Do monkeys understand actions and minds of others? Studies of single cells and eye movements

Lorincz, Erika; Jellema, Tjeerd; Gómez, Juan-Carlos; Barraclough, Nick; Xiao, Dengke; Perrett, David I.

How to cite

LORINCZ, Erika et al. Do monkeys understand actions and minds of others? Studies of single cells and eye movements. In: From Monkey Brain to Human Brain: A Fyssen Foundation Symposium. Dehaene, S.; Duhamel, J.-R.; Hauser, M. D. & Rizzolatti G. (Ed.). Cambridge, Ma : MIT Press, 2005. p. 189–211. (Fyssen Foundation symposium. Fyssen Foundation series.) doi: 10.7551/mitpress/3136.003.0014

This publication URL: https://archive-ouverte.unige.ch/unige:167515
Publication DOI: 10.7551/mitpress/3136.003.0014

Book: ***From monkey brain to human brain***
Editors: S. Dehaene, J.-R. Duhamel, M. D. Hauser and G. Rizzolatti
Publisher: The Fyssen Foundation, MIT Press
Year: 2005
Pages: 189-211
ISBN: 1-84169-349-9

Do monkeys understand actions and minds of others? Studies of single cell and eye movements.

E. N. Lorincz[1], T. Jellema[2], J-C. Gómez[1], N. Barraclough[1], D. Xiao[1], and D. I. Perrett[1]

[1]School of Psychology, University of St. Andrews, Fife KY16 9JU, UK.
[2]Department of Experimental Psychology, Helmholtz Research Institute, Utrecht University, 3584 CS Utrecht, The Netherlands.

Abstract

In our visual analysis of behaviour, we may go beyond a perceptual description of actions and make inferences about mental states (intentions and beliefs). For example, we can see the action of reaching to an object as goal-directed or the action of looking as attention to a specific target. More complexly, we can infer from reaching an intention or desire to get an object (even if the action is unsuccessful), and from attention to an event the acquisition of *knowledge* about that event, or conversely from lack of attention we can infer *ignorance* or even false beliefs. The extent to which non-human primates are able to make such inferences is unclear. We consider here, (i) the visual processing that is available to macaque monkeys to describe behaviour of others, and (ii) experiments on the behavioural capacity of monkeys to exploit such visual processing to predict events based on another's knowledge states. We conclude that monkeys possess the neural machinery for discriminating intentional behaviour and predicting the events (end postures or reappearance) from actions. At the behavioural level we found evidence that monkey eye movements anticipate events based on an awareness of action of others, however, we were unable to demonstrate the capacity to represent another's knowledge states.

Overview

In order to understand the nature of visual processing that might support comprehension of actions it is useful to first review the general scheme of cortical visual processing. This is done in Section 1.1. Action coding appears to be prevalent within one area of the higher-level visual cortex labelled the STS (Superior Temporal Sulcus). Section 1.2 reviews the spectrum of cell types found in this area, while Section 2 reviews the cell properties that apply to comprehension

of the behaviour of others. Section 3 is concerned with the question of how the neural apparatus might translate into observable behaviour guided by the actions of others.

1.1 Different streams for processing the visual stimulus

Existing ideas about where in the brain the processing of the different features of a visual stimulus occur are heavily influenced by the Ungerleider and Mishkin model (1982), and by a subsequent adaptation by Milner and Goodale (1995). The Ungerleider and Mishkin model envisages a separation of visual processing into two distinct cortical streams: a dorsal "where" stream, extending from V1 into the inferior parietal cortex, primarily dealing with the spatial relationships of objects, and a ventral "what" stream, extending from V1 into the inferior temporal cortex (IT) dealing with the shape and identity of objects. The role of the ventral stream in the recognition of complex objects is supported by findings showing a gradual increase in the complexity of stimuli analysed by cells from V1 up to IT (Perrett & Oram, 1993).

Milner and Goodale questioned the strict "what-where" dichotomy, and suggested that space and form are processed in both parietal and temporal areas but for different purposes (Milner & Goodale, 1995). In their view, the ventral stream subserves visual 'perception', i.e. object and scene recognition, requiring allocentric spatial coding to represent the enduring characteristics of objects, while the dorsal stream subserves the visual control of 'action', requiring egocentric spatial coding for short-lived representations (vision for perception *versus* vision for action).

These ideas of joint processing of form and position have been substantiated by studies at the cellular level, with cells coding for object shape in parietal cortex (Murata *et al*., 1996; Sereno & Maunsell, 1998), and cells coding for the object's spatial position in area V4 within the ventral stream (Dobbins *et al*., 1998).

1.2 The anterior part of the superior temporal sulcus (STSa) in the monkey

The anterior part of the superior temporal sulcus (STSa) in the macaque monkey, corresponding to area STPa (Bruce *et al*., 1981), is considered part of the ventral visual stream. STSa cells often maximally respond to the visual appearance of the face and body, and to body actions, most notably of conspecifics and humans (Földiák *et al.*, in press). Gross *et al*. (1972) made the first startling finding of cells that responded selectively to the sight of a specific body part, for example, a monkey's paw. Subsequent work in STSa revealed populations of cells selective for the sight of faces and bodily actions, such as articulations of the limbs and torso, but also whole body actions such as walking (Bruce *et al*., 1981; Perrett *et al*., 1982, 1984, 1985a,b, 1989, 1992; Desimone *et al*., 1984; Jellema & Perrett, 2002; Jellema *et al.*, 2002).

Other STSa cells are tuned to multiple views of the same animate object or the same action (Logothetis *et al*., 1995; Perrett *et al*., 1989; Jellema *et al*., 2002), or are tuned to conceptually related visual stimuli, such as multiple body signals of directed attention (Perrett *et al*., 1985a, 1992). This selectivity is most likely obtained through pooling of the outputs of cells coding for distinct stimuli related by temporal association.

Characteristic for many STSa cells is that they integrate information about form and motion of animate objects (Oram & Perrett, 1994, 1996; Tanaka *et al*., 1999) and, as has only recently become clear, integrate information about the spatial location of animate objects

(Baker *et al*., 2000; Jellema *et al*., 2004). STSa cells often generalise their sensitivity to complex shapes across changes in various other stimulus properties such as size, retinal position, orientation, the species (human or monkey), luminance and colour (e.g. Perrett *et al*., 1984, 1989; Rolls & Baylis, 1986; Ashbridge *et al*., 2000; Földiák *et al.*, in press).

Collectively these summarised response characteristics suggest a role in object recognition and allocate the STSa to the ventral visual stream (Milner & Goodale, 1995). The findings in the macaque monkey have led to the idea that the STSa is primarily involved in the visual analysis of actions performed by other individuals (Perrett *et al*., 1989). This view is supported by recent brain imaging studies which show activation of the human posterior STS, which area is thought to be the homologue of the monkey STSa, for the perception of biologically significant stimuli. Examples of such stimuli are 'biological motion' of human figures, hand actions, static faces, eye-gaze and eye-motion, and meaningful actions; (for reviews see Allison *et al.*, 2000; Puce & Perrett, 2003). A recent study explicitly showed that the human posterior STS represents goal-directed or intentional actions (Saxe *et al*., 2003).

2. Action coding in the STS

2.1 Orientation of body cues

The visual information arising from body cues appears to contribute to STS cell sensitivity in a way that is consistent with the cell's role in analysing the direction of attention. For example, cells tuned to the left profile view of the head are often additionally tuned to the left profile view of the body (Wachsmuth *et al*., 1994). However, the direction of another individual's head or body may not always be a reliable index of where that individual's attention lies (Perrett *et al*., 1992). Gaze direction seems a more powerful guide in this respect and gaze may therefore also be expected to affect STSa cell responses. Indeed tuning to both head view and gaze direction seems relatively common in the STS. Moreover when conjoint sensitivity is observed, the effective gaze directions match the effective head directions, and gaze in an ineffective direction can prevent the response to an otherwise effective head posture. Thus cellular coding of head and gaze direction seems compatible with gaze direction taking precedent (Perrett *et al.* 1985, 1992).

Despite the findings of cellular sensitivity to attention direction in macaques, it still remains a matter of debate whether Old World monkeys are able to use information about the gaze direction of others. For example, Anderson *et al*. (1996) report that macaque monkeys cannot be trained to use human gaze to locate hidden food. Behavioural assessments in our lab, however, have shown that macaques do spontaneously utilise the direction of attention of conspecifics to orient their own attention (Emery *et al*., 1997; Lorincz *et al.*, 1999). We will return to the issue of this discrepancy later in the paper.

2.2. Modulation of Action coding

2.2.1 Modulation by attention

Cells in STSa that are selectively responsive to articulations of limbs or parts of the face or body, such as the mouth, eyes, head, torso, legs, arms, hands and fingers, have also been documented (Perrett *et al*., 1985b, 1990; Mistlin & Perrett, 1990). Cells may respond selectively to arm movements and not to equivalent leg movements, or to leg movements and not to equivalent arm movements (Jellema *et al*., 2000). Additionally, the cells often showed sensitivity for the direction of the motion; some are tuned to reaching toward the observer, others to reaching to the observer's left, etc. Movements directed away from the subject may acquire particular meaning in a given context. For example, if a food tray is kept out of sight to the subject's right, then the sight of the experimenter reaching right can become salient, since this may bring the experimenter's hand to food, which is subsequently given to the subject.

For a subset of the population of STSa cells responding to the sight of arm reaching, the response could be modulated by the direction of attention of the agent performing the action (Jellema et al., 2000). Actions performed when the agent faced and gazed in the direction of reaching were more effective than actions performed with the head and gaze oriented away from the direction of reaching. Body posture, which provides another potential cue to the direction of attention (Perrett *et al*., 1992; Wachsmuth *et al*., 1994), was found to further contribute to the 'modulation' of the response to a reaching movement (Jellema *et al*., 2000).

Cells responses with conjoint selectivity for congruent reaching and attention can be formed by combining the appropriate outputs of cells that respond to directed attention with the outputs of cells that respond to directed limb movements. The significance of someone reaching toward an object while his/her attention is focused on the object clearly differs from an identical arm and hand action performed with attention directed elsewhere. In the former case one is likely to infer that it was this person's intention to reach out for the object in order to pick it up or make contact with it. In the latter case one may infer that the object was incidental to the arm extension. Crucial is that information about the reaching action, such as its direction and the possible presence of a reaching goal, is linked to information about the direction of attention of the performer. The responses of the cells to both the hand actions and to attention direction can be selective for movements of the agent that appear intentional. Actions that are attended to are more likely to be intentional whereas actions accompanied by attention elsewhere can result in accidental effects.

2.2.2 Modulation by goals

Some STSa cells seem to be sensitive to the causal relationship between an action and the object or goal of that action. This has been demonstrated most clearly in cells sensitive to purposeful hand-object interactions, such as reaching for, picking, tearing and manipulating objects (Perrett *et al.,* 1989, 1990; Jellema *et al.,* 2000). These STSa cells are sensitive to the form of the hand performing the action, and are unresponsive to the sight of tools manipulating objects in the same manner as hands. Furthermore, the cells code the spatio-temporal interaction between the agent performing the action and the object of the action. For example, cells tuned to hands manipulating an object cease to respond if (1) the object is removed, (2) the hand action is made in a direction away from the object, or (3) the hands and object move appropriately but remain spatially separated (Perrett *et al*., 1989). This selectivity ensures the cells are more responsive in situations where the agent's motion is *causally* related to the object's motion.

2.2.3 Modulation by location

The brain integrates different features of a visual stimulus, such as its form, colour, motion and location, into a single coherent percept. Milner and Goodale suggested that space and form might be processed in both dorsal (parietal) and ventral (temporal) streams but for different purposes (e.g. Goodale *et al*., 1991; Milner & Goodale, 1995). Recently we discovered cell populations in STSa that are sensitive to the spatial location of animate objects after they moved out of sight behind a screen (Baker *et al*., 2000, 2001).

More generally we found that spatial information is indeed integrated with form and motion information at the single cell level quite extensively within the STS. This capacity may enable STSa cell populations to form representations of goal-directed or socially relevant actions.

Jellema et al. (2004) tested cells responsive to the sight of walking (i.e. tuned to body view and direction of motion, Perrett *et al.*, 1984) for sensitivity to the position of the walking person within the testing room. More than half of the cells were sensitive for the spatial location of the agent. Some cells would respond only to walking at a far distance, others only to walking near. The locations of the walking agent can be referenced to the subject's perspective: near or far from the subject. This assumes an egocentric frame of reference, but in principle the cells could just as well have used an allocentric frame of reference (i.e. spatial descriptions based on environmental landmarks rather than the subject's own position and orientation). Allocentric coding has been observed for STSa cells sensitive to goal-directed actions (Perrett *et al*., 1989), and for an STSa cell coding for occluded agents (Baker *et al*., 2001). Such relative positions are especially relevant in social interactions.

Spatial coding may indeed be widespread in STSa. Previously it was suggested that STSa plays a role not only in animate object identification but also in the visual analysis of the intentions and goals of others' actions, which forms an important aspect of social cognition (Emery & Perrett, 1997; Jellema & Perrett, 2002). The significance of spatial coding in STSa must be seen in this light. The spatial positions that individuals occupy with respect to each other, or with respect to objects, contain essential information for an observer when it comes to determining the goals and intentions of those individuals.

2.3. Temporal associations in action coding

2.3.1 Implying next or past motion from current posture

Actions performed by most animals typically involve articulation. To understand an articulated action performed by another individual, we do not necessarily have to witness the entire action sequence. A single momentary view is often enough to identify the likely action about to be performed, or recently executed. The same momentary view may permit the identification of the probable goal of the action. This is a very useful capacity since it allows us to understand an agent's actions even when we get only one glimpse of an agent or goal before the agent becomes hidden from view.

Articulated motion seems to be preferentially processed in STS, as shown by Beauchamp *et al*. (2003), who made a direct comparison between articulated and non-articulated human

motion, and found that the former activated the STS significantly more than the latter. Jellema and Perrett (2003b) studied the sensitivity of STSa cells to body postures containing implied motion and to actual articulated body movements. It was postulated that if the static articulated posture were to be presented in isolation (i.e. in absence of actual movement), STS cells would respond as if the associated actual motion was presented.

Articulated actions were defined as actions where one body part (e.g. a limb or head) moves with respect to the body part it is attached to; conversely non-articulated actions are actions where the equivalent body parts do not move with respect to each other but move as one. Similarly, articulated static body postures contain a torsion or rotation between parts, while non-articulated postures do not (i.e. the head, chest and pelvis aligned and oriented in the same direction, typical of an 'at rest' posture).

Testing cells that were responsive during particular actions revealed that 55% of cells responded both to the articulated action and to the articulated static posture that formed the end-point of the action presented in isolation. The cells did not respond to the sight of the non-articulated static posture, which formed the starting point of the action. Moreover the cells did not respond to static postures resembling the articulated end point posture, but which were in a more relaxed muscular state (i.e. non-articulated). The cells also did not respond to other articulated body actions that were less often associated with the effective static articulated posture.

The above findings give rise to the intriguing possibility that the STSa cells code for a particular articulated action both when actually presented and when implied in a still image. Previously, STSa cell responses were described that were tuned to the same perspective view of multiple parts of the body (e.g. left profile view of the trunk and left profile view of the head; Wachsmuth *et al*., 1994). The cells described here required different perspective views of body segments (i.e. torsion or twisting of body parts). We suggest that the cells code for the implied motion contained in the static articulated posture, or, in other words, code for the association of motion and posture, rather than for the articulated posture *per se.*

The STS could support recognition of an object and the likely type of movement associated with that object. Representation of such an association could allow inferences as to whether or not the object was likely to be moving when it was briefly seen (or at the time a picture was taken), and its most likely previous and future trajectory.

The neural representations in STSa for *actual* biological motion may also extend to biological motion *implied* from static postures. The data show that the visual processing of static form may contribute to the comprehension of dynamic actions. Sensitivity to associations between image form and motion could form the basis of the ability of the nervous system to retrieve likely motion implied entirely by static images.

2.3.2. Sequences of actions and postures

Under natural viewing conditions, STSa cell responses to the sight of static body postures may be controlled by actions performed by that body in the one or two seconds immediately preceding the onset of the static posture (Jellema & Perrett, 2003a). In other words, the perceptual history can enable or prevent a cell's response to the current retinal input. For example a cell may respond vigorously to the sight of e.g. a face when the face was preceded by action A, but fail to respond to the identical face when preceded by action B and fail to respond

to the sight of a face when presented without any preceding action. For 54% of cells sensitive to the static posture, the nature of the movement preceding a static posture proved critical to the modulation of cell responses.

These results show that the 'vocabulary or grammar' of actions and body postures coded by single STSa cells is much larger than previously thought. Thus cells in temporal cortex could support the formation of expectations about impending behaviour of others, which suggests a role in the understanding of actions. The neural representations for sequences of events may play a role in predicting or anticipating the next move or posture of the animate object. For example, the sight of a body that has just stopped walking forward may invoke an expectation that, should walking commence again, it is likely to resume a forward direction. The same view of a static body that has just stopped walking backward, by contrast, may be expected to move again in a backward direction should walking resume.

2.3.3. Action sequences that become hidden

The actions of others are not always fully visible, for example someone may become hidden from our sight as they move behind a tree, or their hands may not remain fully in view as they reach to retrieve an object. Within STS it is now apparent that specific cell populations are activated when the presence of a hidden agent can be inferred from the preceding visual events (i.e. the agent was witnessed passing out of sight behind a screen and has not yet been witnessed re-emerging into sight, therefore the agent is likely to remain behind the screen; see Baker *et al.*, 2001). The population response of STSa cells to this sequence of events is shown in Fig. 1. It shows that STSa cells responded maximally when individuals were seen to 'hide' behind an occluding screen. In the 3 s following disappearance from sight behind a screen, the population response was larger than in the prior 3 s when the agent was visible and moving toward the screen. Some cells had no detectable response to visible movements but started responding 1-4 s after the agent had become completely hidden.

The cells responding to occlusion additionally showed spatial sensitivity discriminating between locations at which the agent was completely hidden (at the left, right or middle of the room; Baker *et al*., 2001). Cell responses to the experimenter walking in-sight were consistent with the out-of-sight responses. For example, if hiding behind a screen located at the right-hand side of the testing room evoked significantly larger responses than hiding behind a screen at the left-hand side, then walking toward the right-hand screen would also evoke a larger response than walking toward the left-hand screen, with walking in both cases from left to right. These responses are consistent with the idea that the cells coded not only for the presence of the experimenter behind the right-hand screen, but also for the intention of the experimenter to go behind that screen. For this interpretation, we need only assume that walking towards the right screen reflects the intention to move behind that screen.

The directional selectivity and position coding could be used to anticipate the re-emergence of a hidden agent. Thus a monkey watching a human disappear behind a screen, in some sense is aware that the human is hidden from sight and that he or she is likely to reappear from one or other side of the screen.

2.4. Relation of visual coding in STSa to motor planning in premotor cortex

The STSa cell populations coding body and hand actions appear to be predominantly sensory, although information from the motor system does affect other STSa cell populations (Hietanen & Perrett, 1996) and modulates STS activity in humans (Iacoboni *et al.,* 2001; Nishitani & Hari, 2001). Gallese and Goldman (1998) suggested that the 'action detecting' system in STSa could provide an initial 'pictorial' description of the action, and this information is then relayed to frontal motor planning systems. The manner in which temporal STS and frontal systems interconnect is not fully clear, but probably involves intermediate processing steps mediated by parietal areas (Nishitani & Hari, 2000, 2001; Gallese *et al.,* 2002).

The frontal region of primate cortex (inferior area 6) has long been known to be somatotopically organised for the representation and control of movements of the mouth and arm (Rizzolatti *et al.,* 1988). This area can be subdivided into areas F4 and F5 (Rizzolatti & Gentilucci, 1988). Neurons in F5 are activated during specific motor acts performed with the hand or mouth, such as grasping, holding and tearing. Cells with responses related to reaching movements of the arm are typically found in F4.

Cells in F5 discharge during the execution of a particular action and during the sight of the same action. For example, an F5 cell, which responds selectively when the monkey executes a grasping action, may also respond (like STS cells) to the sight of another monkey or the experimenter grasping an object, but not to the sight of different hand actions such as tearing (Di Pellegrino *et al*., 1992; Gallese *et al*., 1996; Rizzolatti *et al*., 1996a,b). An F5 cell selective for the execution of grasping would also respond when the monkey grasps an object in the dark, thereby demonstrating the motor properties of the response. These conjoint properties have led Rizzolatti *et al.* (1996a,b) and Gallese *et al.* (1996) to postulate that the F5 neurons form a system for matching observation and executing actions for the grasping, manipulation and placement of objects. These neurons have now been labelled "mirror" neurons. Cells with response selectivity similar to those in F5 have recently been reported in the inferior parietal lobule (Gallese *et al*., 2002; see Williams *et al.,* 2001).

The experiments in which actions are partially, or totally, occluded from sight have also highlighted the similarities of STS and F5 systems in the processing of actions. F5 cells may respond, in a manner analogous to the STS cells, to the sight of the agent reaching to grasp an object. The same F5 cells are active when the experimenter places an object behind a screen and then reaches as if to grasp it (even though the object is hidden from view during part of the action; Umilta *et al*., 2002). The sight of equivalent reaching when there is no reason to believe an object is hidden from sight fails to activate the F5 cells. Thus F5 and STS cells code the sight of actions on the basis of what is currently visible and on the basis of the recent perceptual history (Jellema & Perrett, 2002, 2003a; Jellema *et al.,* 2002).

Thus, the visual properties of mirror neurons in F5 cell are strikingly similar to those described in STSa. Both F5 and STSa cells respond when the monkey observes the experimenter reaching and grasping an object, but will not respond to the sight of the experimenter's hand motion alone or to the sight of the object alone. In addition, the F5 cells respond to the corresponding sound of actions (Kohler *et al.,* 2002). STS cells too appear to be sensitive to the sound of actions (Barraclough *et al.,* 2003). In addition, the F5 cells but apparently not STS cells respond to the execution of the corresponding motor act. Therefore the F5 mirror system may provide a supra-modal conceptual representation of actions and their

consequences in the world. To some extent the poly-modal sensory representation of actions may be constructed in the posterior and ventral regions of the brain (i.e. temporal lobe). The integration of the sensory representation with the motor representation of actions reflects an additional processing in parietal and frontal systems.

While STS and F5 cells have similar visual properties they may subserve distinct functions; the frontal system perhaps serves to control the behaviour of the self particularly in dealing with objects (Rizzolatti *et al*., 1996a,b), whereas the STS system is specialised for the detection and recognition of the behaviour of others (Hietanen & Perrett, 1996; Mistlin & Perrett, 1990; Perrett *et al*., 1989).

The mirror neuron system might also complement the STSa description of the perceived action by adding information about the motor requirements of the perceived action, which could not easily be obtained from purely visual features. Crucially the properties of the frontal mirror system suggest that we may understand actions performed by others because we can match the actions we sense through vision and audition to our ability to produce the same actions ourselves. At a more speculative level, it has been proposed that the mirror neurons are involved in the ability to 'read' others' minds. The cells may allow an observer to 'experience' and understand an action performed by another through 'simulating' that action (Gallese & Goldman, 1998).

2.5. Summary of STS action coding

Cells in the STSa of the macaque monkey code not only the sight of others' bodily actions but also combinations of other visual cues. These other visual cues derive either from the body, e.g. head/eye gaze direction and articulated body posture, or from the environment, e.g. objects acted upon, and the spatial location of the actor. There are two ways in which the combined sensitivity arises: through spatial conjunctions with other cues that are simultaneously present, and through temporal conjunctions with other cues that are present consecutively.

Combining the sensitivity for actions with sensitivity for other bodily or environmental cues puts the STSa in a position to form representations of the causality or goal-directedness of others' actions. For instance, conjunctions of perceived actions with the perceived eye gaze direction of the agent performing the action may contribute to detecting the accidental or intentional outcome of the perceived action.

The relative spatial locations of the agent and the objects the agent interacts with (including the observer) may give insight to the agent's intention or goal. Populations of STSa cells are well equipped to keep track of such spatial relationships between agents and objects, which may further support comprehending the actions of others.

The object to which an action is directed is particularly important when it comes to interpreting goals. Some STSa cells are sensitive to the spatio-temporal interaction between the agent performing the action and the target-object of the action. Such selectivity ensures the cells are optimally responsive in situations where the agent's motion is causally related to the object's motion.

The sensitivity to sequences might also contribute to representations of goals. The formation of associations between an articulated action and the static articulated end-posture of that action might well underlie the ability of the brain to imply impending or prior action from currently visible static postures. The performance of dexterous manual tasks can easily be

specified as a series of static pictures, each demonstrating particular sub-goals or stages in the action sequence. Based on an understanding of momentary postures during an action sequence, individuals can infer the dynamics of how an action was performed.

Other STS cells are tuned to the perceptual history within action sequences. Witnessing the history prior to viewing a static body allows one to predict the likelihood and nature of the body's future movement with more certainty than from a still image of a person performing a motor act. Natural actions are not isolated postures, but are continuous and complex sequences of postures with linking movements. The STS cells could play a role in recognising complex action sequences and predicting the most likely next stage or consequence of actions.

Sequences of events are also crucial for those STS cells that code for agents hidden behind a screen. The observer needs to witness the agent disappearing behind the screen in order for the cell to produce a response to the hidden agent. The cellular responses to temporarily hidden agents combined with their sensitivity to direction of the agent's motion and spatial position allows again for prediction of when and where the agent may reappear.

Determining the goal or intention of an action not only involves the sensory systems but also emotional (Adolphs, 1999) and motor (Gallese & Goldman, 1998) systems. The ability to determine others' intentions is thus likely to be generated in a widely distributed network, involving many brain areas. The STSa could play a role in this network by providing descriptions of others' actions in terms of goals, intentions or causes. These descriptions are, however, still mechanistic in nature. We have no evidence that descriptions include the attribution of mental states, such as motivational drives and beliefs (Baron-Cohen, 1994; Saxe *et al.*, 2003), to the agent performing the action.

The advantages of being able to determine others' intentions from their actions are clear. The observer may anticipate the nature of the future actions of the other individual and adjust responses accordingly. An obvious question is of course whether there is behavioural evidence that non-human primates indeed discriminate between intentional and non-intentional actions (Byrne & Whiten, 1988). Macaque monkeys spontaneously follow the direction of attention of other monkeys (Emery *et al.*, 1997; Lorincz *et al.*, 1999). Similar gaze following abilities have been described in different species of primates (see, for example, Tomasello, Call and Hare, 1998). Attention following does not, however, guarantee understanding of the consequences of attention, for example that seeing leads to knowing. Studis by Call and Tomasselo (1998), and Hare et al. (2000, 2001) suggest that chimpanzees may understand what others can see, that seeing leads to knowing and the distinction between intentional and accidental actions For monkeys, however, evidence of comprehension of beliefs or intentions is lacking.

In the following section we consider behavioural experiments designed to probe the monkey's understanding of an agent's actions and the possibility of anticipating impending actions on the basis of prior history and the agent's knowledge state.

3. Behavioural studies of actions, intentions and belief

Although electrophysiology data tend to suggest that monkeys can anticipate others' actions and build some expectation of 'what should come next' in a sequence of movements, behavioural data that support the existence of an even partial 'Theory of Mind' in monkeys are scarce and have given negative results. For example, Hare et al. (2003) found that capuchin

monkeys failed to understand what others can see in the same test in which chimpanzees had succeeded.

To get an insight into this issue, we examined the eye gaze behaviour of one macaque monkey as an index of the monkey's ability to predict the behaviour of an experimenter in a feeding situation. We studied the gaze behaviour as the monkey became accustomed to a particular feeding scenario. We reasoned that the monkey might learn to predict the experimenter's actions and that the monkey's gaze could in principle show evidence of anticipating the intentions of the experimenter.

We simulated the 'Sally-Anne' false belief test used with children. In this task, Sally places an object in one position, and leaves the room for a short time. During Sally's absence, Anne moves the object and hides it in a new location. Individuals with autism and typical children under 4 years incorrectly predict that Sally will search at the new shifted location on her return. Such individuals are said to lack a 'Theory of Mind' because they do not understand that Sally's actions will be guided by her false belief that the object is where she left it.

Comprehension of a situation (physical or interpersonal) may outstrip actual overt behavioural performance. The understanding of the actions of others may develop faster than the ability to act on the basis of that understanding. Performance failure can occur because the operant behaviour that is required by the experimental task is one that is guided by a pre-potent response. For example, chimpanzees are unable to withhold a reaching response to the more numerous (or larger) of two arrays of food rewards in a task where the experimental rule was to indicate the smallest array in order to get the biggest (Boysen *et al*., 1995, 1996, 2001). If reaching is required not directly to one of two differently sized food rewards, but to one of two symbols (Arabic numerals) which represent the different quantities, then the chimpanzees were able to learn the rule to point to the symbol signifying the smaller quantity in order to receive the largest reward.

At a speculative level failure may reflect inadequate development of frontal cortex and an inadequate ability to give up short-term hedonistic impulses in exchange for the largest benefit that could come with deferred gratification. 'Grab what you can now and don't care about what tomorrow might bring'. Indeed much of old world monkeys' behaviour might be guided by such short-term rules and not be subject to more strategic planning (Henzi & Barrett, 1999; Barrett & Henzi, 2000, 2002).

Similarly, several studies suggest that children under 4 years of age who fail traditional false-belief tasks may show a better understanding of false-belief if they are tested with simplified procedures with less executive demands. For example, Carlson et al. (1998) found that children are better able to mislead an opponent if instead of pointing to an empty location they are allowed to mark it with a sticker. More pertinent for our study, Clements and Perner (1994) reported that children who fail a standard Sally/Ann false belief test with their verbal answers, however looked at the correct location

Performance failures in tasks may therefore depend on the behavioural index chosen as an indication of comprehension. It may be that a reaching response, particularly for a reward, is too difficult to inhibit. By contrast, eye movements usually do not gain an individual reward. Where an individual looks may therefore be a better guide to potential comprehension. Dissociations between knowledge expressed in gaze behaviour and knowledge expressed in action have been

found in other areas of cognitive development, such as object understanding (see Hauser, 2003, for a review).

Whatever the explanation, it is apparent that eye movements may provide the first indication of a developing 'Theory of Mind'. The duration of looking is often used as an index of the ability of an individual to perceive and understand the difference between two stimuli or situations. It is assumed that the tested individual will look longer at the more unusual of the two stimuli, since habituation leads to decreased looking time and novel situations lead to longer durations of looking. Looking behaviour can be compared in two scenarios in the Sally-Anne task. In one scenario Sally returns, goes and searches at the site where she left the item (i.e. where her behaviour would be guided by the nature of her false belief). In the second scenario Sally returns and goes and searches at the new site, the site where Anne has hidden the object. Here Sally's behaviour is exceptional since she has a knowledge state that is inconsistent with her actual behaviour.

Santos and Hauser (Hauser, 1999; Santos & Hauser, 1999 and personal communication) have studied the looking behaviour of Tamarin monkeys. Tamarins looked longer in the second scenario than in the first Sally-Anne scenario (Hauser, 1999; Santos & Hauser, 1999 personal communication). This would indicate that they perceive the difference between the two situations, but does not necessarily indicate that the monkeys have a 'Theory of Mind'. To conclude that the monkeys can comprehend the actions of Sally based on Sally's false belief requires further tests. For example, if Sally remains present in the room while Anne shifts the location of the object, then Sally would have a true belief about the object's new location. In this case, when Sally attempts to retrieve the object, Sally should search at the new location, not at the location she originally placed the object.

Using an eye-tracking system, we studied the looking behaviour of one monkey in a Sally-Anne situation. The testing situation involved two boxes to the left and right of a central screen (1.02 m from the monkey). Boxes were oriented so that the monkey could always see the contents. By contrast Sally could not see the contents during her approach. On a 'Start' trial, an experimenter (Sally, who normally feeds and interacts with the monkey) would enter the room from one side. There was then a period (average 3, range 2-5 seconds) during which Sally walked towards the monkey while hidden from sight behind the central screen. Sally then emerged on one side of the screen carrying food and a feeding device such as a spoon, which she would use to feed the monkey. Sally would then exit the scene on the same side as she entered, leaving the spoon in the box on that side. On a following 'Anticipation' trial, Sally would always re-enter the room with more food. Sally's approach was again obscured from sight during her return for 3 seconds behind the screen, after which she emerged to retrieve the spoon from where she had placed it and then fed the monkey.

During Start trials the spoon could be placed in the left or right boxes but on ensuing Anticipation trials Sally always returned to the side where she had placed the spoon. We measured the monkey's gaze during the period in which Sally was hidden from view behind the screen. We found that over sessions the monkey's eyes became governed by looking strategies. The first strategy was to look on the side of the screen Sally had been '*last seen*' before she became occluded behind the screen. This *'last seen'* strategy was evident on Start trials where there could be no prediction of the side on which the monkey was next going to be fed. Indeed the '*last seen'* strategy was superstitious, in that the side of Sally's actual re-emergence

following occlusion was random on such Start trials. Thus the monkey's anticipatory looking was not guided by any auditory cues to the approach or reappearance of Sally.

On Anticipation trials two strategies governed the anticipatory looking behaviour during occlusion from sight: the *'last seen'* strategy and a new strategy which predicted Sally's actual behaviour. The monkey would look more on the side Sally had previously left the feeding tool. This was the side from which '*Sally last fed*' the monkey and from which Sally would soon re-emerge to collect the feeding tool and use it to feed the monkey.

This baseline behaviour was established over 10 daily sessions (inc. 62 Start and 220 Anticipatory trials). In essence it showed that the monkey developed expectations about Sally's behaviour in the feeding scenario. With this baseline established we progressed to 'Critical trials'. Here a second experimenter (Anne) entered after Anticipation trials and approached occluded from view behind the screen and then re-emerged randomly on the left or right side and searched in the box for the spoon. If Anne did not find the spoon, she moved to the other box and searched there. Once Anne had found the spoon, she moved the spoon from the box from where Sally had placed it to the box on the opposite side. During all of this time Sally was away from the scene and therefore ignorant of the re-location of the spoon. On the next trial Sally returned and after the 3 seconds hidden approach behind the screen re-emerged on the side where she had placed the spoon (where the spoon is according to her false belief). She searched there but did not find the spoon. She then searched the box on the opposite side, found the spoon and without pause returned to the original side to feed the monkey. In these critical trials the monkey's looking strategy (during Sally's temporary occlusion behind the screen) was significantly biased to the side where Sally should return guided by Sally's false belief (i.e. the side that did not contain the spoon).

We conducted various control trials to investigate whether the monkey's looking behaviour was guided by Sally's knowledge state or some other rule. In 'Critical Control' trials, Anne moved the spoon as before and exited. Anne then returned with food and after a 2-3 second pause hidden behind the screen, Anne emerged to search where she herself had left the spoon, which she picked up and then used to feed the monkey. On these trials the monkey looked differentially at the scene during the time Anne was hidden. The monkey's eyes, however, did not anticipate where Anne would emerge (i.e. the monkeys fixation pattern was not guided by Anne's knowledge state). Indeed the monkey looked more to the side from which Anne had last been seen (*Anne last seen*) which could either be the side Sally last fed the monkey or the opposite side. On trials with Anne about to reappear, the monkey's gaze was governed by the prior visible behaviour of Anne and not the prior behaviour of Sally. So the monkey's gaze and anticipation depended upon the identity of the experimenter.

The monkey's gaze was, however, <u>not</u> governed by Anne's knowledge (i.e. Anne's intention to collect the spoon and her 'true belief' of its location). On more careful consideration (of all trial types, see below), Anne used the spoon to feed the monkey on only 17% of the times she was seen by the monkey to pick up the spoon. One can argue that the monkey had little reason to anticipate feeding by Anne, so the monkey's gaze behaviour was perhaps appropriately unconnected with Anne's feeding actions.

We used two further control trials. In 'Visible swap' trials Sally witnessed Anne moving the spoon (here Anne has a true belief of the location of the spoon). In 'Touch trials' Anne touched but did not move the spoon. Combining the data from all unusual trial types (i.e. trial

sequences involving Anne: Critical, Critical Control, Visible swap and Touch trials), we found that the monkey's pattern of looking most reflected the side on which Sally had last fed her rather than Sally's or Anne's knowledge state as to where the spoon was and hence where they would search (Fig 2).

Thus the monkey's looking behaviour revealed an understanding of the actions of others, but it was not governed by their knowledge state.

Nonetheless the monkey's looking behaviour was subtly affected by the actions of Anne and Sally. We compared Critical trials (in which Sally is ignorant of Anne's behaviour) with Visible swap trials where Sally sees Anne's naughty tool relocation behaviour. This comparison revealed that the monkey's looking changed between the two trial types. In the Critical trials looking was governed solely by a *last fed* strategy which correctly anticipated Sally's re-emergence, whereas in the Visible swap trials looking was governed by both *last fed* and *last seen* strategies. Recall that the *last seen* strategy had been used on trials where Sally's impending behaviour was unpredictable. This is tantalizing evidence that the monkey was less able to predict Sally's next step behaviour on the basis of what had been witnessed (Anne relocating the tool and Sally seeing Anne doing this).

These studies were exploratory and the conclusions are tentative. We might be tempted to conclude that there was no evidence for 'Theory of Mind' in monkeys. This may be premature because we do not yet know what humans with or without 'Theory of Mind' would do in exactly the same control scenarios that we explored. Indeed, the studies of looking behaviour by young children (Clements & Perner, 1994) and of Tamarin monkeys (Hauser, 1999) concluded in favour of an understanding of belief on the basis of the same type of evidence we obtained in the Critical trials. Further studies are thus needed in human and non-human primates to investigate the strategies they use to guide their behaviour in response to actions, intentions and beliefs of others.

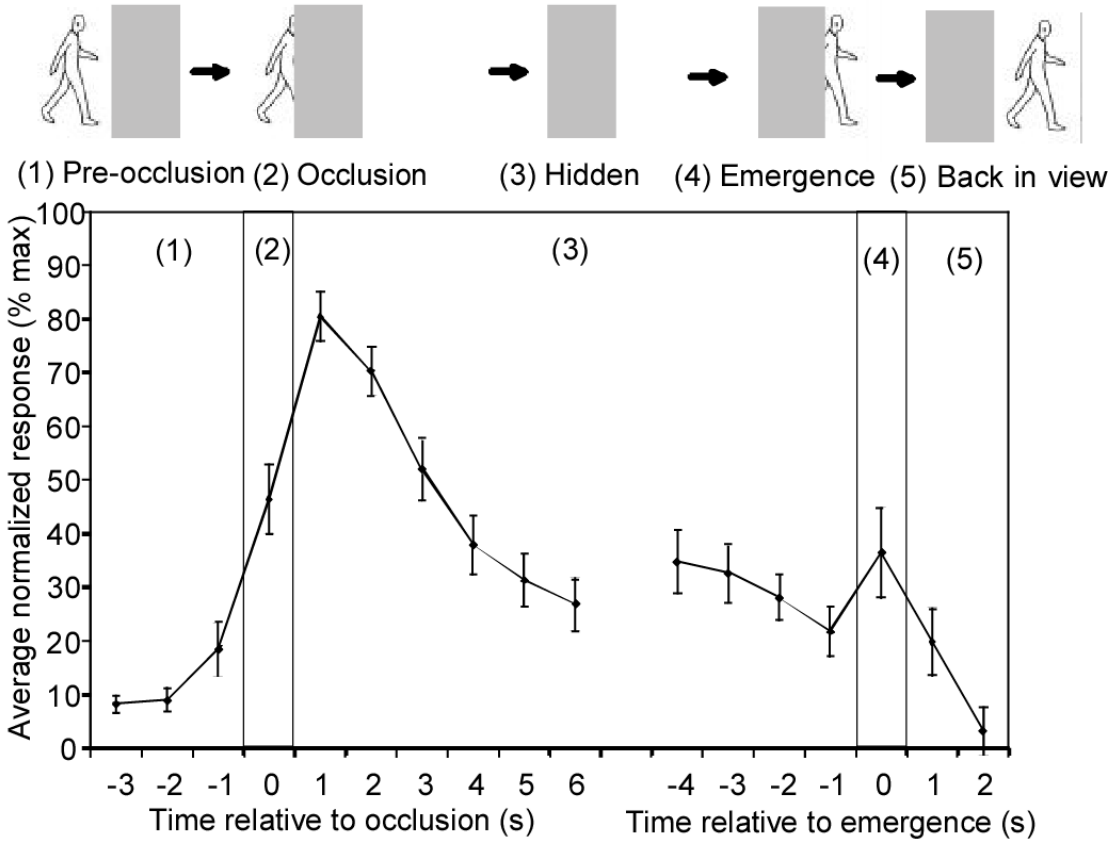

Figure 1. STS population response before, during and after an occlusion from sight. (Upper) Periods of the visual stimulus. The experimenter moved towards the occluding screen (1), was gradually occluded (2), remained hidden from view with only the screen visible (3), and gradually re-emerged (4), until the experimenter was once again fully in view (5). Filled arrows show the progression of events. (Lower) Activity profile during the disappearance and subsequent emergence of the experimenter. The graph shows the average normalized population response of 26 cells recorded in STSa. On the left of the graph, responses are aligned with respect to the occlusion period. On the right, the responses are aligned with respect to the emergence period. (Adapted from Baker *et al*., 2001).

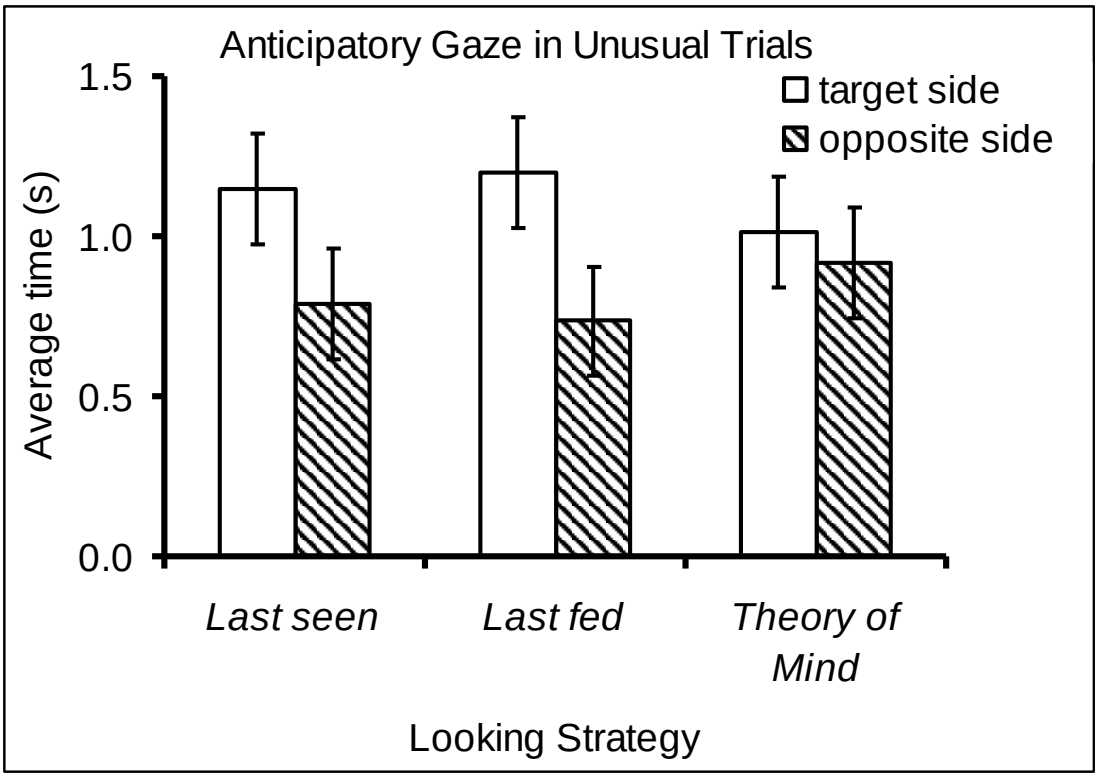

Figure 2. Gaze of one monkey during unusual trials in the Sally-Anne false belief scenario. Mean duration (s) of anticipatory gaze to the target and the opposite side of a central occluding screen (where the target side was defined by the looking strategy). Gaze was measured in the 3-5 s period during which the experimenter (Sally, or Anne) was hidden from sight prior to reappearing to search for a spoon to be used in feeding the monkey. The monkey's gaze was goverened overall by the strategy of looking to the side on which she had been *last fed* by Sally ($F_{1,40}$=5.0, p=0.03). There was a trend for looking to be guided by a *last seen* strategy (i.e. the side on which Sally had been seen prior to disappearance behind the screen, $F_{1,40}$=3.2, p=0.08). The *Theory of Mind* strategy, where looking would be guided by true or false beliefs of Sally (or Anne), did not predict looking ($F_{1,40}$=0.2, p=0.6).

References

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

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4,* 267-278.

Anderson, J.R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioral Processes, 37,* 47-55.

Ashbridge, E., Perrett, D.I., Oram, M.W., & Jellema, T. (2000). Effect of image orientation and size on object recognition: responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology, 17,* 13-34.

Baker, C.I., Keysers, C, Jellema, T., & Perrett, D.I. (2000). Coding of spatial position in the superior temporal sulcus of the macaque. *Current Psychology Letters: Behaviour Brain and Cognition, 1,* 71-87.

Baker, C.I., Keysers, C., Jellema, T., Wicker, B., & Perrett, D.I. (2001). Neuronal representation of disappearing and hidden objects in temporal cortex of the macaque, *Experimental Brain Research, 140,* 375-381.

Barraclough, N. E., Xiao, D., Oram, M. W., & Perrett, D.I. (2003, 7-13 November 2003). Primate superior temporal sulcus (STS) neurons integrate visual and auditory information for biological motions. *Paper presented at the 33th Annual Meeting of the Society For Neuroscience, New-Orleans, La.*

Barrett, L., & Henzi, S. P. (2002). Constraints on relationship formation among female primates. *Behaviour, 139,* 263-289.

Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E., & Hill, R. A. (2000). Female baboons do not raise the stakes but they give as good as they get. *Animal Behaviour, 59,* 763-770.

Beauchamp, M.S., Lee, K.E., Haxby, J.V., & Martin, A. (2003). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron, 34,* 149-159.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity - perceptual versus cognitive mechanisms in chimpanzees (Pan-Troglodytes). *Journal of Experimental Psychology-Animal Behavior Processes, 21,* 82-86.

Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity-based interference and symbolic representations in chimpanzees (Pan troglodytes). *Journal of Experimental Psychology-Animal Behavior Processes, 22,* 76-86.

Boysen, S. T., Berntson, G. G., & Mukobi, K. L. (2001). Size matters: Impact of item size and quantity on array choice by chimpanzees (Pan troglodytes). *Journal of Comparative Psychology, 115,* 106-110.

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.

Byrne, R.W., & Whiten, A. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Clarendon Press,

Call, J., & Tomasselo, M. (1998). Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology, 112,* 192-206.

Carlson, S. M., Moses, L. J., & Hix, H. R. [1998]. The role of inhibitory processes in young children's difficulties with deception and false belief. Child Development, 69(3), 672-691

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9,* 377-395.

Desimone, R., Albright, T.D., Gross, C.G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051-2062.

di Pellegrino, G., Fadiga, L., Fogassi, V., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91,* 176-180.

Dobbins, A.C., Jeo, R.M., Fiser, J., & Allman, J.M. (1998). Distance modulation of neural activity in the visual cortex. *Science, 281,* 552-555.

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.

Földiák, P. Xiao, D., Keysers, C. Edwards, R., & Perrett, D.I. (in press). Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Progress in Brain Research*.

Gallese, V., Fadiga, L, Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119,* 593-609.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2,* 493-501.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz and B. Hommel (eds) *Attention and Performance XIX*, 334-355. Oxford: O.U.P.

Gross, C.G., Rocha-Miranda, C.E., & Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35,* 96-111. Hare, B., Call, J. & Tomasello, M. 2001. Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61, 139-151.

Hare, B., Call, J., Agnetta, B. & Tomasello, M. 2000. Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59, 771-785

Hauser, M.D. (1999). Primate representations and expectations: mental tools for navigating in a social world. In P. Zelazo, J. Astington and D. Olson (Ed.), *Developing Theories of Intention: Social Understanding and Self-Control* (pp. 169-194). Hillsdale: Erlbaum.

Hauser, M. D. (2003). Knowing about knowing. Dissociations between perception and action systems over evolution and during development. *Annals of the New York Academy of Sciences*. In press.

Henzi, S.P., & Barrett, L. (1999). The value of grooming to female primates. *Primates, 40,* 47-59.

Hietanen, J.K., & Perrett, D.I. (1996). Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self- and externally generated pattern motion. *Behavioural Brain Research, 76,* 155-167.

Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.C., Mazziotta, J.C., & Rizzolatti, G. (2001). Reafferent copies of imitated actions in the right superior temporal cortex. *Proceedings of the National Academy of Sciences USA, 98,* 13995-

13999.

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.

Jellema, T., Baker, C.I., Oram, M.W., & Perrett, D.I. (2002). Cell populations in the banks of the superior temporal sulcus of the macaque and imitation. In W. Prinz and A. Meltzoff (eds) *The Imitative Mind. Development, Evolution, and Brain Bases*, pp. 267-290. Cambridge: Cambridge University Press.

Jellema, T., & Perrett, D.I. (2002). Coding of visible and hidden objects. *Common mechanisms in Perception and Action. Attention and Performance, 19,* 356-380.

Jellema, T., & Perrett, D.I. (2003a). Perceptual history influences neural responses to face and body postures. *Journal of Cognitive Neuroscience, 15***,** 961-971.

Jellema, T., & Perrett, D.I. (2003b). Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia, 41,* 1728-1737.

Jellema, T., Maassen G. & Perrett, D.I. (2004). Integration of animate motion, form, and location in the superior temporal sulcus of the macaque monkey. *Cerebral Cortex, 14,* 781-790.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science, 297,* 846-848.

Logothetis, N.K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys, *Current Biology, 5,* 552-563.

Lorincz, E.N., Baker, C.I., & Perrett, D.I. (1999). Visual cues for attention following in rhesus monkeys. *Current Psychology of Cognition, 18,* 973-1001.

Mistlin, A.J., & Perrett D.I. (1990). Visual and somatosensory processing in the macaque temporal cortex: the role of 'expectation'. *Experimental Brain Research, 82,* 437-450.

Milner A.D., & Goodale M.A. (1995). *The Visual Brain in Action*. Oxford: O.U.P.

Murata, A., Gallese, V., Kaseda M. & Sakata, H. (1996). Parietal neurons related to memory-guided hand manipulation. *Journal of Neurophysiology, 75,* 2180-2186.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the USA, 97,* 913-918.

Nishitani, N., & Hari, R. (2001). Sign language and mirror neuron system. *Neuroimage, 13,* S452

Oram, M.W., & Perrett D.I. (1994). Responses of anterior superior temporal polysensory (STPa) neurones to 'biological motion' stimuli. *Journal of Cognitive Neuroscience, 6,* 99-116.

Oram, M.W., & Perrett, D.I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology, 76,* 109-129.

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

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

Perrett, D.I., Smith, P.A.J., Mistlin, A.J., Chitty, A.J., Head, A.S., Potter, D.D., Broennimann, R., Milner, A.D., & Jeeves, M.A. (1985a). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: A preliminary report. *Behavioral Brain Research, 16,* 153-170.

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves M.A. (1985b). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London: Series B, 223,* 293-317.

Perrett, D.I., Harries, M.H., Bevan R., Thomas, S., Benson, P.J., Mistlin, A.J., Chitty, A.J., Hietanen, J.K., & Ortega, J.E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology, 146,* 87-113.

Perrett, D.I., Mistlin, A.J., Harries, M.H., & Chitty, A.J. (1990). Understanding the visual appearance and consequences of hand actions. In M.A. Goodale (Ed.), *Vision and Action: the control of grasping.* (pp. 163-180). Norwood, NJ: Ablex Publishing.

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, 335,* 23-30.

Perrett, D.I., & Oram, M.W. (1993). Neurophysiology of shape processing. *Image and Vision Computing, 11,* 317-333.

Puce, A., & Perrett, D.I. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London, B, 358,* 435-445.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Experimental Brain Research, 71,* 491-507.

Rizzolatti, G., & Gentilucci, M. (1988). Motor and visual-motor functions of the premotor cortex. In P. Rakic and W. Singer (eds). *Neurobiology of Neocortex, Vol 42*: John Wiley & Sons.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu E., Perani, D., & Fazio, F. (1996a). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research, 111,* 246-252.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996b). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3,* 131-141.

Rolls, E.T., & Baylis, G.C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research, 65,* 38-48.

Santos, L. R., & Hauser, M. D. (1999). How monkeys see the eyes: cotton-top tamarins' reaction to changes in visual attention and action. *Animal Cognition, 2,* 131-139.

Saxe, R., Xiao, D.-K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2003). Distinct representations of bodies, actions and thoughts in posterior superior temporal sulcus. (Submitted).

Sereno, A.B., & Maunsell, J.H.R. (1998). Shape selectivity in primate lateral intraparietal cortex. *Nature, 395,* 500-503.

Tanaka, Y.Z., Koyama, T., & Mikami, A. (1999). Neurons in the temporal cortex changed their preferred direction of motion dependent on shape. *Neuroreport, 10,* 393-397.

Tomasello, M., Call, J. & Hare, B. 1998. Five primate species follow the visual gaze of conspecifics. *Animal Behaviour, 55,* 1063-1069

Umilta, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron, 31,* 155-165.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A Goodale, R.J.W. Mansfield (eds) Analysis of Visual Behavior, pp. 549-586. Cambridge, MA: MIT Press.

Wachsmuth, E., Oram, M.W., & Perrett, D.I. (1994). Recognition of objects, & their component parts: responses of single units in the temporal cortex of the macaque, *Cerebral Cortex, 5,* 509-522.

Williams, J.H.G., Whiten, A., Suddendorf, T., & Perrett, D.I. (2001). Imitation, mirror neurons, & autism. *Neuroscience, & Behavioural Reviews, 25,* 287-295.