attention grades naturally into symbolic communication: use of conventionalized, arbitrary, and public signals.

### Joint Attention and Deixis in Animals

Animals, too, direct the attention of others. Many animals respond to specific threats through coordinated counterattack: for example, blue jays and mockingbirds mob larger birds such as owls and ravens in order to drive them off. Coordinating these responses, many animals have sophisticated systems of alarm calls: vervet monkeys, for example, distinguishing between (roughly) avian, mammalian, and ophidian threats (Cheney & Seyfarth, 1990). Producing certain types of alarm calls—upon seeing a snake or rival social group, for example—monkeys perform stereotyped “pointing” movements in which they shift position and abruptly freeze, staring intently at the target (personal observation, SS). Macaques similarly solicit third-party support in intragroup conflicts, using a series of alternating glances while loudly threatening a foe, and do so more often when the result of one-on-one combat is uncertain (“appeal-aggression”; de Waal, 2003). Because they help coordinate defense against adversaries, these signals are likely widespread among group-living animals.

However, while these signals appear to be both ostensive and deictic in character, they contrast with human pointing in several respects (Tomasello, 2008, pp. 34–41). Crucially, they are not arbitrary, as each signal intrinsically and specifically implies some immediate goal or motivation: for example, the ostensive signals generally convey affect (e.g., alarm calls), and the deictic signals, motor intention (e.g., staring for surveillance, tensely freezing in readiness to move). In several nonhuman primate species, evidence exists that hand gestures have been ritualized: capuchin monkeys make a food-cupping gesture to beg, even when the food is out of reach (de Waal, 2003). Moreover, the great apes appear to use open-handed reaching as a begging gesture even in nonfood contexts or situations in which the desired object is mechanically abstracted from the gestural target (de Waal, 2003). Chimpanzees, in particular, can point to direct another's attention (Leavens & Hopkins, 1998; Leavens, Hopkins, & Bard, 2005), but this behavior has been observed primarily in soliciting a human recipient to bring them some food they can't reach. Orangutans can also point to indicate a tool necessary to retrieve food (Call & Tomasello, 1994), but—differently from human children—they do it only if the object is of interest to their own behavioral goals rather than to their companion's (Bullinger, Zimmermann, Kaminski, & Tomasello, 2011). These findings suggest that even among the apes, deictic gestures rarely serve as an altruistically informative or communally declarative gesture (but see Leavens, this volume). Instead, these behaviors appear to be unified in their specific linkage to concrete, well-defined behavioral

responses. Such deictic behaviors can be termed “instrumental” or “imperative,” in that they result in specific physically observable (and generally selfish) end points.

This absence of altruistic/communal, informative/declarative gesture in nonhuman signalers may explain otherwise curious behavior. While many animals follow gaze, few use gaze appropriately in choice tasks. Apes appear to be able to read a human’s direction of selfish interest (Tomasello, Hare, & Agnetta, 1999), but in object choice tasks, in which food is hidden in one out of many possible locations, a human’s informative pointing is simply not understood (Tomasello, Call, & Gluckman, 1997). This fact is imputed by Tomasello (2008, p. 41, 2009) to the apes’ inability to attribute altruism. Interestingly, when informative/declarative pointing is distinguished from imperative/instrumental pointing as a predictor of human language ability, it is only the former sort—the sort rare or absent among nonhuman animals—that predicts language acquisition (Mundy et al., 2007).

Nonprimates, too, exhibit pointing behaviors. For example, dogs appear to actively direct human attention (Miklósi, Polgardi, Topal, & Csanyi, 2000; see also Miklósi et al., 2003), perhaps using skills crucial to group-hunting wolf ancestors (Miklósi & Soproni, 2006) and repurposed by human breeding (Hare, Brown, Williamson, & Tomasello, 2002) or early-life enculturation (Udell, Dorey, & Wynne, 2010). While formal studies have focused on pointing to acquire food rewards, dogs have been bred and trained to point for more abstract purposes (e.g., to spot birds or bombs).

Moreover, we may easily underestimate the use of deictic signals: it is challenging to distinguish the communicative significance of a signal which is not consistently coupled with triggering situations or behavioral responses. Instead, we might judge the signal to be merely a generic “greeting,” “contact call,” or “territorial marker.” For example, through a lengthy investigation, Teufel and colleagues (Teufel, Gutmann, Pirow, & Fischer, 2010) found that a generally overlooked “comment” call—most often produced by third parties witnessing a social interaction—appears to increase the likelihood their gaze will be followed by a monkey audience.

If it is true that we underestimate the rate of ostensive signaling by nonhuman animals, might we also underestimate the extent to which they have cooperative intent? Consider, for example, the case with which we opened this chapter. Among Barbary macaques, an extended audiovisual facial expression involving repetitive lip/jaw movements (the “chatter”) is “part of a typical behavioral pattern called triadic interaction or agonistic

buffering ... where usually two adults sit in body contact, hold one infant or little juvenile, and *chatter*. The adults approach their heads normally to the child while *chattering*, often lifting and lowering the head or turning the head from side to side (*head flag*). Frequently the animals *hug* and *knead* each other. Often they *chatter at* the genitalia of the child and touch the genitalia of the infant with lips or hands (*touch genitalia*)" (Hesler & Fisher, 2007, pp. 170–171). This behavioral pattern (see figure 9.4) is particularly interesting because in contrast to many triadic behaviors, it appears to have an abstractly prosocial goal. Nonetheless, these cases of triadic signaling among monkeys appear to lack arbitrariness and to be deployed in relatively limited and species-typical contexts rather than being used flexibly for informative or declarative purposes. In particular, there is little evidence to date that signaling behaviors are sensitive to the knowledge state of potential recipients (Cheney & Seyfarth, 1990). This negative finding challenges the notion that nonapes signal with the intention of informing or declaring to others. By contrast, great apes do appear somewhat sensitive to recipient knowledge: they generate more deictic gestures to guide a



**Figure 9.4**

Barbary macaques sit together, inspecting a nursing infant. Image courtesy Sébastien Wautelet.

human's behavior if there is evidence that the human is ignorant (Zimmermann, Zemke, Call, & Gomez, 2009). However, typically reared apes appear to do so only to obtain food rewards rather than as part of an altruistic, communal, or otherwise prosocial goal. The core finding seems to be that nonhuman animals lack the expectation that their peers will collaborate with them in flexible, rather than limited and species-typical, contexts. Even where sensitivity to knowledge states exists, nonhuman animals appear to have little incentive to shape others' knowledge for communal (as opposed to personal) ends.

Finally, the distinction between imperative and informative/declarative *production* of pointing implies a subtle difference in how humans and nonhumans *respond* to social attention cues. Human signalers actively shape their gestures so that they may be easily decoded by their target. As a result, an optimal human recipient needs to understand not only the *communicator's* intent but also the communicator's understanding of the recipient's own intent! According to Tomasello, our pervasive tendency to see ourselves from other points of view is crucial to our ability to cooperate in arbitrary endeavors. It is interesting to consider the possibility that attentional coordination through hand gesture may have been a crucial step in the evolution of "full blooded" joint attention (level 4). Manual gestures, unlike gaze gestures, can be simultaneously perceived by both signaler and recipient, affecting both in a similar (one might say mutual) fashion. Pointing, unlike gaze, is expressly ostensive in that it does not directly achieve a personal goal. Finally, pointing frees the eyes to move independently between the attentional target and any coattender (see Hopkins & Tagliabata, this volume, for more on this theme).

## Conclusions

Human joint attention abilities are crucial to our social behavior and appear fully integrated with mentalistic representation, language, and other cognitive abilities thought uniquely human. These abilities did not appear *de novo*, however, and are constructed atop underlying behaviors and neural circuits we share with other species. In particular, many organisms have the ability to respond behaviorally to another individual's attention—for example, by following its gaze into distant space. Moreover, many species (including corvid birds, domesticated mammals, and nonhuman primates) appear to follow gaze with geometrical sophistication, and, moreover, can do so in a flexible, goal-directed, and context-sensitive fashion. Perception of gaze may guide behavior, permitting uniquely social

affordances to arise: social animals often act jointly to drive away adversaries. These features suggest animals recognize gaze as a kind of intentional relationship and can coordinate species-typical social behaviors in response to these relationships.

Different animals may use very different neural systems to perceive and respond to gaze. Subcortical mechanisms may play a role in fast and stereotyped gaze responses. However, mammalian cortex plays a clear role in perceiving, interpreting, and responding flexibly to observed gaze. In particular, Shepherd and colleagues (2009) identified parietal sensorimotor neurons which fire both when attending to a region of space and when seeing another individual gaze there. Neurons in this area would be well situated to interpret gaze in three-dimensional space and to flexibly integrate information about observed gaze into ongoing attentional decisions. These “mirror” responses may play a role in translating social contingencies into species-typical behaviors such as *mob our attacker*. Such uniquely social affordances suggest a kind of nonrepresentational joint attention which coordinates joint action but is limited to species-typical social repertoires.

These mechanisms are likely the foundation upon which human joint attention evolved. While nonhuman animals can manipulate the attentional states of others, humans do so through specialized gestures used for no other purpose. While other animals appear to engage in joint attention, only humans have a flexible repertoire of cooperative behaviors facilitating informative and declarative pointing gestures. The development of the human pointing gesture appears intimately connected with representational joint attention abilities and likely played a pivotal role in the development of symbolic signaling including human speech.

### Acknowledgments

We thank Ipek Kulahci and Matthew Slayton for their helpful comments. SS is supported by a Princeton University Training Grant in Quantitative Neuroscience, NRSA T32 MH065214–1. MC is supported by a Postdoctoral Research Fellowship from Valente Center for Arts and Sciences, Bentley University—Waltham, Massachusetts.

### References

Akiyama, T., Kato, M., Muramatsu, T., Saito, E., Nakachi, R., Kashima, H., et al. (2006). A deficit in discriminating gaze direction in a case with right superior temporal gyrus lesion. *Neuropsychologia*, *44*, 161–170.

- Akiyama, T., Kato, M., Muramatsu, T., Saito, F., Umeda, S., Kashima, H., et al. (2006). Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. *Neuropsychologia*, *44*, 1804–1810.
- Akiyama, T., Kato, M., Muramatsu, T., Umeda, S., Saito, F., Kashima, H., et al. (2007). Unilateral amygdala lesions hamper attentional orienting triggered by gaze direction. *Cerebral Cortex*, *17*, 2593–2600.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.
- Allman, J. M. (1999). *Evolving brains*. New York: Freeman.
- Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience*, *118*, 1099–1120.
- American Psychiatric Association. (1994). *Autistic disorder*. Washington, DC: American Psychiatric Association.
- Amici, F., Aureli, F., Visalberghi, E., & Call, J. (2009). Spider monkeys (*Ateles geoffroyi*) and capuchin monkeys (*Cebus apella*) follow gaze around barriers: Evidence for perspective taking? *Journal of Comparative Psychology*, *123*, 368–374.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223–229.
- Bedny, M., Pascual-Leone, A., & Saxe, R. R. (2009). Growing up blind does not change the neural bases of Theory of Mind. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 11312–11317.
- Beets, I. A., Rösler, F., & Fiehler, K. (2010). Nonvisual motor learning improves visual motion perception: Evidence from violating the two-thirds power law. *Journal of Neurophysiology*, *104*, 1612–1624.
- Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*, 260–267.
- Borg, E. (2007). If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies*, *14*(8), 5–19.
- Borjon, J., Shepherd, S. V., Todorov, A., & Ghazanfar, A. A. (in press). Eye-gaze and arrow cues influence elementary sound perception. *Proceedings, Biological Sciences*.
- Bräuer, J., Call, J., & Tomasello, M. (2005). All great ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology*, *119*, 145–154.
- Bugnyar, T., Stöwe, M., & Heinrich, B. (2004). Ravens, *Corvus corax*, follow gaze direction of humans around obstacles. *Proceedings, Biological Sciences*, *271*, 1331–1336.

- Bullinger, A., Zimmermann, F., Kaminski, J., & Tomasello, M. (2011). Different social motives in the gestural communication of chimpanzees and human children. *Developmental Science*, *14*, 58–68.
- Burkart, J., & Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, *120*, 120–130.
- Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of minds* (pp. 223–232). Oxford: Blackwell.
- Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., et al. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology*, *17*, 20–25.
- Calder, A., Lawrence, A., Keane, J., Scott, S., Owen, A., Christoffels, I., et al. (2002). Reading the mind from eye gaze. *Neuropsychologia*, *40*, 1129–1138.
- Call, J., Hare, B. A., & Tomasello, M. (1998). Chimpanzee gaze following in an object-choice task. *Animal Cognition*, *1*, 89–99.
- Call, J., & Tomasello, M. (1994). The social learning of tool use by orangutans (*Pongo pygmaeus*). *Human Evolution*, *9*(4), 297–313.
- Campbell, J. (2005). Joint attention and common knowledge. In N. M. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds: Issues in philosophy and psychology* (pp. 205–229). Oxford: Clarendon Press.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1–143.
- Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass: Counter-mirror activation following incompatible sensorimotor learning. *European Journal of Neuroscience*, *28*, 1208–1215.
- Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., et al. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 17825–17830.
- Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world: Inside the mind of another species*. Chicago: University of Chicago Press.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*, 319–349.
- Corballis, M. C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, *26*, 199–208, discussion 208–260.

- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, 3, 201–215.
- Csibra, G. (2007). Action mirroring and action understanding: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Attention and performance* (Vol. 22, pp. 435–459) Oxford: Oxford University Press.
- Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York: Norton.
- Deaner, R. O., & Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, 13, 1609–1613.
- De Souza, W. C., Eifuku, S., Tamura, R., Nishijo, H., & Ono, T. (2005). Differential characteristics of face neuron responses within the anterior superior temporal sulcus of macaques. *Journal of Neurophysiology*, 94, 1252–1266.
- de Waal, F. B. (2003). Darwin's legacy and the study of primate visual communication. *Annals of the New York Academy of Sciences*, 1000(1), 7–31.
- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39(1), 39–45.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91, 176–180.
- Doherty, M. J., Anderson, J. R., & Howieson, L. (2009). The rapid development of explicit gaze judgment ability at 3 years. *Journal of Experimental Child Psychology*, 104, 296–312.
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6, 509–540.
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43, 1035–1045.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 9602.
- Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, 85, 199–212.
- Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15, 447–452.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, 308, 662–667.

- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*, 490–495.
- Gallese, V. (2005a). The intentional attunement hypothesis: The mirror neuron system and its role in interpersonal relations. In S. Wermter, G. Palm, & M. Elshaw (Eds.), *Biomimetic neural learning for intelligent robots: Intelligent systems, cognitive robotics, and neuroscience* (pp. 19–30). Berlin: Springer-Verlag.
- Gallese, V. (2005b). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, *4*, 23–48.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593–609.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance* (Vol. 19, pp. 247–266). Oxford: Oxford University Press.
- Gazzola, V., Van Der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., & Keysers, C. (2007). Aphasics born without hands mirror the goal of hand actions with their feet. *Current Biology*, *17*, 1235–1240.
- Ghashghaei, H. T., & Barbas, H. (2002). Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Science*, *115*, 1261–1279.
- Goossens, B., Dekleva, M., Reader, S., Sterck, E., & Bolhuis, J. (2008). Gaze following in monkeys is modulated by observed facial expressions. *Animal Behaviour*, *75*, 1673–1681.
- Gothard, K., Battaglia, F., Erickson, C., Spitler, K., & Amaral, D. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal of Neurophysiology*, *97*, 1671–1683.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, *66*, 377–388.
- Grosbras, M., Laird, A. R., & Paus, T. (2005). Cortical regions involved in eye movements, shifts of attention, and gaze perception. *Human Brain Mapping*, *25*(1), 140–154.
- Hafed, Z. M., & Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research*, *42*, 2533–2545.
- Hare, B., & Tomasello, M. (1999). Domestic dogs (*Canis familiaris*) use human and conspecific social cues to locate hidden food. *Journal of Comparative Psychology*, *113*, 173–177.
- Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*, 571–581.

- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science*, *298*, 1634–1636.
- Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, *32*, 203–218.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science*, *2*(3), 96–99.
- Hesler, N., & Fisher, J. (2007). Gestural communication in Barbary macaques (*Macaca sylvanus*): An overview. In J. Call & M. Tomasello (Eds.), *The gestural communication of apes and monkeys* (pp. 159–195). Mahwah, NJ: Erlbaum.
- Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*, 1229–1243.
- Hobson, R. P., & Bishop, M. (2003). The pathogenesis of autism: Insights from congenital blindness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*, 335–344.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80.
- Hoffman, K. L., Gothard, K. M., Schmid, M. C., & Logothetis, N. K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology*, *17*, 766–772.
- Horton, K. E., & Caldwell, C. A. (2006). Visual co-orientation and expectations about attentional orientation in pileated gibbons (*Hylobates pileatus*). *Behavioural Processes*, *72*, 65–73.
- Jacob, P. (2008). What do mirror neurons contribute to human social cognition? *Mind & Language*, *23*(2), 190–223.
- Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition*, *44*, 280–302.
- Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: Dissociating subsystems for facial-information processing. *Current Biology*, *16*, 2023–2029.
- Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews. Neuroscience*, *6*, 766–774.
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L., Feinberg, D. R., et al. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proceedings. Biological Sciences/The Royal Society*, *277*, 617–624.

- Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2005). Domestic goats, *Capra hircus*, follow gaze direction and use social cues in an object choice task. *Animal Behaviour*, *69*, 11–18.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *361*, 2109–2128.
- Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., et al. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain*, *122*, 779–783.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendrick, K. M., da Costa, A. P., Leigh, A. E., Hinton, M. R., & Peirce, J. W. (2001). Sheep don't forget a face. *Nature*, *414*(6860), 165–166.
- Kingstone, A., Friesen, C., & Gazzaniga, M. (2000). Reflexive joint attention depends on lateralized cortical connections. *Psychological Science*, *11*, 159–166.
- Kita, S. (2003). *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Erlbaum.
- Klein, J. T., Deaner, R. O., & Platt, M. L. (2008). Neural correlates of social target value in macaque parietal cortex. *Current Biology*, *18*, 419–424.
- Kuhn, G., & Kingstone, A. (2009). Look away! Eyes and arrows engage oculomotor responses automatically. *Attention, Perception & Psychophysics*, *71*, 314–327.
- Langton, S., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541–567.
- Leavens, D. A., & Hopkins, W. D. (1998). Intentional communication by chimpanzees: A cross-sectional study of the use of referential gestures. *Developmental Psychology*, *34*, 813–822.
- Leavens, D. A., Hopkins, W. D., & Bard, K. A. (2005). Understanding the point of chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in Psychological Science*, *14*(4), 185–189.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*, 297–307.
- Loretto, M., Schloegl, C., & Bugnyar, T. (2010). Northern bald ibises follow others' gaze into distant space but not behind barriers. *Biology Letters*, *6*(1), 14–17.
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology, Paris*, *102*(1–3), 59–70.

Meltzoff, A. N., & Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds* (pp. 335–366). Oxford: Oxford University Press.

Miklósi, A., & Soproni, K. (2006). A comparative analysis of animals' understanding of the human pointing gesture. *Animal Cognition*, *9*, 81–93.

Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., Csányi, V., et al. (2003). A simple reason for a big difference: Wolves do not look back at humans, but dogs do. *Current Biology*, *13*, 763–766.

Miklósi, A., Polgardi, R., Topál, J., & Csányi, V. (2000). Intentional behavior in dog–human communication: An experimental analysis of “showing” behaviour in the dog. *Animal Cognition*, *3*, 159–166.

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*, 1–9.

Moore, C., & Dunham, P. J. (1995). Joint attention: Its origins and role in development. Hillsdale, NJ: Lawrence Erlbaum.

Moore, T., & Fallah, M. (2004). Microstimulation of the frontal eye field and its effects on covert spatial attention. *Journal of Neurophysiology*, *91*, 152–162.

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating “unseen” fear. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 1680–1685.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., Parlade, M. V., et al. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*, 938–954.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., Rizzolatti, G., et al. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, *78*, 2226–2230.

Neiworth, J. J., Burman, M. A., Basile, B. M., & Lickteig, M. T. (2002). Use of experimenter-given cues in visual co-orienting and in an object-choice task by a new world monkey species, cotton top tamarins (*Saguinus oedipus*). *Journal of Comparative Psychology*, *116*, 3–11.

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*, 135–143.

Okada, T., Sato, W., Kubota, Y., Usui, K., Inoue, Y., Murai, T., et al. (2008). Involvement of medial temporal structures in reflexive attentional shift by gaze. *Social Cognitive and Affective Neuroscience*, *3*(1), 80–88.

Okamoto-Barth, S., Call, J., & Tomasello, M. (2007). Great apes' understanding of other individuals' line of sight. *Psychological Science*, *18*, 462–468.

Op De Beeck, H. P., Dicarlo, J. J., Goense, J. B., Grill-Spector, K., Papanastassiou, A., Tanifuji, M., et al. (2008). Fine-scale spatial organization of face and object selectivity in the temporal lobe: Do functional magnetic resonance imaging, optical imaging, and electrophysiology agree? *Journal of Neuroscience*, *28*, 11796–11801.

Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, *8*, 315–324.

Pack, A. A., & Herman, L. M. (2004). Bottlenosed dolphins (*Tursiops truncatus*) comprehend the referent of both static and dynamic human gazing and pointing in an object-choice task. *Journal of Comparative Psychology*, *118*, 160–171.

Paukner, A., Anderson, J. R., Fogassi, L., & Ferrari, P. F. (2007). Do facial gestures, visibility or speed of movement influence gaze following responses in pigtail macaques? *Primates*, *48*(3), 241–244.

Pelphrey, K. A., Singerman, J., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: The influence of context. *Neuropsychologia*, *41*, 156–170.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*, 23–30.

Perrett, D., Rolls, E., & Caan, 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., et al. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *223*, 293–317.

Pika, S., & Zuberbühler, K. (2008). Social games between bonobos and humans: Evidence for shared intentionality? *American Journal of Primatology*, *70*(3), 207–210.

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., et al. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology*, *101*, 2581–2600.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*, 233–238.

Platt, M. L., Lau, B., Glimcher, P. W., Hall, W. C., & Moschovakis, A. (2004). Situating the superior colliculus within the gaze control network. In W. C. Hall & A. Moschovakis (Eds.), *The superior colliculus: New approaches for studying sensorimotor integration* (pp. 1–34). Boca Raton: CRC Press.

Ricciardi, E., Bonino, D., Sani, L., Vecchi, T., Guazzelli, M., Haxby, J. V., et al. (2009). Do we really need vision? How blind people “see” the actions of others. *Journal of Neuroscience*, *29*, 9719–9724.

- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews. Neuroscience*, *11*, 264–274.
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Research*, *3*, 131–141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews. Neuroscience*, *2*, 661–670.
- Rosa, M. G. P., & Tweedale, R. (2005). Brain maps, great and small: lessons from comparative studies of primate visual cortical organization. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 665–691.
- Rosati, A. G., & Hare, B. (2009). Looking past the model species: Diversity in gaze-following skills across primates. *Current Opinion in Neurobiology*, *19*, 45–51.
- Saalmann, Y. B., & Kastner, S. (2009). Gain control in the visual thalamus during perception and cognition. *Current Opinion in Neurobiology*, *19*, 408–414.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*, 235–239.
- Schloegl, C., Kotrschal, K., & Bugnyar, T. (2007). Gaze following in common ravens, *Corvus corax*: Ontogeny and habituation. *Animal Behaviour*, *74*, 769–778.
- Seltzer, B., & Pandya, D. N. (1991). Post-rolandic cortical projections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology*, *312*, 625–640.
- Seltzer, B., & Pandya, D. (2004). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology*, *281*, 97–113.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology: CB*, *18*, 668–671.
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*, 127–134.
- Sewards, T. V., & Sewards, M. A. (2002). Innate visual object recognition in vertebrates: Some proposed pathways and mechanisms. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, *132*, 861–891.
- Shepherd, S. V. (2010). Following gaze: Gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, *4*(5), 1–13.
- Shepherd, S. V., & Platt, M. L. (2008). Spontaneous social orienting and gaze following in ringtailed lemurs (*Lemur catta*). *Animal Cognition*, *11*, 13–20.

- Shepherd, S. V., & Platt, M. L. (2009). Neuroethology of attention in primates. In L. Tommasi, M. A. Peterson, & L. Nadel (Eds.), *Cognitive biology: Evolutionary and developmental perspectives on mind, brain, and behavior (Vienna Series in Theoretical Biology)* (Vol. d, pp. 281–308). Cambridge, MA: MIT Press.
- Shepherd, S. V., Deaner, R. O., & Platt, M. L. (2006). Social status gates social attention in monkeys. *Current Biology*, *16*, R119–R120.
- Shepherd, S. V., Klein, J. T., Deaner, R. O., & Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 9489–9494.
- Shepherd, S. V., Steckenfinger, S. A., Hasson, U., & Ghazanfar, A. A. (2010). Human–monkey gaze correlations reveal convergent and divergent patterns of movie viewing. *Current Biology*, *20*, 649–656.
- Schiffer, S. (1972). *Meaning*. Oxford: Clarendon Press.
- Sperber, D., & Wilson, D. (1986). *Relevance, communication and cognition*. Cambridge, MA: Harvard University Press.
- Tamietto, M., Castelli, L., Vighetti, S., Perozzo, P., Geminiani, G., Weiskrantz, L., et al. (2009). Unseen facial and bodily expressions trigger fast emotional reactions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 17661–17666.
- Tanner, J. E., & Byrne, R. W. (2010). Triadic and collaborative play by gorillas in social games with objects. *Animal Cognition*, *13*, 591–607.
- Teufel, C., Gutmann, A., Pirow, R., & Fischer, J. (2010). Facial expressions modulate the ontogenetic trajectory of gaze-following among monkeys. *Developmental Science*, *13*, 913–922.
- Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*, 314–318.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (Vol. 16, pp. 103–130). Hillsdale, NJ: Erlbaum.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, *68*, 1067–1080.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*, 1063–1069.

- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–691.
- Tomasello, M., Hare, B., & Agnetta, B. (1999). Chimpanzees, Pan troglodytes, follow gaze direction geometrically. *Animal Behaviour*, *58*, 769–777.
- Tomasello, M., Hare, B., & Fogleman, T. (2001). The ontogeny of gaze following in chimpanzees, *Pan troglodytes*, and rhesus macaques, *Macaca mulatta*. *Animal Behaviour*, *61*, 335–343.
- Tootell, R. B. H., Tsao, D., & Vanduffel, W. (2003). Neuroimaging weighs in: humans meet macaques in “primate” visual cortex. *Journal of Neuroscience*, *23*, 3981–3989.
- Tsao, D. Y., Schweers, N., Moeller, S., & Freiwald, W. A. (2008). Patches of face-selective cortex in the macaque frontal lobe. *Nature Neuroscience*, *11*, 877–879.
- Udell, M. A., Dorey, N. R., & Wynne, C. D. (2010). What did domestication do to dogs? A new account of dogs’ sensitivity to human actions. *Biological Reviews of the Cambridge Philosophical Society*, *85*, 327–345.
- Umiltà, M. A., Escola, L., Intskirveli, I., Rochat, M., Caruana, F., Jezzini, A., et al. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 2209–2213.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., et al. (2001). I know what you are doing. A neurophysiological study. *Neuron*, *31*(1), 155–165.
- Vecera, S. P., & Rizzo, M. (2004). What are you looking at? Impaired “social attention” following frontal-lobe damage. *Neuropsychologia*, *42*, 1657–1665.
- Vuilleumier, P. (2002a). Facial expression and selective attention. *Current Opinion in Psychiatry*, *15*, 291–300.
- Vuilleumier, P. (2002b). Perceived gaze direction in faces and spatial attention: A study in patients with parietal damage and unilateral neglect. *Neuropsychologia*, *40*, 1013–1026.
- Wang, Y., Brzozowska-Prechtel, A., & Karten, H. J. (2010). Laminar and columnar auditory cortex in avian brain. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 12676–12681.
- Zimmermann, F., Zemke, F., Call, J., & Gomez, J. C. (2009). Orangutans (*Pongo pygmaeus*) and bonobos (*Pan paniscus*) point to inform a human about the location of a tool. *Animal Cognition*, *12*, 347–358.