Social perception

- Risky decisions
- Evolution of speech
- Body self-calibration
Social perception refers to initial stages in the processing of information that culminates in the accurate analysis of the dispositions and intentions of other individuals. Single-cell recordings in monkeys, and neurophysiological and neuroimaging studies in humans, reveal that cerebral cortex in and near the superior temporal sulcus (STS) region is an important component of this perceptual system. In monkeys and humans, the STS region is activated by movements of the eyes, mouth, hands and body, suggesting that it is involved in analysis of biological motion. However, it is also activated by static images of the face and body, suggesting that it is sensitive to implied motion and more generally to stimuli that signal the actions of another individual. Subsequent analysis of socially relevant stimuli is carried out in the amygdala and orbitofrontal cortex, which supports a three-structure model proposed by Brothers. The homology of human and monkey areas involved in social perception, and the functional interrelationships between the STS region and the ventral face area, are unresolved issues.

Consider the predicament of Barbara Ehrenreich, who is considering a vacation out West:

It would be nice to go on a vacation where I didn’t have to worry about being ripped limb from limb by some big, urinous sloth...All right, I know the ecologically correct line: “They won’t bother you if you don’t bother them.” But who knows what bothers a bear?...So instead of communing with the majestic peaks and flower-studded meadows, I spend my hikes going over all the helpful tips for surviving an Encounter. Look them in the eye? No, that was mountain lions. Bears just hate it when you stare at them, so keep your gaze fixed dreamily on the scenery. Play dead? Let’s see, that works for grizzlies but not for black bears. So do you take off the backpack, get out the wildlife guidebook, do a quick taxonomic determination and then play dead?

If it is difficult to infer the intentions of other humans from their facial gestures and body language, it is even harder,
In monkeys, gaze direction is an important component of facial expressions, particularly those related to dominance and submission. Perrett and colleagues have studied the responsiveness of monkey STS cells to gaze and head direction. The results summarized in Fig. 1 are representative. In general, cells that were most responsive to the full face preferred eye contact, and cells that were tuned to the profile view of a face preferred averted gaze, but some cells showed independent sensitivity to head and eye direction.

Such cells appear to have a role in ‘social attention’, that is, cells that signal the direction of another individual’s attention. A dramatic example is illustrated in Fig. 2. A young woman faces the viewer (a), but her gaze is directed to her right. The reason for her averted gaze becomes clear when we view more of the scene (b) and see that she is carefully but surreptitiously attending to the dupe on her right and stealing his jewelry while his attention is captured by the young woman’s confederate, a fortune teller. Such deceptive
behavior is not limited to humans. A young gorilla has been observed hugging and looking into the eyes of a human while surreptitiously stealing his watch. Three neuroimaging studies in humans have examined the activity evoked by viewing eye movements or direction of gaze. In a PET study, subjects viewed videos of actors looking towards the subject (mutual gaze condition) or looking away (averted gaze condition). Compared with a condition in which the actor looked down such that the eyes appeared to be closed (no-gaze condition), several regions of activation were found, including parts of the middle temporal gyri and the STS itself, but the relatively poor spatial resolution of PET and the consequent use of across-subject averaging to detect activation might have obscured possible activity within the STS.

---

**Fig. 2.** An example of social attention. (a) Why are this woman’s eyes averted? (b) Because she is paying careful attention to the direction of gaze of the man on her right, while he is paying attention to the old woman to their left. The man’s gullibility and misdirected attention will cost him dearly. (Reproduced, with permission, from Georges de La Tour’s ‘The Fortune-Teller’, ca. 1630.)

---

**Fig. 3.** Activation of the superior temporal sulcus (STS) region in the left hemisphere (a) and right hemisphere (b) during the perception of biological motion. All activations are to moving stimuli except for three studies of implied motion. Activations are shown in the coordinate system of Talairach and Tournoux. There is variability in the location and configuration of sulci and gyri between individuals, hence the centers of activation (circles) are only approximately correct in relation to brain structures. The center of activation during the perception of American Sign Language (ASL) by expert deaf signers are correct but misleading, as the entire STS region is activated bilaterally. In some cases activations in different conditions or experiments of a study have been combined. Some centers have been moved slightly to allow visualization of overlapping centers. Activations of other brain regions are not shown. Abbreviations: AG, angular gyrus; STG, superior temporal gyrus. (Centers of activation are taken from Refs 17, 21, 22, 43, 48–50, 52, 56, 57, 59–62, 92.)
case, a region of cortex near and perhaps partly within the STS was activated during perception of gaze. We will use the term ‘STS region’ to refer to cortex within the STS, to adjacent cortex on the surface of the superior and middle temporal gyri (near the straight segment of the STS), and to adjacent cortex on the surface of the angular gyrus (near the ascending limb of the STS). Similar results have been obtained in a functional MRI (fMRI) study in which subjects viewed a face in which the eyes averted to the left or right.21 Alternations of eye aversion and eyes looking at the observer activated portions of the STS region (Fig. 3). These regions were anterior to area MT/V5, as determined by the activation produced by nonbiological motion in the same parts of the visual field. Another fMRI study found that the perception of eye gaze in static facial images activated similar portions of the STS region.22

In scalp recordings,23 event-related potentials (ERPs) were recorded while subjects viewed the same stimuli used in the fMRI study described above.21 An N170 ERP, previously shown to be responsive to face and eye stimuli,24–26 was larger in response to eye aversion than to eyes returning to gaze at the observer. This effect occurred whether the eyes were viewed in isolation or in the context of a full face. The ERP effects evoked by eye movement were not produced by movement per se. The exact regions of cortex that generate N170 cannot be determined from scalp recordings, but the results suggest that regions of posterior temporal cortex are part of a system sensitive to eye movement and direction of gaze.

The results reviewed above suggest that the STS region has a role in the perception of gaze. If so, a lesion of the region should impair judgements of gaze direction, as indeed occurs in monkeys following lesions restricted to the STS (Refs 27,28). Three prosopagnosic individuals also exhibited deficits in the perception of gaze direction,27,29, but the location and extent of the abnormality is unknown in a developmental prosopagnosic (patient A.B.) and might involve, but is not limited to, the STS region in the other two (patients R.B. and K.D.). It will be important to study the perception of gaze direction in individuals with lesions restricted mainly to the STS region and who are not prosopagnosic, as it is likely that perception of gaze direction and perception of facial identity can be dissociated.22

Of the behavioral deficits seen in autistic children,30–32 two are in the use of gaze direction and the comprehension of mental states. Baron-Cohen and colleagues asked whether these two abnormalities might be related.33 Two of the experiments they carried out are summarized in Fig. 4. When normal or mentally retarded children were shown pictures like the one illustrated in Fig. 4a, and asked which candy ‘Charlie’ prefers, they typically pointed to the Polo Mints. By contrast, autistic children were significantly less likely to point to the Polo Mints. This deficit was not due simply to an inability to perceive the direction of gaze; autistic children scored as well as normal or mentally retarded children when shown faces like those illustrated in Fig. 4b and asked, ‘Which one is looking at you?’ In other words, they were able to perceive the direction of gaze, but were unable to use such information to infer the mental state of another person. Anatomical abnormalities have been described in the temporal lobes of autistic individuals,34 but no studies have yet specifically implicated the STS region (for a review see Ref. 34).

Head movement

Head movements convey several types of information.11 Rolls and colleagues studied the responsiveness of STS cells as monkeys viewed human head movements.35 The responsiveness of a cell located in the lower bank of the STS (Fig. 5d) is particularly interesting in this context. This cell fired vigorously during ventral flexion of the head (movement towards the chest), but responded poorly during dorsal flexion of the head (movement away from the chest) or to a static face (Fig. 5a). The responsiveness of the cell to head movements performed in different orientations is shown in Fig. 5b. The cell responded strongly to ventral flexion whether the head was viewed full face, inverted, in profile or from behind. The response of the cell to the inverted head is informative, because in this case the direction of movement on the retina is the opposite of that to

![Fig. 4. A direction-of-gaze experiment in normal and autistic children. (a) Example of a display. When asked which candy ‘Charlie’ prefers, most normal children point to the Polo Mints, but autistic children are less likely to do so. (b) Example of schematic faces. When asked ‘which one is looking at you?’, autistic children score as well as normal children. (Reproduced, with permission, from Ref. 33.)](image-url)
Indeed, it is likely that any mouth movement that is related and non-speech movements. First we consider the Mouth movement

Mouth movements can be broadly divided into speech-related and non-speech movements. We consider the latter category of movement.

In monkeys, mouth movements are an important component of facial gestures. For example, mouth opening and teeth baring are components of threat or fear displays, whereas 'smiling' denotes submission or a positive emotional state. Some cells in monkey STS respond specifically to mouth opening and teeth baring are components of threat or fear displays, whereas 'smiling' denotes submission or a positive emotional state. Some cells in monkey STS respond specifically to mouth opening or to lowering of the eyebrows. Note that the cell was not responsive to the eyes or to direction of gaze. It fired only weakly in response to downward or upward movements of the eyes alone, and recall from Fig. 5c that it fired strongly when the eyes were not present, for example, when the back of the head moved ventrally. Rolls et al. point out that this combination of responses is of interest because downward eye movements and ventral flexion of the head often appear together as part of the behavioral response of breaking contact with another monkey during dominance interactions.

The neuronal activity evoked in the human temporal lobe by moving heads has not been studied. However, ERPs evoked by static faces and heads have been recorded from the posterior STS region using methods described in Box 1. Static images evoked face-specific N200 ERPs that were smaller in response to profile views of the head than to three-quarter or full-face views, but whether the difference was related to mechanisms involved in face recognition or to analysis of implied head movement is unknown.

In humans, intracranial ERPs were recorded from the temporal lobe to determine whether non-linguistic mouth movement activated cells in the temporal lobe analogous to the activation found in monkeys. An example of the results is shown in Box 1. Long-latency ERPs were generated in or near the STS in response to mouth opening. Similar activity was recorded from the STS region in six additional individuals; the centers of activation are summarized in Fig. 3. An fMRI study of the activation produced by mouth movement used stimuli similar to those shown in Box 1 (Ref. 21). Computation of the activation produced by motion of the radial circles, compared with a baseline condition with no movement, allowed the authors to determine the brain regions activated by movement per se. Such activation was centered in area MT/V5. By contrast, mouth movement activated a more-anterior
Box 1. Electrophysiological recordings from the human temporal lobe

As part of their evaluation for possible surgery, individuals with medically intractable epilepsy might have electrodes implanted on the surface of the brain, or within the brain, to determine the locus of seizure onset (Refs a,b). During the period of electrode implantation, which usually lasts from 7 to 14 days, continuous EEG and behavioral video monitoring is carried out. ERPs can also be recorded without interfering with clinical monitoring.

A 19-year-old female had partial complex seizures. Scalp EEG recordings and behavioral features of her seizures suggested a seizure onset in the left frontotemporal region. An 8 × 8 grid of electrodes was placed over the region (Fig. Ia,b). In addition, depth probes were inserted into the brain (Fig. Ic,d) to monitor EEG activity in the hippocampus and in other medial temporal lobe structures that are often epileptogenic.

Four days after electrode implantation she viewed static and moving images of faces, face parts and control stimuli presented on a computer screen at 1.5–2.2 s intervals (Fig. IIa). ERPs were recorded simultaneously from 64 electrodes. ERPs at the locations shown in Figs IIb–d are shown in Figs IIb–d, and demonstrate that focal ERPs specific to mouth movement were recorded from sites within the STS and from adjacent sites on the cortical surface near the STS. These and similar recordings allow the conclusion that regions of the STS respond preferentially to mouth movement.

References


Fig. 6. Sensitivity of a cell in the superior temporal sulcus to facial gestures. Illustration of different facial gestures by a Rhesus monkey (left). Recordings of spike activity of the cell during viewing of these gestures (right). (Reproduced, with permission, from Ref. 11.)
The subject viewed a computer monitor while sitting upright in his hospital bed. The images were 60 cm from the subject’s face and subtended 10.7° × 10.7° of visual angle. Eyes were directed at the viewer and the mouth was closed (1), eyes were averted to the right (2), eyes were averted to the left (3), or the mouth opened (4). Rapid transition from the image at left to the other image yields the strong illusion of eye or mouth movement. (b) Event-related potential (ERP) recordings from the surface electrodes shown in Fig. 1b. Movement onset was 100 ms after the beginning of the traces. Electrodes 2 and 3, adjacent to the superior temporal sulcus (STS), recorded a positive ERP to mouth movement but not to eye movement. (c) Electrodes 2 and 3 adjacent to the superior temporal sulcus (STS) region distinct from those activated by non-speech mouth movements. Polarity inversion at nearby electrodes 4–6, within the floor of the STS, recorded a negative ERP to mouth movement but not to eye movement. (d) Recordings from surface electrode 3 and depth electrode 6 are juxtaposed to illustrate the polarity inversion of ERPs evoked by mouth movement. Polarity inversion at nearby electrodes provides strong evidence that the activity is locally generated in cortex of the STS or adjacent surface cortex. (Adapted, with permission, from Ref. d.)
centrally been reviewed. Congenitally blind children gesture when they speak, a topic that has received much attention, mainly in the left hemisphere. Recent studies have shown that viewing meaningful hand movements activates the STS region (Fig. 3). Thus, the responsiveness of these cells suggests that they encode goal-directed hand movement.

Activations produced by similar types of hand movement have been studied in humans using PET. Rizzolatti and colleagues found that observation of objects being grasped by an investigator activated a portion of the left STS region (Fig. 3). In a control condition, subjects viewed static scenes of the investigator holding the same objects. Subtracting the activation produced by the control condition from that of the grasping condition thus removed activation that was due to object recognition, suggesting that activation of the STS region was due to analysis of meaningful hand movement. Graf von der Schulenburg and colleagues, using the same stimuli as Rizzolatti et al., confirmed activation of the left STS region during observation of grasping movements, although in this study the center of activation was anterior to the center reported by Rizzolatti et al. (Fig. 3). Similarly, Grezes and colleagues studied the activation produced when subjects viewed meaningful hand actions (e.g., opening a bottle) compared with a control condition in which the hand was stationary. Bilateral activation of the STS region was observed (Fig. 3).

Johansson has developed a useful technique in which isolated points of light are attached to relevant body parts and viewed in darkness (see Ref. 5). Walking and other biological movements are readily perceived, in the absence of form cues and even though static point-light displays are meaningless. Bonds and colleagues used a point-light display of the act of reaching towards a glass, picking it up, and bringing it to the mouth. Viewing this motion activated a posterior portion of the left STS region (Fig. 3). Thus, viewing meaningful hand movements activates the STS region, mainly in the left hemisphere.

Humans from all cultural and linguistic backgrounds produce hand gestures when they speak, a topic that has recently been reviewed. Congenitally blind children gesture as much when they speak as do sighted children, even when they know that the person they are speaking to is also blind and thus unable to profit from information conveyed by their gestures. These results suggest that gesture is an integral part of the speaking process itself. Indeed, Corbalis makes a plausible argument that spoken language might have evolved from hand gestures. He concludes that the spontaneous emergence of sign language among deaf communities confirms that gestural communication is as natural to humans as is spoken language. If so, one would expect that speech-related hand gestures would activate the same region of the STS activated by speech and by lip-reading. This prediction has been confirmed in an fMRI study by Neville and colleagues. They studied the activation evoked in congenitally deaf subjects whose native language was American Sign Language (ASL), and found that the STS region of both hemispheres was activated during ASL sentence processing, compared with a baseline condition in which the subjects viewed nonsense gestures that were physically similar to ASL signs (Fig. 3). In some parts of the STS region the right hemisphere was more activated than was the left hemisphere, but the hemispheric differences were not statistically significant. These results demonstrate that the STS region is involved in analysis of ASL. Such activation, particularly in the right hemisphere, might result from analysis of this type of biological motion; in the left hemisphere the activation could equally well be regarded as reflecting linguistic processing that is independent of the mode of transmission.

In monkeys, responsiveness of STS cells was greater to a hand making a movement than to a bar of the same size making the same movement, demonstrating that the cells are preferentially responsive to biological motion. In humans, hearing subjects who did not know ASL showed no activation of the STS region in either hemisphere while viewing ASL sentences. Bonds et al. found no activation of the STS region when subjects viewed random motion of point-light displays. These results suggest that the STS region of monkeys and humans is activated primarily by communicative or meaningful hand gestures, but not by meaningless hand movements. However, Grezes et al. have reported activation of the STS region when subjects viewed meaningful or meaningless hand movements.

Body movement

If movements of body parts activate the STS region, it stands to reason that movement of the whole body would do so as well. Several studies have investigated the responsiveness of cells in the monkey STS to moving bodies. For example, a population of cells in the STS responded selectively to views of a human walking in different directions.

In humans three PET or fMRI studies have examined the activation produced by point-light displays of humans performing movements such as walking, dancing and throwing. In these studies, a control condition consisted of the same lights moving randomly at the same velocity. Compared with the control condition, body movements activated the posterior STS region or a more-anterior part of the superior temporal gyrus that probably included part of the STS region (Fig. 3). Thus, parts of the STS region are involved in the perception of complex body movements.

Implied motion

The woman depicted by Georges de La Tour is looking to her right (Fig. 2a). Because this is an atypical view of a face we infer that she has just averted her gaze. Recent studies suggest that implied biological motion is also analyzed in part in the STS region. Kourtzi and Kanwisher tested subjects while they viewed static images of implied motion (for example, a discus thrower photographed at rest). They found stronger fMRI activation of the STS during viewing of implied motion than when viewing images without implied motion (Fig. 3). Senior and colleagues tested subjects with MRI while they viewed actual movements of hands and objects, and while they viewed static images that implied...
similar movements. Implied motion primarily activated area MT/V5, but in the right hemisphere a part of the STS region was also activated (Fig. 3).

The concept of implied motion allows us to reconcile the results of some of the studies reviewed above. The stimuli used in the monkey single-cell recordings (e.g., Figs 1,5,6) were static images. Furthermore, the human STS region is activated by static views of eyes, mouths, hands and faces. Thus, static views of bodies or body parts might activate the STS region when they imply motion. However, many of these stimuli do not imply motion in any obvious way; it will be important to investigate by behavioral and imaging studies the properties of stimuli that do and do not imply motion.

It is also possible that actual and implied motion are only two of several categories of cues used by the STS region in social perception. In an fMRI study, subjects viewed static ‘theory-of-mind’ and ‘non-theory-of-mind’ cartoons and stories. Some of the theory-of-mind stimuli implied motion while others did not; theory-of-mind stimuli evoked activation of parts of the STS region. The authors suggested that ‘...this region...is sensitive not merely to biological motion but, more generally, to stimuli which signal intentions or intentional activity’.

Brain regions important for social perception

For the purposes of this article the term social perception is used to include early stages of the analysis of actual or implied bodily movements and related cues that provide socially relevant information. Social perception is thus conceptualized as part of a larger domain of cognitive skills referred to as theory of mind, metacognition, social attention, and social cognition, which are defined as the ‘processing of information which culminates in the accurate perception of the dispositions and intentions of other individuals’.

The role of the STS region in social perception has been emphasized. Parts of the monkey STS receive input both from the ventral object recognition system (the ’what’ system) and from the dorsal spatial location-movement system (the ’where’ system), suggesting that this region integrates information about form and movement. Other brain regions are also involved in social perception, particularly in more broadly defined social cognition. The amygdala and orbitofrontal cortex (OFC), to which the STS region projects, are important components of the system (Fig. 7). The role of these structures in social cognition has recently been reviewed and will not be discussed except to note that: (1) the amygdala contains cells responsive to complex body movements; (2) the amygdala is activated during gaze monitoring; (3) perception of gaze direction is impaired following amygdalotomy; and (4) the role of the OFC in analysis of biological motion has not been studied, but it is activated by faces as determined by single-cell recordings in monkeys. ERP recordings in humans and PET studies in humans faces can convey information important for social reinforcement, a suggested role of the OFC (Ref. 79). The three-part system summarized in Fig. 7 has been proposed as the basis of social cognition in monkeys and the ‘mindreading system’ in humans.

The manner in which these structures interact to guide social behavior is unknown, but previous studies suggest useful working hypotheses. As illustrated in Fig. 7, the STS not only sends feedforward projections to the amygdala but receives feedback projections from it. In anatomically earlier regions of visual cortex, which have been more intensively studied, the anatomy of feedback projections and ERP recordings in monkeys and humans suggest that feedback projections to superficial cortical layers act to enhance the responses of pyramidal cells to feedforward sensory input. It is plausible to assume that a similar mechanism enhances the responses of STS cells. The amygdala attaches emotional salience to sensory input. In the encounter imagined by Barbara Ehrenreich, for example, cells that are responsive to perceived eye contact would be additionally activated by amygdalar feedback, based on cues that the bear is not only making eye contact but is growling and running rapidly towards her. This additional neuronal activity occurs several hundred milliseconds after initial activation of STS cells, can discriminate between different emotional expressions and might be due to feedback from the amygdala. Such feedback might thus induce ‘attentional amplification’ of STS activity evoked by salient social stimuli. At higher levels of social perception and cognition,
Facial gestures in monkeys, as illustrated, for example, in Fig. 6, are vocalizations can be used to infer the intentions of other individuals. What cues does the visual system use to categorize movement as ascending limb of the STS. (Adapted, with permission, from Ref. 64.)

This coronal slice is approximately at the junction of the straight segment and the ascending limb of the STS. (Adapted, with permission, from Ref. 64.)

Fusiform gyrus and the lateral activation (apparent in images. The ventral activation is in the fusiform gyrus and the lateral activation (apparent in this slice only in the right hemisphere, left side of image) is in the superior temporal sulcus (STS). This coronal slice is approximately at the junction of the straight segment and the ascending limb of the STS. (Adapted, with permission, from Ref. 64.)

Specific vocalization of STS cells responsive to faces, biological motion and other social stimuli, it seems likely that social stimuli are also represented by a sparse population code; this possibility has not as yet been tested.

Another region of visual cortex also needs to be considered. It is now well established that part of the ventral occipitotemporal cortex is involved in face perception. It is composed of cortex in the fusiform gyrus and adjacent inferior temporal and occipital gyri, and has been referred to as the ‘fusiform face area’ or the ‘ventral face area’. Both the ventral face area and the STS region are activated by faces (Fig. 8; see also Ref. 87). The existence of two anatomically separate temporal areas involved in face perception raises three issues.

First, the anatomical organization of human face-related cortex (Fig. 8) differs from monkey face-related cortex, which is a single continuous region extending from the upper bank of the STS to the ventral part of inferior temporal (IT) cortex (Fig. 5d; see also Ref. 15). Before the studies summarized in Fig. 3 were carried out, it was proposed that human face-related cortex ‘migrated’ as a unit from its location in monkeys, with the occipitotemporal sulcus (OTS) being analogous to the STS and the fusiform gyrus being analogous to IT cortex. This hypothesis is probably wrong because none of the studies summarized in Fig. 3 reported activation of the OTS by biological motion. Instead the human STS region apparently retains the functionality of the monkey STS. The region between the lateral and ventral face areas (Fig. 8) might be involved in the perception of non-face objects.

Second, the ventral face area is important for face recognition because, among other reasons, prosopagnosia is produced by ventral lesions but not by lesions restricted to the STS region. It is tempting to suppose that the ventral face area analyzes permanent facial features, whereas the STS region analyzes moment-to-moment changes in facial configuration. This is probably an oversimplification because: (1) the ventral face area is sometimes responsive to moving eyes and mouths; and (2) the perception of eye gaze preferentially activates the STS region and the perception of face identity preferentially activates the ventral face area, but both regions are activated by both types of stimuli. The role of the ventral face area in social perception remains to be investigated.

Third, as this article has demonstrated, many studies of the neural substrates of social perception in humans have been guided by results obtained in monkeys. However, direct comparison of human and monkey results is difficult for several reasons: (1) the anatomical methods used to differentiate areas of monkey cortex are not available in humans; hence little is known about homologies between monkey and human temporal cortex; (2) some of the STS region lies within language-related cortex (for a review see Ref. 83) that probably has no useful analog in monkey cortex; and (3) inferences about human function are based mainly on single-cell recordings. However, the recent development of fMRI techniques for use in monkeys (including a demonstration of activation in the STS by faces) will allow systematic study of social perception in humans and monkeys using the same stimuli and analytical techniques.

**Outstanding questions**

- What cues does the visual system use to categorize movement as biological or nonbiological? Is all movement analyzed by area MT/V5 (an area that myelinates early in development) in infants, with a later segregation of analysis by MT/V5 (nonbiological motion) and the STS region (biological motion)? Alternatively, are there sites within the STS region that are specialized for some types of nonbiological motion, and sites within MT/V5 that are specialized for some types of biological motion, as some imaging data suggest?

- Facial gestures in monkeys, as illustrated, for example, in Fig. 6, are often accompanied by specific vocalizations. Parts of the STS receive both visual and auditory input. Do some STS cells respond best to a face gesture and its accompanying vocalization, but not (or not as well) to either stimulus in isolation? The brain activations produced by such combined stimuli have not as yet been studied either by single-cell recordings in monkeys or by ERP or imaging studies in humans.

- Vocalizations can be used to infer the intentions of other individuals. What are the neural substrates of the auditory social perception system? Part of this putative system might reside in the STS region, which is activated more by vocalizations than by non-vocal sounds, suggesting that voice-selective parts of the STS region are the auditory counterpart of face-selective regions of visual cortex. Are there olfactory and somatosensory social perception systems, and if so, are they instantiated partly in the STS?
Conclusions
The actions and intentions of other individuals are important enough that brain mechanisms have evolved to provide rapid and accurate assessment of the actual and implied motions of other individuals. The neural substrates of these mechanisms have been studied in monkeys for two decades, but only in the past five years have neuroimaging and electrophysiological techniques been used to study the neural basis of social perception in humans. These studies converge to suggest that initial analysis of social cues occurs in the STS region, which is anatomically well situated to integrate information derived from both the ventral ‘what’ and the dorsal ‘where’ visual pathways. The STS region is large and functionally complex. A challenge for the future will be to integrate human and monkey studies better, in order to provide a clearer understanding of the locations and functions of the neural network involved in social perception, including downstream structures such as the amygdala and orbitofrontal cortex, which are also involved in social perception and cognition.

Acknowledgements
Preparation of this review was supported by the Veterans Administration and by NIH Grant NH-02868. We thank J. Jolierski and M. Jensen for assistance, and Drs D.O. and S.S. Spencer, and the staff of the Yale Epilepsy Surgery Program for their support of the ERP recordings discussed. We also thank three anonymous referees for their helpful comments on earlier versions of this article.

References
Allison et al. – Social perception

84 Sugano, Y. et al. (1998) Global and fine information coded by single neurons in the temporal visual cortex. Nature 400, 860–873
87 Chao, L.L. et al. (1998) Are face-responsive regions selective only for faces? Neuroreport 10, 2940–2943