

## 8 Engaging Neocortical Networks with Dynamic Faces

Stephen V. Shepherd and Asif A. Ghazanfar

Dynamism is *the* rule that governs our pursuit of goals and patterns the environment we perceive. To understand perception of faces, we must do more than imagine a series of discrete, static, stimuli and responses. Brains exist to coordinate behavior, not merely to perceive; they have evolved to guide interaction with their environment and not merely to produce a representation of it. Brains expect our environment to change, both in response to our actions and of its own accord, and they extrapolate both the predictable motion of physical objects and the more capricious, goal-directed motion of other animate beings. Finally, our brains have adapted to a world that tests us continuously in real time, not discretely, in separable questions and answers.

After all, primates are animate; our faces translate and rotate in space relative to our bodies and to our larger environment. Furthermore, onto the static, bone-structure-defined configuration of any individual's face are layered dynamic processes, including (from longer to shorter time scales) aging processes; changes in health, hormonal status, hunger, hydration, and exertion; blinks and expression shifts occurring irregularly every several seconds; irregular but sustained mouth and postural changes associated with vocalizations; microexpression shifts several times per second; shifts in sensory orientation several times per second (most obviously by the eyes, but also ears in some primates); and irregular but rhythmic patterns associated with mastication and human speech. Each of these steadily changing features alters the perception of the observer and contributes to both basic perception (e.g., biological motion) and cognitive attributions (e.g., mental states, social relationships). Earlier study of these dynamic features has focused on cortical responses to static gaze direction, static expression, and pictures or brief videos depicting biological motion. Although some of the consequences of observing facial dynamics are known—including a tendency to physically mimic both observed actions and attentional or emotional states—the details and mechanisms of these processes remain largely obscure.

This dynamism presents incredible technical difficulties, but recent advances make it possible to generate and analyze increasingly complex stimuli and, in parallel, to record from large-scale brain networks and analyze their relationships with their external environment. Thus we believe that despite the myriad challenges, we are at the verge of crossing another Rubicon beyond which social processes will not be conceived merely as stimulus behaviors and perceptual responses. It seems likely that the most consequential features of natural social behavior—including vocal communication, such as speech—cannot be successfully evoked using static, unimodal, non-interacting, spatially abstracted stimuli. Nonetheless, the transition to an interactive paradigm of social interaction will necessarily be bridged by using data collected with static faces. As such, this chapter will review some of what we know about face processing in dyadic interactions with humans and nonhuman primates. When possible, we will reveal how dynamic faces change what we know about face processing and its underlying neural substrates.

#### **Facial Motion and Vocal Communication**

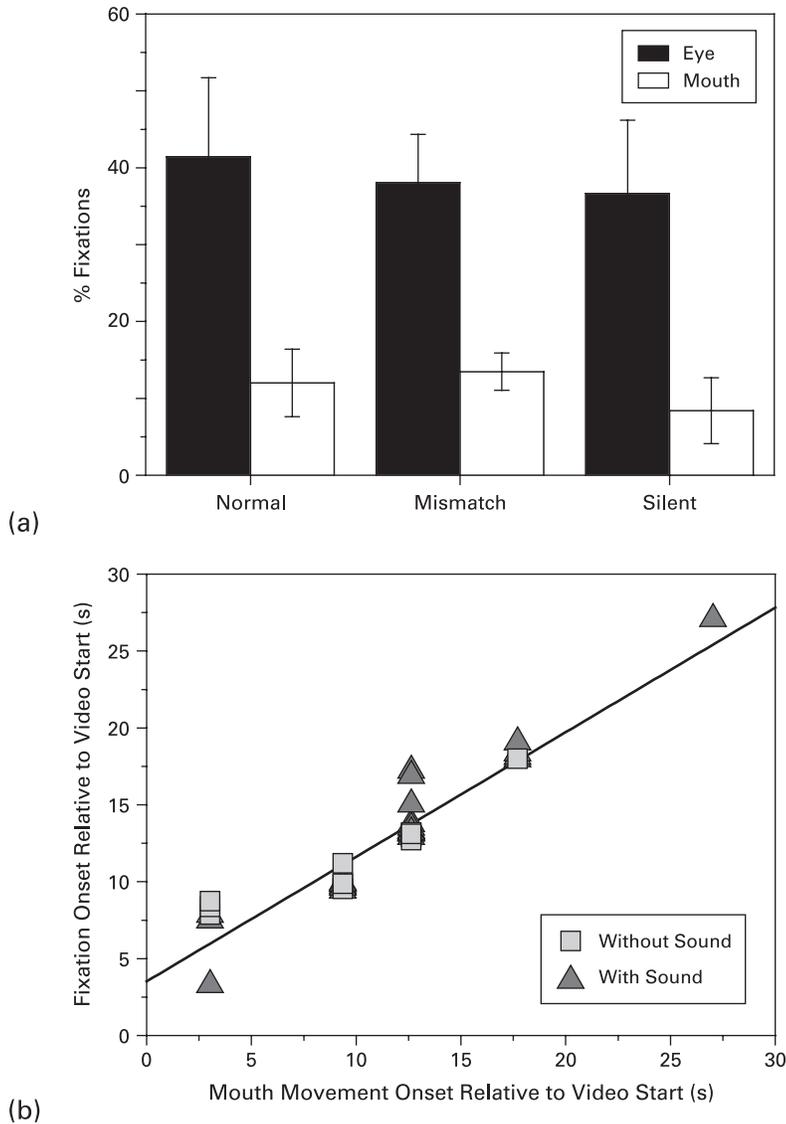
Primates spend much of their time looking at the faces of other individuals and in particular at their eyes. If the eye-movement strategies of monkeys viewing vocalizing faces are made primarily to glean social information, then why fixate on the eyes? Many previous experiments have shown that monkeys prefer to look at the eyes when viewing neutral or expressive faces (Keating & Keating, 1982; Nahm, Perret, Amaral, & Albright, 1997; Guo, Robertson, Mahmoodi, Tadmor, & Young, 2003) and the attention directed at the eyes often seems to be used to assess the intention of a conspecific or other competitor (Ghazanfar & Santos, 2004). Humans likewise tend to focus on the eye region (Yarbus, 1967; Birmingham, Bischof, & Kingstone, 2007; Fletcher-Watson, Findlay, Leekam, & Benson, 2008). Thus both humans and monkeys may focus on the eyes when observing a conspecific to glean information about that individual's intentions. (This is reviewed in detail in the following section.)

During vocal communication, patterns of eye movements of the observer can be driven by both the demands of the task as well as the dynamics of the observed face. Recent studies of humans examined observers' eye movements during passive viewing of movies (Klin, Jones, Schultz, Volkmar, & Cohen, 2002) and under different listening conditions, such as varying levels of background noise (Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998), competing voices (Rudmann, McCarley, & Kramer, 2003), or silence (i.e., speech-reading with no audio track) (Lansing & McConkie, 1999, 2003). When typical human subjects are given no task or instruction regarding what acoustic cues to attend to, they will consistently look at the eye region more than the mouth when viewing videos of human speakers (Klin et al.,

2002). However, when subjects are required to perform a specific task, then eye-movement patterns are task dependent (see Land & Hayhoe, 2001). For example, when they are required to attend to speech-specific aspects of a communication signal (e.g., phonetic details in high background noise, word identification, or segmental cues), humans will make significantly more fixations on the mouth region than on the eye region (Vatikiotis-Bateson et al., 1998; Lansing & McConkie, 2003). In contrast, when subjects are asked to focus on prosodic cues or to make social judgments based on what they see or hear, they direct their gaze more often toward the eyes than the mouth (Lansing & McConkie, 1999; Buchan, Pare, & Munhall, 2004, 2007). The sensorimotor mechanisms that analyze and integrate facial and vocal expressions are most likely an early innovation that is not specific to perception of human speech (Ghazanfar & Santos, 2004).

The eye-movement patterns of rhesus monkeys viewing dynamic vocalizing faces share many of the same features as human eye-movement patterns (Ghazanfar, Nielsen, & Logothetis, 2006). Monkeys viewing video sequences of other monkeys vocalizing under different listening conditions (silent, matched, or mismatched) spent most of their time inspecting the eye region relative to the mouth under all conditions (figure 8.1a). When they did fixate on the mouth, it was highly correlated with the onset of mouth movements (figure 8.1b). These data show that the pattern of eye fixations is driven at least in part by the dynamics of the face and are strikingly similar to what we know about human eye-movement patterns during speech-reading. In both species, a greater number of fixations fall in the eye region than in the mouth region when subjects are required simply to view conspecifics (Klin et al., 2002), to attend vocal emotion cues, or to make social judgments (Buchan et al., 2007). Even during visual speech alone (no auditory component), when subjects are asked to attend to prosodic cues, they will look at the eyes more than the mouth (Lansing & McConkie, 1999). Furthermore, like human observers (Lansing & McConkie, 2003), monkeys look at the eyes *before* they look at the mouth and their fixations on the mouth are tightly correlated with mouth movement (Ghazanfar et al., 2006). For instance, Lansing and McConkie (2003) reported that regardless of whether it was visual or audiovisual speech, subjects asked to identify words increased their fixations on the mouth region with the onset of facial motion. The same was true for rhesus monkeys; they fixate on the mouth upon the onset of movement in that region (Ghazanfar et al., 2006).

The dynamics of the face can give us clues as to why both monkeys and humans focus primarily on the eye region even during audiovisual vocal communication. One possibility is that gaze deployments may be optimized to extract socially relevant cues that are relatively higher in spatial frequency near the eyes and temporal frequency near the mouth, matching the relative spatiotemporal precision of the foveal and peripheral retina. In addition, the angular size of faces may be too small at

**Figure 8.1**

Eye movements of monkey observers viewing vocalizing conspecifics. (a). The average fixation on the eye region versus the mouth region across three subjects while viewing a 30-second video of a vocalizing conspecific. The audio track had no influence on the proportion of fixations falling onto the mouth or the eye region. Error bars represent the S.E.M. (b). We also find that when monkeys do saccade to the mouth region, it is tightly correlated with the onset of mouth movements ( $r = 0.997$ ,  $p < 0.00001$ ).

conversational distances for there to be a large cost to speech-reading in fixating on the eyes rather than the mouth.

As proposed by Vatikiotis-Bateson et al. (1998), it is possible that perceivers acquire vocalization-related information that is distributed broadly on the vocalizer's face. Facial motion during speech is in part a direct consequence of the vocal tract movements necessary to shape the acoustics of speech; indeed, a large portion of the variance observed in vocal tract motion can be estimated from facial motion (Yehia, Kuratate, & Vatikiotis-Bateson, 2002). Humans, therefore, can identify vocal sounds when the mouth appears outside the fovea or is masked, presumably by using these larger-scale facial motion cues (Preminger, Lin, & Levitt, 1998). Head movement can also be an informative cue, linked to both the fundamental frequency ( $F_0$ ) and the voice amplitude of the speech signal (Yehia et al., 2002; Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004). When head movements are eliminated or distorted in speech displays, speech perception is degraded (Munhall et al., 2004). Finally, it is possible that saccades to the mouth are epiphenomenal and merely a reflexive response to detection of motion in the visual periphery (Vatikiotis-Bateson et al., 1998).

As in humans, different rhesus monkey vocalizations are produced with unique facial expressions and the motion of articulators influences the acoustics of the signal (Hauser, Evans, & Marler, 1993; Hauser and Ybarra, 1994). Such articulatory postures could influence facial motion beyond the mouth region. For example, grimaces produced during scream vocalizations cause the skin folds around the eyes to increase in number (Hauser, 1993). Thus, for many of the same reasons suggested for human perceivers, rhesus monkeys may simply not need to look directly at a mouth to monitor visual aspects of vocalization.

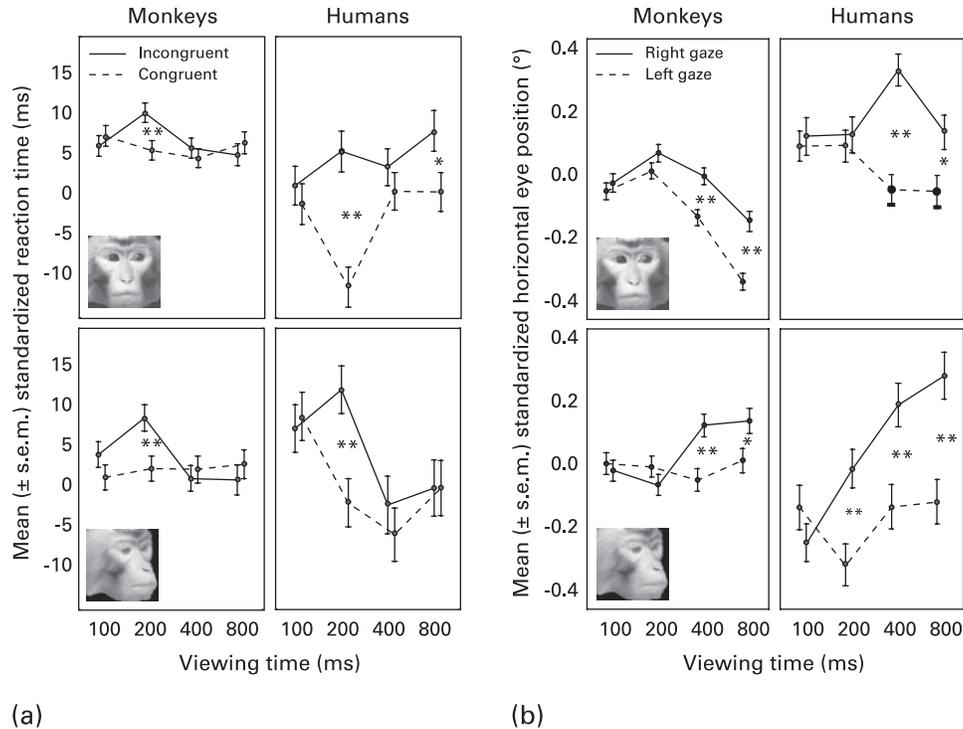
### Gaze and Attention

Faces are salient because they are readily identifiable indicators of animacy and because they strongly distinguish species and individuals. More than this, faces typically lead the body in movement; they contain the ingestive apparatus by which competitors consume and predators prey; and they combine the major sensory organs that both define an animal's attentional orientation and connote its current intent. The orientation of our most important sensory organ—our eyes—is particularly revealing, providing a crucial context for facial expressions and revealing both our attentional state and our likely future intentions. The information encoded in the direction of a gaze appears to consist of two related signals: first, and urgently, whether the observed individual gazes toward the observer; second, whether (and which) object or area has captured the attention of the observed individual.

It is quite likely that the first manner of sensitivity to gaze direction—sensitivity to being watched—occurs early in ontogeny and is phylogenetically widespread. Human infants prefer full-face to nonface configurations within 72 hours (Macchi Cassia, Simion, & Umiltà, 2001), and prefer direct to averted gaze shortly thereafter (within 2–5 days; Farroni, Csibra, Simion, & Johnson, 2002), by gestation age 10 months (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000). This sensitivity is widespread, being shared with diverse vertebrates, reported variously in fish (Coss, 1979), primitive primates (Coss, 1978), marine mammals (Xitco, Gory, & Kuczaj, 2004), lizards (Burger, Gochfeld, & Murray, 1992), snakes (Burghardt & Greene, 1988), and birds (Ristau, 1991).

The second manner of sensitivity—the use of gaze as a deictic (pointing) cue—has a more uncertain origin both in the brain and in the course of evolution. It is known, for example, that gaze perception interacts bidirectionally with emotional and social perception, both in the perception of facial expressions (Adams & Kleck, 2005; Ganel, Goshen-Gottstein, & Goodale, 2005) and in environmental interactions (Bayliss & Tipper, 2006; Bayliss, Frischen, Fenske, & Tipper, 2007). Furthermore, perceiving another's direction of gaze does not merely inform us but actively steers our own attention. Fletcher-Watson and colleagues found that while the first saccades to a social scene were to the eyes (Fletcher-Watson et al., 2008), subsequent saccades tended to follow gaze. Increasingly, evidence suggests that these responses are not fully under conscious control. Friesen and Kingstone (1998) discovered, and others quickly replicated (Driver et al., 1999; Langton & Bruce, 1999), that humans reflexively follow the gaze of others. We respond faster and more accurately when detecting, localizing, or discriminating targets that appear in the direction viewed by another individual. These responses occur to static and dynamic gaze cues, to head and eye orientation, to real or cartoon faces (reviewed by Frischen, Bayliss, & Tipper, 2007), and persist even when gaze cues oppose the demands of a task four times out of five.

However important such a response may be to human's sophisticated social cognition (Baron-Cohen, 1994), it appears it is not unique. Apes (Brauer, Call, & Tomasello, 2005), monkeys (Emery, Lorincz, Perrett, Oram, & Baker, 1997; Tomasello, Call, & Hare, 1998), and perhaps even lemurs (Shepherd & Platt, 2008) follow gaze. Intriguingly, similarities between macaque and human gaze-following share similar dynamics (figure 8.2) (Deaner & Platt, 2003), suggesting that our behaviors may share common, evolutionarily ancient mechanisms. In monkeys, as in humans, these fast and stereotyped responses are nonetheless nuanced and context dependent [e.g., influenced in humans by emotion (Mathews, Fox, Yiend, & Calder, 2003; Putman, Hermans, & van Honk, 2006), gender and familiarity (Deaner, Shepherd, & Platt, 2007); and in monkeys by dominance status (Shepherd, Deaner, & Platt, 2006)]. The mechanisms that mediate and modulate gaze-following may not, however, be unique

**Figure 8.2**

Gaze-following tendencies are shared by humans and macaques, and their similar time course suggests they may derive from similar mechanisms. (a). Savings in reaction time when responding to congruent versus incongruent stimuli. At 200 ms after cue onset, both macaque and human subjects responded faster to targets that appeared in the direction of their gaze. (b). Bias in fixation position as a function of observed gaze direction. From 200 ms onward, a systematic drift in eye position suggested that evoked shifts of covert attention had biased the subject's microsaccades in the same direction toward which they'd seen another primate look. \* $p = 0.05$ ; \*\* $p = 0.001$ .

to perceived attention: instead, they may represent just one aspect of a more general tendency, alternately termed “mimicry,” “mirroring,” or “embodied perception.”

### Activating Motor Responses through Dynamic Faces

Faces do not reflect the attention and intention of other animals by accident. In primates, faces have become a primary vehicle for the active communication of visuo-social signals, a trend that has reached its apogee in the naked faces of *Homo sapiens*. Anthropoid primates, in whom visual signaling has supplanted olfactory cues, devote large swaths of the cortex to the production of facial expressions (Allman, 1999). Researchers have identified at least five categories of human facial expression that

appear to be intelligible across cultures (Ekman, 1993). Although these emotions can be readily identified in static images, the canonical photographs sometimes appear caricatured to observers, seeming unnatural when they are divorced from a dynamic context. Indeed, some facial expressions—such as embarrassment—cannot be fully captured except in dynamic sequence: dropping gaze, smiling (often suppressed), then turning away or touching the face (Keltner, 1995).

As with gaze, the responses evoked by dynamic social stimuli are not merely perceptual. William James (James, 1890) articulated the ideomotor theory of action, in which every mental representation evokes the represented behavior—evoked traces that are not mere concomitants of perception of emotion but necessary intermediaries in perceptual experience. Since that time, ample support has arisen that emotional facial expressions (e.g., Hess & Blairy, 2001) induce mimicry and are thus in some sense contagious. This mimicry is automatic and does not require conscious awareness (Dimberg, Thunberg, & Elmehed, 2000). Furthermore, interference with motor mimicry can disrupt recognition of facial expression (Oberman, Winkielman, & Ramachandran, 2007).

Again, however, this seemingly reflexive process nonetheless presents subtle context dependence. Mimicry may not be tied to any specific physical effector, suggesting that we mirror emotions, rather than purely physical states (Magnee, Stekelenburg, Kemner, & de Gelder, 2007), and facial expressions may induce complementarity rather than mimicry when social dominance is at stake (Tiedens & Fragale, 2003). Crucially, this mimicry does not just aid in perception, it may actively regulate social interaction. Mimicry and affiliation mutually reinforce one another, with mimicry promoting affiliation (Chartrand & Bargh, 1999; Wiltermuth & Heath, 2009, although note van Baaren, Holland, Kawakami, & van Knippenberg, 2004) and affiliation likewise enhancing mimicry (Lakin & Chartrand, 2003; Likowski, Muehlberger, Seibt, Pauli, & Weyers, 2008). While presentation of static faces can sometimes evoke these responses, dynamic faces are much more effective. For example, muscle contraction is much more evident when viewing videotaped than static expressions in both the corrugator supercilii (contracted in anger) and the zygomatic major (contracted in joy) (Sato, Fujimura, & Suzuki, 2008).

The attentional and emotional contagion produced by observing a face may be part of a more general trend toward mirroring as a perceptual process. Mimicry appears to occur at multiple levels of abstraction, reflecting action goals, overall motor strategies, and specific effectors and movements. Furthermore, mimicry is not only triggered by gaze and perception of expression, it also plays a role in the perception of speech; mechanical manipulation of listeners' faces biases their auditory perception toward congruently mouthed words (Ito, Tiede, & Ostry, 2009). Although the data may not require that perceived actions and emotions be embodied by the observer before they can be comprehended [individuals with facial paralysis can

nonetheless recognize emotion (Calder, Keane, Cole, Campbell, & Young, 2000)], they do suggest that physical or simulated embodiment contributes to normal face perception, empathy, communication, and social affiliation.

#### **Neural Networks Activated by Dynamic versus Static Faces**

It is likely that social pressure is a major factor in the evolution of larger brains (Barton and Dunbar, 1997; Reader and Laland, 2002). Consistent with this idea, we note that many—perhaps the majority—of cortical areas can be activated by faces under some conditions and may thus be considered part of an extended face perception network. For example, in addition to core face perception areas (reviewed by Tsao and Livingstone, 2008), cortical areas involved in attention (Haxby, Hoffman, & Gobbini, 2000; Calder et al., 2007) and somatosensation (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000) are active during observation of faces. Moreover, a range of other areas involved in emotional, mnemonic, and goal-directed processes are commonly engaged by socially significant faces (e.g., Vuilleumier, Armony, Driver, & Dolan, 2001; Ishai, Schmidt, & Boesiger, 2005).

Most of what we know about neural processing of faces comes from research on passive perception or simple categorization of static, cropped faces. Dynamic faces appear to be much more effective than static ones at activating neural tissues throughout the extended face perception network (Fox, Iaria, & Barton, 2008), and we speculate that socially relevant, interactive faces would be more effective still. Although current models suggest that facial transients are processed through a separate stream from permanent features such as identity (Haxby et al., 2000 but see also Calder & Young, 2005), dynamism can be expected to enrich the information flowing through both streams. Dynamic stimuli make explicit which facial features are transitory and which are permanent; furthermore, patterns of movement are often idiosyncratic and provide information about underlying structure. Conversely, because features such as bone structure and musculature constrain facial movements, and because recognition of a person likewise constrains the posterior probabilities of underlying mental states, identity-related computations are likely to modulate the processing of more dynamic facial features. For further discussion of identity processing and facial dynamics, see chapter 4.

The dynamic visual features described here—including orofacial movement, gaze shifts, and emotional expressions—are thought to be analyzed primarily by cortical areas located near the superior temporal sulcus (Allison, Puce, & McCarthy, 2000). The detailed pathways by which the brain mediates social perceptions and behavioral responses is incompletely understood. Although faces do not remain long in one position, they do dynamically orbit a resting state. When muscles are relaxed, all other things being equal, the mouth gently closes, the eyes align with the head and the head

with the body, and the facial expression becomes placidly neutral. Static photographs of any other pose can thus be considered to imply a temporary shift from this hypothetical default. Perhaps for this reason, relatively little data explicitly contrast cortical responses and dynamic and static stimuli.

#### **Gazed Direction**

Regions of the human STS (Puce, Allison, Bentin, Gore, & McCarthy, 1998; Wicker, Michel, Henaff, & Decety, 1998) and amygdala (Kawashima et al., 1999) have both been implicated in visual processing of observed gaze. In particular, the posterior STS is reciprocally interconnected with posterior parietal attention areas, and both regions are activated when subjects specifically attend gaze direction (Hoffman & Haxby, 2000). In monkeys, 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 neurons in the caudal STS respond symmetrically to gaze averted to either the right or the left, anterior neurons respond differentially to different gaze directions (De Souza, Eifuku, Tamura, Nishijo, & Ono, 2005). Although human imaging studies have most consistently reported activations in the caudal STS, recent studies have suggested that only neuronal activity in the inferior parietal lobule and the anterior STS represent gaze deaxis, that is, the specific direction toward which gaze is directed (Calder et al., 2007). Nonetheless, a fast subcortical pathway that routes visual information to the amygdala may also play a role in social attention shifts (Adolphs et al., 2005), potentially including those cued by observed gaze (Akiyama et al., 2007).

#### **Emotion and Mimicry**

Emotional expressions are likewise believed to be processed initially in the STS but have been reported to involve a broad array of areas, including the amygdala; parietal somatosensory regions; insula; and frontal motor, reward, and executive function centers. These activations in the extended face perception system often vary from task to task and emotion to emotion (reviewed by Adolphs, 2002; Vuilleumier & Pourtois, 2007). The significance of an emotional signal and the neural activity it evokes is no doubt profoundly affected by dynamic social and nonsocial environmental cues. For example, it is crucial whether your ally became angry at your enemy, or at you. To understand brain responses to dynamic faces, we must understand how contextual and motivational effects interact with these perceptions. In short, we must extend our paradigm to reflect the interactive and contextually nuanced nature of social relationships.

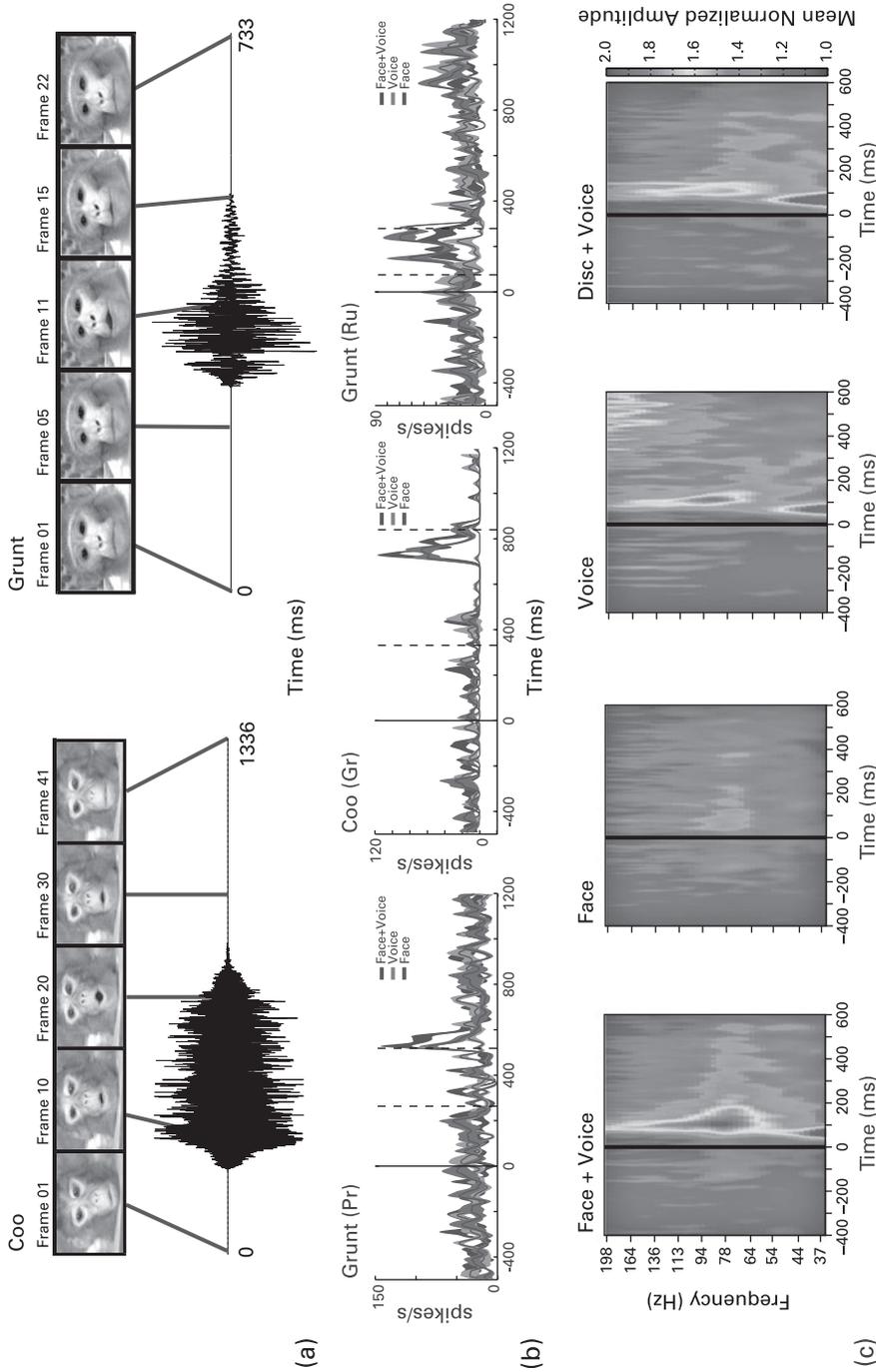
The role of mimicry in face perception is echoed by findings that somatosensory and motor cortices are involved in processing dynamic faces. In particular, parietal somatosensory and frontoparietal motor areas have been found to contain individual

neurons that respond both when performing an action and when observing, either visually or auditorily, the same action being performed by others (reviewed by Rizzolatti & Craighero, 2004). Furthermore, the same neural tissues that generate motor activity are activated during passive perception. For example, transcranial magnetic stimulation (TMS) pulses evoke activity at a lower threshold when subjects view congruent motor activations (Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Strafella & Paus, 2000); and the observation of dynamic facial emotions activates motor tissues governing congruent facial displays (Sato, Kochiyama, Yoshikawa, Naito, & Matsuura, 2004). Such mirror activations may mediate observed mimicry effects. In the lateral intraparietal areas that govern attention, for example, observation of averted gaze in another evokes neuronal activity that tracks the dynamics of gaze-following (Shepherd, Klein, Deane, & Platt, 2009). These data suggest that independent of whether the observed stimuli are physically embodied, the tissues that produce and detect an individual's own behavioral state are likewise involved in perceiving the behavioral state of another.

#### **Vocal Communication**

Beyond the STS and inferior regions of the temporal lobe, dynamic faces influence how voices are processed in the auditory cortex (Ghazanfar, Maier, Hoffman, & Logothetis, 2005; Ghazanfar, Chandrasekaran, & Logothetis, 2008). The vast majority of neural responses in this region show integrative (enhanced or suppressed) responses when dynamic faces are presented with voices than with unimodal presentations (figure 8.3 a and b). Furthermore, these dynamic face and voice responses were specific: replacing the dynamic faces with dynamic disks that mimicked the aperture and displacement of the mouth did not lead to integration (Ghazanfar et al., 2005, 2008). This parallels findings from human psychophysical experiments in which such artificial dynamic disk stimuli can lead to enhanced speech detection but not to the same degree as a dynamic face (Bernstein, Auer, & Takayanagi, 2004; Schwartz, Berthommier, & Savariaux, 2004).

Although there are multiple possible sources of visual input to the auditory cortex (Ghazanfar & Schroeder, 2006), the STS is likely to be the major region through which facial images influence the auditory cortex. This is first, because there are reciprocal connections between the STS and the lateral belt and other parts of the auditory cortex (described earlier; see also Barnes & Pandya, 1992; Seltzer & Pandya, 1994; second, because neurons in the STS are sensitive to both faces and biological motion (Harries & Perrett, 1991; Oram & Perrett, 1994); and finally, because the STS is known to be multisensory (Bruce, Desimone, & Gross, 1981; Schroeder & Foxe, 2002). One mechanism for establishing whether the auditory cortex and the STS interact at the functional level is to measure their temporal correlations as a function of stimulus condition. Concurrent recordings of oscillations and single



**Figure 8.3** Dynamic faces engage the auditory cortex and its interactions with the superior temporal sulcus. (a) Example of a coo call and a grunt call from rhesus monkeys. The top panel shows frames at five intervals from the start of the video (the onset of mouth movement) until the end of mouth movement. The x-axes depict time in milliseconds. The bottom panel displays the time waveform of the vocalization where the blue lines indicate the temporally corresponding video frames. (b) Examples of multisensory integration in auditory cortex neurons. Peristimulus time histograms and rasters in response to a grunt vocalization (left and right panels) and coo vocalization (middle panel) according to the face+voice (F+V), voice alone (V), and face alone (F) conditions. The x-axes show time aligned to onset of the face (solid line). Dashed lines indicate the onset and offset of the voice signal. The y-axes depict the firing rate of the neuron in spikes per second. (c) Time-frequency plots (cross-spectrograms) illustrate the modulation of functional interactions (as a function of stimulus condition) between the auditory cortex and the STS for a population of cortical sites. The x-axes depict the time in milliseconds as a function of onset of the auditory signal (solid black line). The y-axes depict the frequency of the oscillations in hertz. The color bar indicates the amplitude of these signals normalized by the baseline mean.

neurons from the auditory cortex and the dorsal bank of the STS reveal that gamma band correlations significantly increased in strength during presentation of bimodal face and voice videos compared with unimodal conditions (Ghazanfar et al., 2008) (figure 8.3c). Because the phase relationships between oscillations were significantly less variable (tighter) when dynamic faces were paired with voices, these correlations are not merely due to an increase in response strength but also reflect a tighter temporal coordination between the auditory cortex and the STS (Varela et al., 2001). This relationship is elaborated further in the following chapter by Puce and Schroeder.

### **To See Past the Mask**

For humans, the most important part of face perception is the insight it gives into the underlying person. The significance of human faces in our environment is that they belong to people—friends, enemies, coworkers, competitors, prospects, spouses, and relatives—about whom we maintain richly detailed histories. These individuals form the cast among whom we give our aspirations play. Diverse neural tissues help guide our interaction with these characters, and thus mnemonic, emotional, and executive areas are included in the extended network activated by observation of dynamic faces.

Although we did not review these extended processes in detail here, we note that dynamic faces most likely potentiate these processes, just as they potentiate the lower-level perceptions we have dwelled upon. Medial temporal areas recalling diverse personal associations (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005) are most likely more strongly activated as dynamic videos make explicit the temporal invariants within a face, as well as characteristic or otherwise memorable expression trajectories. Likewise, the circuits involved in perspective taking, mentalistic attribution, empathy, and goal processing can be cued by dynamic faces, and the evoked processing seems to resemble that which we outlined earlier. Again, tissues that are specialized for self-processing (e.g., of perceptual sets and working memory in the right temporoparietal junction and inferior frontal gyrus, and of goal sets and cognitive and emotional load in the medial prefrontal cortex) seem to be recruited to aid in the comprehension of others (Saxe, 2006; Frith & Frith, 2007). Thus, dysfunction in these systems can have dire consequences for our ability to relate to others (see de Gelder and Van den Stock in this volume). Both the cortical and the conceptual breadth of activity evoked by dynamic faces speak to their primacy in our lives; our humanity hinges on our relationships with other people.

In nature, faces do not appear and disappear as still images, unannounced, unmoving, and unaccompanied. Nor did our brains evolve merely to perceive passive scenes or to respond to the interrogations of curious researchers. The stage across

which we struggle to survive and multiply is not composed of static scenes, and the players are not mere passive props: we are *interactors*.

## References

- Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotions. *Emotion, 5*, 3–11.
- Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1*, 21–61.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *Journal of Neuroscience, 20*, 2683–2690.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*, 68–72.
- Akiyama, T., Kato, M., Muramatsu, T., Umeda, S., Saito, F., & Kashima, H. (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 Science, 4*, 267–278.
- Allman, J. M. (1999). *Evolving brains*. New York: W.H. Freeman.
- Barnes, C. L., & Pandya, D. N. (1992). Efferent cortical connections of multimodal cortex of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology, 318*, 222–244.
- Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Current Psychology of Cognition, 13*, 513–552.
- Barton, R. A., & Dunbar, R. L. M. (1997). Evolution of the social brain. In A. Whiten and R. Byrne (eds.), *Machiavellian intelligence*. Cambridge, UK: Cambridge University Press.
- 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*, 223–229.
- Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition, 104*, 644–653.
- Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science, 17*, 514–520.
- Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*, 5–18.
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61*, 986–998.
- Brauer, 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.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*, 369–384.
- Buchan, J. N., Pare, M., & Munhall, K. G. (2004). The influence of task on gaze during audiovisual speech perception. *Journal of the Acoustical Society of America, 115*, 2607.
- Buchan, J. N., Pare, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2*, 1–13.
- Burger, J., Gochfeld, M., & Murray, B. (1992). Risk discrimination of eye contact and directness of approach in black iguanas (*Ctenosaura similis*). *Journal of Comparative Psychology, 106*, 97–101.
- Burghardt, G., & Greene, H. (1988). Predator simulation and duration of death feigning in neonate hog-nose snakes. *Animal Behavior, 36*, 1842–1844.
- Calder, A. J., Beaver, J. D., Winston, J. S., Dolan, R. J., Jenkins, R., Eger, E., & Henson, R. N. (2007). Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Current Biology, 17*, 20–25.

- Calder, A. J., Keane, J., Cole, J., Campbell, R., & Young, A. W. (2000). Facial expression recognition by people with Mobius syndrome. *Cognitive Neuropsychology*, *17*, 73–87.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641–651.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.
- Coss, R. (1978). Perceptual determinants of gaze aversion by the lesser mouse lemur (*Microcebus murinus*). The role of two facing eyes. *Behavior*, *64*, 248–267.
- Coss, R. (1979). Delayed plasticity of an instinct: Recognition and avoidance of 2 facing eyes by the jewel fish. *Developmental Psychobiology*, *12*, 335–345.
- 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.
- Deaner, R. O., & Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, *13*, 1609–1613.
- Deaner, R. O., Shepherd, S. V., & Platt, M. L. (2007). Familiarity accentuates gaze cuing in women but not men. *Biology Letters*, *3*, 64–67.
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*, 86–89.
- 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.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.
- Emery, N. J., Lorincz, E. N., Perrett, D. I., Oram, M. W., & Baker, C. I. (1997). Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology*, *111*, 286–293.
- Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 9602–9605.
- Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*, 571–583.
- Fox, C. J., Iaria, G., & Barton, J. J. (2008). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, *30*, 1637–1651.
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review*, *5*, 490–495.
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*, 694–724.
- Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*, *17*, R724–732.
- Ganel, T., Goshen-Gottstein, Y., & Goodale, M. A. (2005). Interactions between the processing of gaze direction and facial expression. *Vision Research*, *45*, 1191–1200.
- Ghazanfar, A. A., Chandrasekaran, C., & Logothetis, N. K. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *Journal of Neuroscience*, *28*, 4457–4469.
- Ghazanfar, A. A., Nielsen, K., & Logothetis, N. K. (2006). Eye movements of monkeys viewing vocalizing conspecifics. *Cognition*, *101*, 515–529.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience*, *25*, 5004–5012.
- Ghazanfar, A. A., & Santos, L. R. (2004). Primate brains in the wild: The sensory bases for social interactions. *Nature Reviews Neuroscience*, *5*, 603–616.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*, 278–285.

- Guo, K., Robertson, R. G., Mahmoodi, S., Tadmor, Y., & Young, M. P. (2003). How do monkeys view faces? A study of eye movements. *Experimental Brain Research*, *150*, 363–374.
- Harries, M. H., & Perrett, D. I. (1991). Visual processing of faces in temporal cortex. Physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience*, *3*, 9–24.
- Hauser, M. D. (1993). Right-hemisphere dominance for the production of facial expression in monkeys. *Science*, *261*, 475–477.
- Hauser, M. D., Evans, C. S., & Marler, P. (1993). The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations. *Animal Behavior*, *45*, 423–433.
- Hauser, M. D., & Ybarra, M. S. (1994). The role of lip configuration in monkey vocalizations—Experiments using xylocaine as a nerve block. *Brain and Language*, *46*, 232–244.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, *4*, 223–233.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*, 129–141.
- 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*, 80–84.
- Ishai, A., Schmidt, C. F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin*, *67*, 87–93.
- Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *Proceedings of the National Academy of Sciences U.S.A.*, *106*, 1245–1248.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt.
- Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., Fukuda, H., Kojima, S., & Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain*, *122*, 779–783.
- Keating, C. F., & Keating, E. G. (1982). Visual scan patterns of rhesus monkeys viewing faces. *Perception*, *11*, 211–219.
- Keltner, D. (1995). Signs of appeasement—Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology*, *68*, 441–454.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*, 809–816.
- Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*, 334–339.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565.
- Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*, 541–567.
- Lansing, C. R., & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech Language and Hearing Research*, *42*, 526–539.
- Lansing, I. R., & McConkie, G. W. (2003). Word identification and eye fixation locations in visual and visual-plus-auditory presentations of spoken sentences. *Perception and Psychophysics*, *65*, 536–552.
- Likowski, K. U., Muehlberger, A., Seibt, B., Pauli, P., & Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, *44*, 1065–1072.
- Macchi Cassia, V., Simion, F., & Umiltà, C. (2001). Face preference at birth: The role of an orienting mechanism. *Developmental Science*, *4*, 101–108.
- Magnee, M. J., Stekelenburg, J. J., Kemner, C., & de Gelder, B. (2007). Similar facial electromyographic responses to faces, voices, and body expressions. *Neuroreport*, *18*, 369–372.
- Mathews, A., Fox, E., Yiend, J., & Calder, A. (2003). The face of fear: Effects of eye gaze and emotion on visual attention. *Visual Cognition*, *10*, 823–835.

- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility—Head movement improves auditory speech perception. *Psychological Science, 15*, 133–137.
- Nahm, F. K. D., Perret, A., Amaral, D. G., & Albright, T. D. (1997). How do monkeys look at faces? *Journal of Cognitive Neuroscience, 9*, 611–623.
- Oberman, L. M., Winkelman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience, 2*, 167–178.
- Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (Stpa) neurons to biological motion stimuli. *Journal of Cognitive Neuroscience, 6*, 99–116.
- 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.
- Preminger, J. E., Lin, H.-B., Payen, M., & Levitt, H. (1998). Selective visual masking in speechreading. *Journal of Speech, Language, and Hearing Research, 41*, 564–575.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18*, 2188–2199.
- Putman, P., Hermans, E., & van Honk, J. (2006). Anxiety meets fear in perception of dynamic expressive gaze. *Emotion, 6*, 94–102.
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*, 1102–1107.
- Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 4436–4441.
- Ristau, C. (1991). Attention, purposes, and deception in birds. In A. Whiten (ed.), *Natural theories of mind*. Oxford, UK: Oxford University Press, pp. 209–222.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.
- Rudmann, D. S., McCarley, J. S., & Kramer, A. F. (2003). Bimodal displays improve speech comprehension in environments with multiple speakers. *Human Factors, 45*, 329–336.
- Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *International Journal of Psychophysiology, 70*, 70–74.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Brain Research Cognitive Brain Research, 20*, 81–91.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology, 16*, 235–239.
- Schroeder, C. E., & Foxe, J. J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Cognitive Brain Research, 14*, 187–198.
- Schwartz, J.-L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: Evidence for early audiovisual interactions in speech identification. *Cognition, 93*, B69–B78.
- Seltzer, B., & Pandya, D. N. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *Journal of Comparative Neurology, 343*, 445–463.
- Shepherd, S. V., Deaner, R. O., & Platt, M. L. (2006). Social status gates social attention in monkeys. *Current Biology, 16*, R119–120.
- 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 U.S.A., 106*, 9489–9494.
- Shepherd, S. V., & Platt, M. L. (2008). Spontaneous social orienting and gaze following in ringtailed lemurs (*Lemur catta*). *Animal Cognition, 11*, 13–20.
- Strafella, A. P., & Paus, T. (2000). Modulation of cortical excitability during action observation: A transcranial magnetic stimulation study. *Neuroreport, 11*, 2289–2292.

- Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive non-verbal behavior. *Journal of Personality and Social Psychology, 84*, 558–568.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behavior, 55*, 1063–1069.
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31*, 411–437.
- van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*, 71–74.
- Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*, 229–239.
- Vatikiotis-Bateson, E., Eigsti, I. M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception and Psychophysics, 60*, 926–940.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30*, 829–841.
- Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia, 45*, 174–194.
- Wicker, B., Michel, F., Henaff, M. A., & Decety, J. (1998). Brain regions involved in the perception of gaze: A PET study. *Neuroimage, 8*, 221–227.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1–5.
- Xitco, M. J., Jr., Gory, J. D., & Kuczaj, S. A., 2nd (2004). Dolphin pointing is linked to the attentional behavior of a receiver. *Animal Cognition, 7*, 231–238.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.
- Yehia, H. C., Kuratate, T., & Vatikiotis-Bateson, E. (2002). Linking facial animation, head motion and speech acoustics. *Journal of Phonetics, 30*, 555–568.