

Mirror neurons and the simulation theory of mind-reading

Vittorio Gallese and Alvin Goldman

A new class of visuomotor neuron has been recently discovered in the monkey's premotor cortex: mirror neurons. These neurons respond both when a particular action is performed by the recorded monkey and when the same action, performed by another individual, is observed. Mirror neurons appear to form a cortical system matching observation and execution of goal-related motor actions. Experimental evidence suggests that a similar matching system also exists in humans. What might be the functional role of this matching system? One possible function is to enable an organism to detect certain mental states of observed conspecifics. This function might be part of, or a precursor to, a more general mind-reading ability. Two different accounts of mind-reading have been suggested. According to 'theory theory', mental states are represented as inferred posits of a naive theory. According to 'simulation theory', other people's mental states are represented by adopting their perspective: by tracking or matching their states with resonant states of one's own. The activity of mirror neurons, and the fact that observers undergo motor facilitation in the same muscular groups as those utilized by target agents, are findings that accord well with simulation theory but would not be predicted by theory theory.

How do we understand other people's behavior? How can we assign goals, intentions, or beliefs to the inhabitants of our social world? A possible way to answer these challenging questions is to adopt an evolutionary frame of reference, both in phylogenetical and ontogenetical terms, envisaging these 'mind-reading' capacities as rooted in antecedent, more 'ancient' and simple mechanisms. This approach can capitalize on the results of different fields of investigation: neurophysiology can investigate the neural correlates of precursors of these mechanisms in lower species of social primates such as macaque monkeys. Developmental psychology can study how the capacity to attribute propositional attitudes to others develops.

In the present article we will propose that humans' mind-reading abilities rely on the capacity to adopt a simulation routine. This capacity might have evolved from an action execution/observation matching system whose neural correlate is represented by a class of neurons recently discovered in the macaque monkey premotor cortex: mirror neurons (MNs).

The macaque monkey premotor area F5 and mirror neurons

Converging anatomical evidence (see Matelli and Luppino¹ for review) supports the notion that the ventral premotor

cortex (referred to also as inferior area 6) is composed of two distinct areas, designated as F4 and F5 (Ref. 2) (Fig. 1A). Area F5 occupies the most rostral part of inferior area 6, extending rostrally within the posterior bank of the inferior limb of the arcuate sulcus. Area F5 is reciprocally connected with the hand field of the primary motor cortex³⁻⁵ and has direct, although limited, projections to the upper cervical segments of the spinal cord⁶. Microstimulation in F5 evokes hand and mouth movements at thresholds generally higher than in the primary motor cortex^{7,8}. The functional properties of F5 neurons were assessed in a series of single unit recording experiments⁹⁻¹¹. These experiments showed that the activity of F5 neurons is correlated with specific hand and mouth motor acts and not with the execution of individual movements like contractions of individual muscle groups. What makes a movement into a motor act is the presence of a goal. This distinction is very important since it allows one to interpret the role of the motor system not just in terms of the control of the dynamic variables of movement (like joint torques, etc.), but rather as a possible candidate for the instantiation of mental states such as purpose or intention. Using the effective motor act as the classification criterion, the following types of neurons were described: 'Grasping neurons', 'Holding neurons', 'Tearing neurons' and 'Manipulation neurons'. Grasping neurons discharge when

V. Gallese is at the
Istituto di Fisiologia
Umana, Università
di Parma, Via
Volturno 39,
I-43100 Parma,
Italy.

tel: +39 521 903879
fax: +39 521 903900
e-mail: fisioum@
symbolic.pr.it

A. Goldman is at the
Department of
Philosophy, University
of Arizona, Tucson,
AZ 85721-0027,
USA.

tel: +1 520 621 3120
fax: +1 520 621 9559
e-mail: goldman@u.
arizona.edu

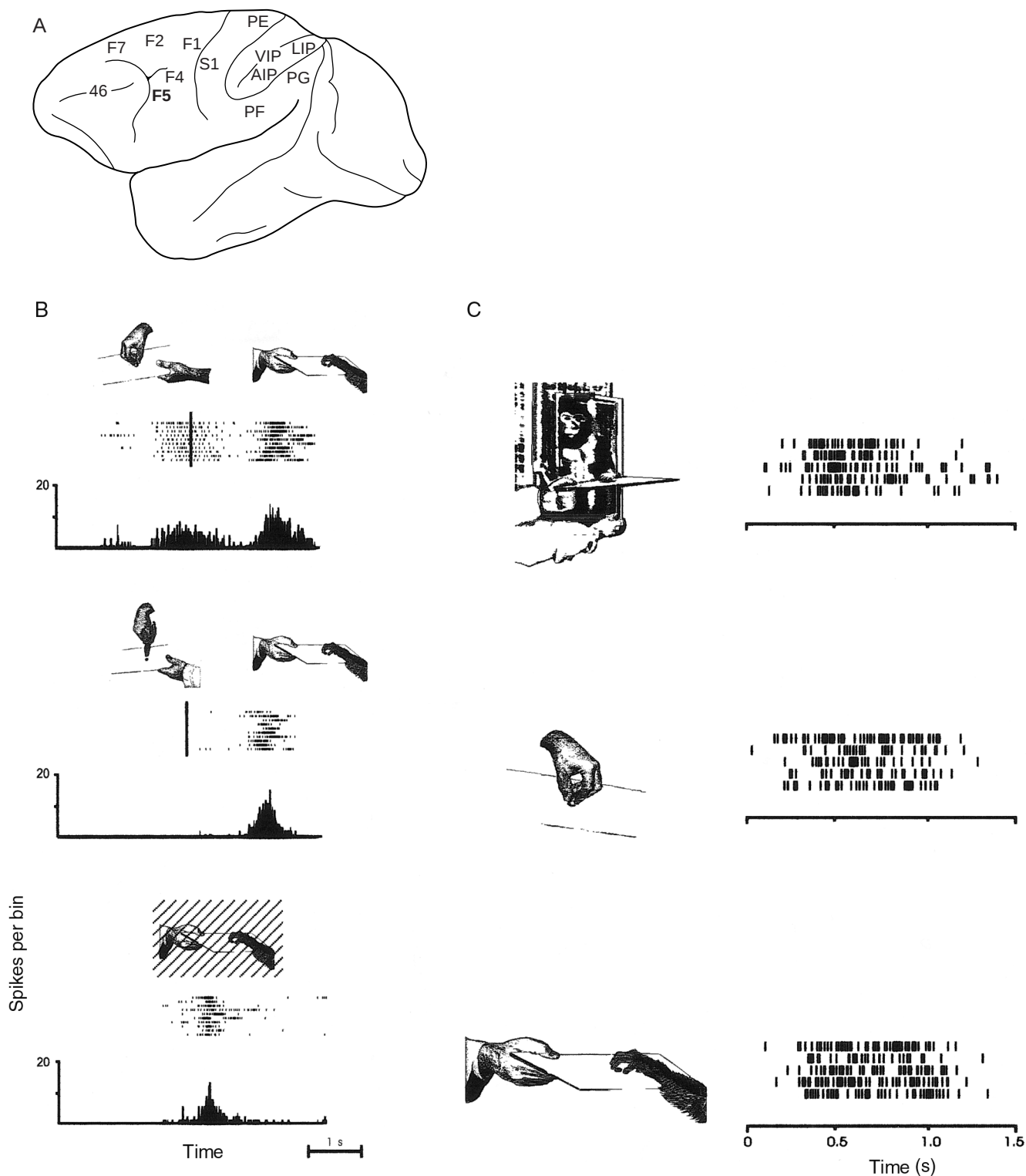


Fig. 1 Anatomical location and functional properties of mirror neurons. (A) Lateral view of the macaque brain showing the cytoarchitectonic parcellation of the agranular frontal cortex and of the posterior parietal cortex. Motor and premotor areas, indicated by the letter F, are defined according to Matelli *et al.*² Mirror neurons were all recorded from area F5 (shown in bold). (B) Visual and motor responses of a mirror neuron. In the upper part of each panel the behavioral context in which the neuron was studied is shown. In the lower part of each panel a series of consecutive rasters and the relative peristimulus response histograms are shown. In the upper panel the experimenter grasps a piece of food with his hand and moves it towards the monkey, who grasps it. The neuron discharges during grasping observation, is silent when the food is moved, and discharges again when the monkey grasps it. In the middle panel the experimenter grasps the food with a tool. Subsequent series of event as in the previous panel. During grasping observation the neuron is silent. In the lower panel the monkey grasps the food in complete darkness. In the upper and middle panels rasters and histograms are aligned (vertical bar) with the moment in which the experimenter grasps the food. In the lower panel alignment is with the beginning of the grasping movement. (Histograms bin width, 20 ms. Ordinates, spikes/bin. Abscissae, time.) (C) Visual and motor responses of a mirror neuron. In the upper panel the recorded monkey observes another monkey grasping food. In the middle panel the recorded monkey observes the experimenter grasping food. In the lower panel the recorded monkey actively grasps food. Each panel illustrates five consecutive trials. The spontaneous activity of the neuron was virtually absent. (Panels B and C are modified from Ref. 15.)

the monkey performs movements aimed to take possession of objects with the hand ('Grasping-with-the-hand neurons'), with the mouth ('Grasping-with-the-mouth neurons'), or with both. Grasping-with-the-hand neurons form the largest class of F5 neurons. Most neurons of this class are selective for different types of grip. The role of these neurons has been conceptualized by Rizzolatti¹² as a 'motor vocabulary' of actions related to prehension.

The study of F5 neurons' responsiveness to visual stimuli led to the discovery of two distinct classes of neurons: canonical neurons¹³, which are activated during observation of graspable objects, and MNs (see Refs 14,15) which discharge when the monkey observes another individual performing an action. We will describe in more detail the functional properties of this class of neurons. Figure 1B and C illustrate two examples of the activity of MNs.

MNs respond both when a particular action is performed by the recorded monkey and when the same action performed by another individual is observed. All MNs, as mentioned above, discharge during specific goal-related motor acts. Grasping, manipulating and holding objects are by far the most effective actions triggering their motor response. About half of them discharge during a specific type of prehension, precision grip (prehension of small objects by opposing the thumb and the index finger) being the most common one. The most effective visual stimuli triggering MNs' visual responses are actions in which the experimenter (Fig. 1B), or a second monkey (Fig. 1C), interacts with objects with their hand or with their mouth. Neither the sight of the object alone nor of the agent alone is effective in evoking the neuronal response. Mimicking the action without a target object, or performing the action by using tools (middle panel of Fig. 1B) is similarly ineffective. In over 90 percent of MNs a clear correlation between the most effective observed action and their motor response was observed. In many neurons this correlation was strict both in terms of the general goal of the action (e.g. grasping) and in terms of the way in which it was executed (e.g. precision grip)^{14,15}.

On the basis of their functional properties, here summarized, MNs appear to form a cortical system that matches observation and execution of motor actions. What could be the possible functional role of this matching system? Before addressing this issue it is important to stress that the existence of an equivalent system has also been demonstrated in humans.

The mirror system in humans

Two lines of evidence strongly suggest that an action/observation matching system similar to that discovered in monkeys also exists in humans. The first refers to an elegant study by Fadiga *et al.*¹⁶ in which the excitability of the motor cortex of normal human subjects was tested by using Transcranial Magnetic Stimulation (TMS). The basic assumption underlying this experiment was the following. If the observation of actions activates the premotor cortex in humans, as it does in monkeys, this mirror effect should elicit an enhancement of the motor evoked potentials (MEPs) induced by TMS of the motor cortex, given its strong anatomical links to premotor areas. TMS was performed

during four different conditions: observation of an experimenter grasping objects; observation of an experimenter doing aimless movements in the air with his arm; observation of objects; detection of the dimming of a small spot of light. The results of this study showed that during grasping observation MEPs recorded from the hand muscles markedly increased with respect to the other conditions, including the attention-demanding dimming detection task. Even more intriguing was the finding that the increase of excitability was present only in those muscles that subjects would use when actively performing the observed movements. This study provided for the first time evidence that humans have a mirror system similar to that in monkeys. Every time we are looking at someone performing an action, the same motor circuits that are recruited when we ourselves perform that action are concurrently activated.

These results posed the question of the anatomical location of the mirror system within the human brain. This issue has been addressed by two brain-imaging experiments utilizing the technique of Positron Emission Tomography (PET) (Refs 17,18). These two experiments, although different in many respects, shared a condition in which normal human subjects observed the experimenter grasping 3-D objects. Both studies used the observation of objects as a control condition. The results showed that grasping observation significantly activates the cortex of the left superior temporal sulcus (Brodmann's area 21), of the left inferior parietal lobule (Brodmann's area 40) and of the anterior part of Broca's region (Brodmann's area 45). The activation, during action observation, of a cortical sector of the human brain traditionally linked with language raises the problem of the possible homologies between Broca's region and the premotor area F5 of the monkey, in which MNs have been discovered. This issue is outside the scope of the present article and will not be dealt with here (for discussion, see Ref. 19).

Mirror neurons and mind-reading

What is the function of the mirror system? One possible function could be to promote learning by imitation. When new motor skills are learned, one often spends the first training phases trying to replicate the movements of an observed instructor. MNs could in principle facilitate that kind of learning. We do not favor this possible role of MNs, at least in non-human primates (see Box 1). Here we explore another possibility: that MNs underlie the process of 'mind-reading', or serve as precursors to such a process.

Mind-reading is the activity of representing specific mental states of others, for example, their perceptions, goals, beliefs, expectations, and the like. It is now agreed that all normal humans develop the capacity to represent mental states in others, a system of representation often called folk psychology. Whether non-human primates also deploy folk psychology is more controversial (see last section of this article), but it certainly has not been precluded. The hypothesis explored here is that MNs are part of – albeit perhaps a rudimentary part of – the folk psychologizing mechanism.

Like imitation learning, mind-reading could make a contribution to inclusive fitness. Detecting another agent's

Box 1. Mirror neurons and imitation

The ability of non-human primates to imitate the behavior of conspecifics is a highly controversial issue. Tomasello *et al.*^a identify three strict criteria to delimit imitational learning: (1) the imitated behavior should be novel for the imitator; (2) it should reproduce the behavioral strategies of the model; (3) it should share with it the same final goal. Behaviors not satisfying these criteria should not be considered as true imitational ones, and are rather to be explained by means of other mechanisms such as stimulus enhancement, emulation, or response facilitation. By applying these strict criteria to the extant literature, Tomasello *et al.*^a exclude the possibility that wild animals may display true imitative behavior. A different perspective is offered by Byrne and Russon^b. These authors start from the concept that behaviors display a hierarchical structure, and can be therefore described at several levels of increasing complexity. Manual skills represents a good example. Because complex behaviors are hierarchically structured, '...there exists a range of possibilities for how imitation might take place, beyond the simple dichotomy of imitation versus no imitation'. Byrne and Russon^b single out an action-level imitation in which a detailed specification of the various motor sequences composing a complex action is made, and a program-level imitation in

which the broader, more highly structured component of a complex skill is retained, with subjective solutions to the low-level specifications. Byrne and Russon^b conclude that imitational learning in non-human primates might have been overlooked by the exclusive application of the action-level strategy as the defining criterion.

What is the relevance of MNs for imitation in non-human primates? First of all, it should be stressed that imitation behavior has never been observed in association with MN activity. Furthermore, even adopting Byrne and Russon's criteria, we are not aware of any clearcut evidence of imitation of grasping behavior among adult macaque monkeys, although this possibility is not precluded for young monkeys during development. On the basis of these considerations we are inclined not to favor the hypothesis that MNs in area F5 promote grasping imitation learning in adult monkeys.

References

- a Tomasello, M., Kruger, A.C. and Ratner, H.H. (1993) Cultural learning *Behav. Brain Sci.* 16, 495–511
- b Byrne, R.W. and Russon, A.E. Learning by imitation: a hierarchical approach *Behav. Brain Sci.* (in press)

goals and/or inner states can be useful to an observer because it helps him anticipate the agent's future actions, which might be cooperative, non-cooperative, or even threatening. Accurate understanding and anticipation enable the observer to adjust his responses appropriately. Our discussion of mind-reading will initially and primarily focus on humans; later we will return to its possible realization in non-human primates.

Two theories of mind-reading

There is a large literature concerned with the nature of (human) mind-reading. Two types of approaches have dominated recent discussion: theory theory (TT) and simulation theory (ST) (Refs 20–22). The fundamental idea of TT is that ordinary people accomplish mind-reading by acquiring and deploying a commonsense theory of the mind, something akin to a scientific theory. Mental states attributed to other people are conceived of as unobservable, theoretical posits, invoked to explain and predict behavior in the same fashion that physicists appeal to electrons and quarks to predict and explain observable phenomena. On the standard presentation, the theory of mind possessed by ordinary people consists of a set of causal/explanatory laws that relate external stimuli to certain inner states (e.g. perceptions), certain inner states (e.g. desires and beliefs) to other inner states (e.g. decisions), and certain inner states (e.g. decisions) to behavior. This picture has been articulated by functionalist philosophers of mind^{23–26} as well as by developmental psychologists^{27,28}. According to TT, attributing particular mental states to others arises from theoretical reasoning involving tacitly known causal laws.

Much on this subject has been done by developmentalists, eager to determine how the mind-reading capacity is acquired in childhood²⁹. Many interpret children's changes in mind-reading skills as evidence in favor of TT because

the skill changes are construed as manifestations of changes in theory^{30,31}. Theory theorists differ among themselves as to whether theory of mind is acquired by a general-purpose scientizing algorithm³² or by the maturation of a domain-specific module or set of modules^{33,34}. This debate will not concern us here.

ST arose partly from doubts about whether folk psychologists really represent, even tacitly, the sorts of causal/explanatory laws that TT typically posits. ST suggests that attributors use their own mental mechanisms to calculate and predict the mental processes of others. For example, Kahneman and Tversky³⁵ gave subjects a description of two travellers who shared the same limousine en route to the airport and were caught in a traffic jam. Their planes were scheduled to depart at the same time, but they arrived 30 minutes late. Mr A was told that his flight left on time; Mr B was told that his flight was delayed and just left five minutes ago. Who was more upset? Ninety-six percent of the experimental subjects said that Mr B was more upset. How did they arrive at this answer? According to TT there must be some psychological law they exploited to infer the travellers' relative upsetness. According to ST, on the other hand, each subject would have put himself in each of the imaginary traveller's 'shoes' and imagined how he would have felt in their place³⁶. Another example concerns the prediction of decisions. To predict White's next move in a chess match ST suggests that you try to simulate White's thought processes and arrive at a decision which you then attribute to him^{36–38}. First you create in yourself *pretend* desires, preferences, and beliefs of the sort you take White to have; for example, preferences among chess strategies. These pretend preferences and beliefs are fed into your own decision-making mechanism, which outputs a (pretend) decision (see Fig. 2). Instead of acting on that decision, it is taken 'off-line' and used to predict White's decision.

According to this simulation account, you need not know or utilize any psychological laws.

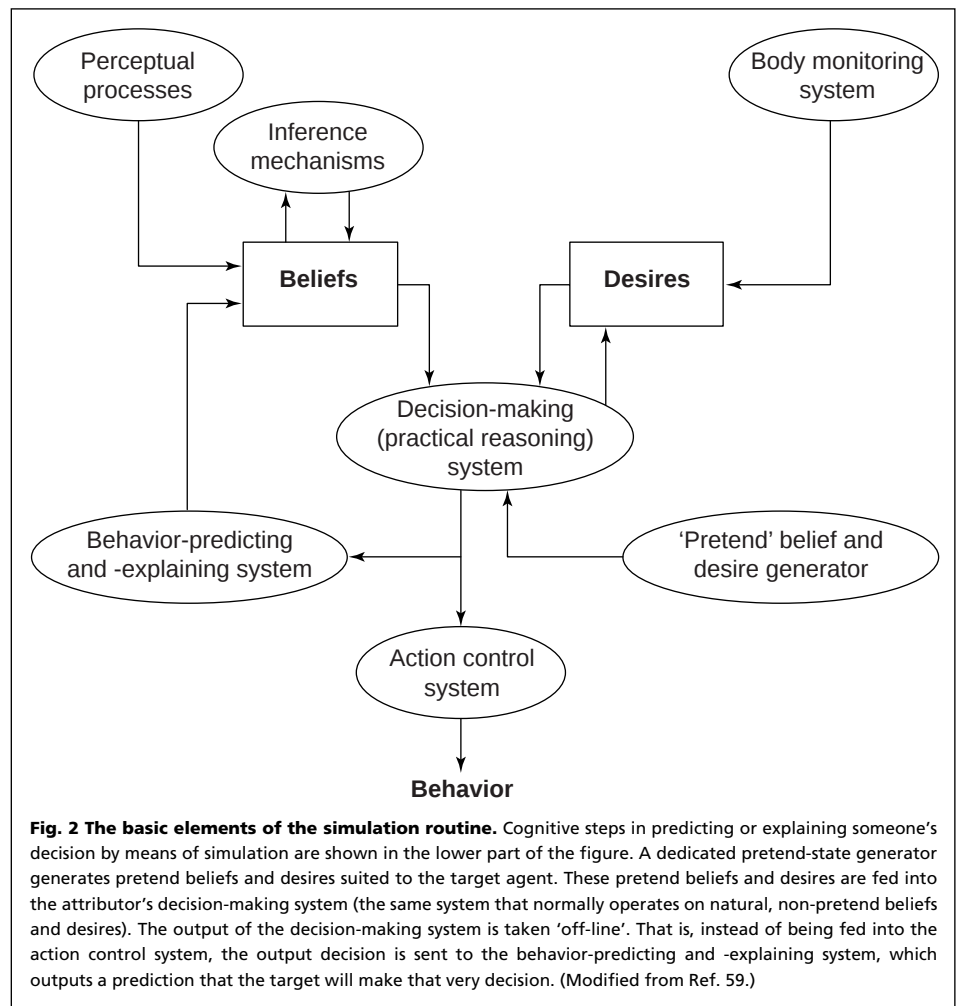
If simulation is going to make accurate predictions of targets' decisions, pretend desires and beliefs must be sufficiently similar to genuine desires and beliefs that the decision-making system operates on them the same way as it operates on genuine desires and beliefs. Are pretend states really similar enough to the genuine articles that this will happen? Homologies between pretend and natural (i.e. non-pretend) mental states are well documented in the domains of visual and motor imagery^{39–43}. (We assume here that visual and motor imaging consist, respectively, in pretending to see and pretending to do; see Currie and Ravenscroft⁴⁴.) These visual and motor homologies do not show, of course, that other pretend mental states, for example, desires and beliefs, also functionally resemble their natural counterparts, but informal evidence suggests this (see Goldman⁴⁵).

The difference between TT and ST

The core difference between TT and ST, in our view, is that TT depicts mind-reading as a thoroughly 'detached' theoretical activity, whereas ST depicts mind-reading as incorporating an attempt to replicate, mimic, or impersonate the mental life of the target agent⁴⁶. This difference can be highlighted diagrammatically, as shown in Fig. 3.

In the simulation scenario there is a distinctive matching or 'correspondence' between the mental activity of the simulator and the target. This is highlighted by the similar state-sequences the two undergo (Fig. 3, A and B), the only exception being that the simulator uses pretend states rather than natural states. The attributor in the TT scenario (Fig. 3C) does not utilize any pretend states that mimic those of the target; nor does he utilize his own decision-making system to arrive at a prediction. Thus, ST hypothesizes that a significant portion of mind-reading episodes involves the process of mimicking (or trying to mimic) the mental activity of the target agent. TT predicts no such mimicking as part of the mind-reading process. This contrast presents a potential basis for empirically discriminating between ST and TT. If there is evidence of mental mimicry in the mind-reading process, that would comport nicely with ST and would not be predicted by TT.

Before turning to such evidence, however, we should note that simulation can be used to retrodict as well as predict mental states, that is, to determine what mental states of a target have *already* occurred. Figure 4 depicts a retrodictive use of simulation. The attributor starts with the question, 'What goal did the target have that led him to perform action *m*?' He conjectures that it was goal *g*, and tries



out this conjecture by pretending to have *g* as well as certain beliefs about the effectiveness or ineffectiveness of the action *m vis-à-vis* goal *g*. This simulation leads him to form a (pretend) decision to do *m*. He therefore uses this result to conclude that the target did indeed have goal *g*. In this fashion, the attributor ultimately makes a 'backward' inference from the observed action to a hypothesized goal state.

Mirror neurons and simulation

In a similar fashion, it is conceivable that externally-generated MN activity serves the purpose of 'retrodicting' the target's mental state, moving backwards from the observed action. Let us interpret internally generated activation in MNs as constituting a plan to execute a certain action, for example, the action of holding a certain object, grasping it, or manipulating it. When the same MNs are externally activated – by observing a target agent execute the same action – MN activation still constitutes a plan to execute this action. But in the latter case the subject of the MN activity knows (visually) that the observed target is concurrently performing this very action. So we assume that he 'tags' the plan in question as belonging to that target. In fact, externally generated MN activity does not normally produce motor execution of the plan in question. Externally generated plans are largely inhibited, or taken 'off-line', precisely as ST postulates. Thus MN activity seems to be nature's way of getting the observer into the same 'mental shoes' as

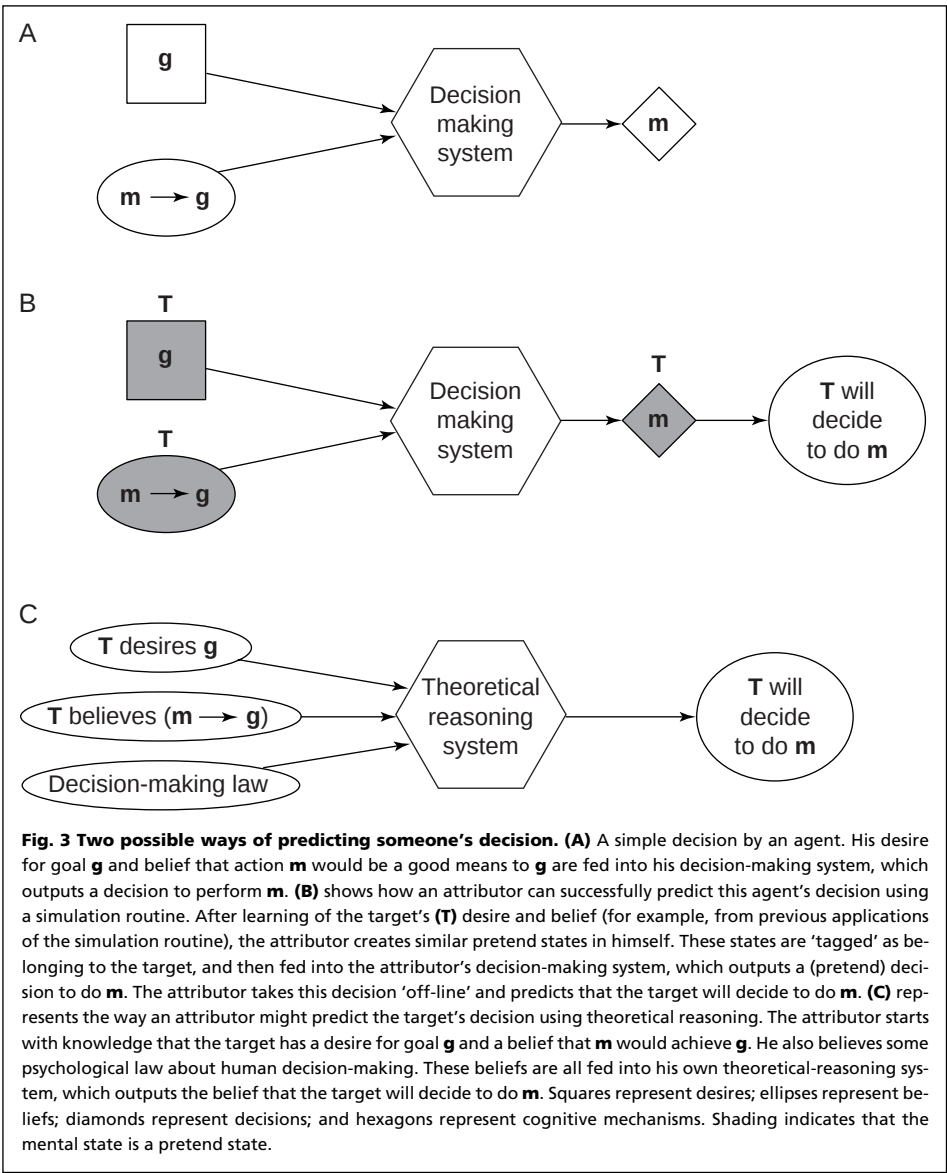


Fig. 3 Two possible ways of predicting someone's decision. (A) A simple decision by an agent. His desire for goal *g* and belief that action *m* would be a good means to *g* are fed into his decision-making system, which outputs a decision to perform *m*. **(B)** shows how an attributor can successfully predict this agent's decision using a simulation routine. After learning of the target's (*T*) desire and belief (for example, from previous applications of the simulation routine), the attributor creates similar pretend states in himself. These states are 'tagged' as belonging to the target, and then fed into the attributor's decision-making system, which outputs a (pretend) decision to do *m*. The attributor takes this decision 'off-line' and predicts that the target will decide to do *m*. **(C)** represents the way an attributor might predict the target's decision using theoretical reasoning. The attributor starts with knowledge that the target has a desire for goal *g* and a belief that *m* would achieve *g*. He also believes some psychological law about human decision-making. These beliefs are all fed into his own theoretical-reasoning system, which outputs the belief that the target will decide to do *m*. Squares represent desires; ellipses represent beliefs; diamonds represent decisions; and hexagons represent cognitive mechanisms. Shading indicates that the mental state is a pretend state.

the target – exactly what the conjectured simulation heuristic aims to do.

Although we compare externally generated MN activity with what transpires in Fig. 4, there clearly are differences. One difference is that the real attributor does not go back to a distal goal or set of beliefs. He only goes back to a motoric plan. Still, this seems to be a 'primitive' use of simulation with the same structure as that depicted in Fig. 4. It also bears a resemblance to the motor theory of speech percep-

agent. It is as if the tracking process in the observer is not taken entirely off-line. This might appear to be a violation of ST, but actually it is wholly within ST's spirit. ST postulates mental occurrences in the mind-reader that are analogous to mental occurrences in the target, so it is not surprising that downstream motor activity is not entirely inhibited. If TT were correct, and an observer represents a target's behavior in purely theoretical fashion, it would not be predicted that the same muscle groups would be facilitated in the observer as in the target. But if ST were correct, and a mind-reader represents an actor's behavior by recreating in himself the plans or movement intentions of the actor, then it is reasonable to predict that the same muscular activation will occur in the mind-reader. As matching muscular activation is actually observed in the observer, this lends support to ST as opposed to TT.

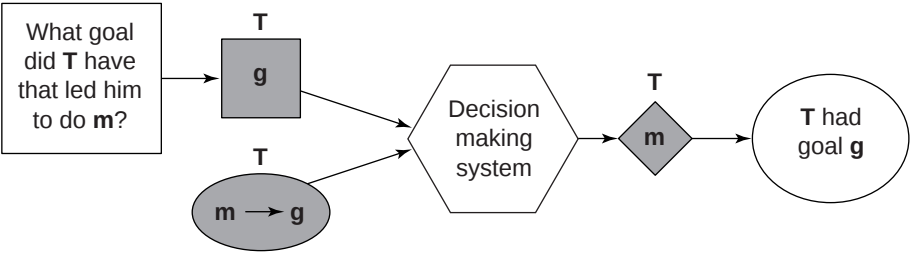


Fig. 4 A retrodictive use of simulation. After observing the target agent (*T*) perform action *m*, the attributor uses simulation to test whether goal *g* would have fitted with the choice of *m*. Goal *g* is re-created and fed into his decision-making system, which does output *m*.

tion advocated by Liberman⁴⁷, in which the common link between the sender and the receiver is not sound but the neural mechanism, shared by both, allowing the production of phonetic gestures.

A proponent of TT might say that TT also has ways of accounting for retrodictive attributions of mental states. Is it clear that anything similar to simulation occurs in externally generated MN activity? The point is that MN activity is not mere theoretical inference. It creates in the observer a state that matches that of the target. This is how it resembles the simulation heuristic. Nothing about TT leads us to expect this kind of matching. It should be emphasized that the hypothesis being advanced here is not that MNs themselves constitute a full-scale realization of the simulation heuristic. In particular, we do not make this conjecture for MNs in monkeys. Our conjecture is only that MNs represent a primitive version, or possibly a precursor in phylogeny, of a simulation heuristic that might underlie mind-reading.

A further link between mirror neuron activity and simulation can be inferred from the fact that, as the TMS experiment by Fadiga *et al.*¹⁶ demonstrates, the human equivalent matching system facilitates in the observer the same muscle groups as those utilized by the target. This supports the idea that even when one is observing the action of another, one undergoes a neural event that is qualitatively the same as an event that triggers actual movement in the observed agent. It is as if the tracking process in the observer is not taken entirely off-line. This might appear to be a violation of ST, but actually it is wholly within ST's spirit. ST postulates mental occurrences in the mind-reader that are analogous to mental occurrences in the target, so it is not surprising that downstream motor activity is not entirely inhibited. If TT were correct, and an observer represents a target's behavior in purely theoretical fashion, it would not be predicted that the same muscle groups would be facilitated in the observer as in the target. But if ST were correct, and a mind-reader represents an actor's behavior by recreating in himself the plans or movement intentions of the actor, then it is reasonable to predict that the same muscular activation will occur in the mind-reader. As matching muscular activation is actually observed in the observer, this lends support to ST as opposed to TT.

Clinical evidence of a similar phenomenon is found in so-called 'imitation behavior'⁴⁸. A group of patients with

Box 2. Neural coding of complex biological stimuli

Neurons responding to complex biological stimuli have been previously described in the macaque brain. A series of studies showed that in the inferior temporal cortex there are neurons that discharge selectively to the presentation of faces or hands^{a-c}. More recently it has been shown that some of these neurons respond to specific features of these stimuli^d. Neurons responding to complex biological visual stimuli such as walking or climbing were reported also in the amygdala^e. Even more relevant to the issues addressed in the present paper is the work of Perrett and co-workers^{b,f,g}. These authors showed that in the cortex buried within the superior temporal sulcus (STS) there are neurons selective to the observation of hand movements. These properties resemble the visual properties of F5 MNs very closely: both populations of neurons code the same types of actions; they both generalize their responses to the different instances of the same action; they both are not responsive to mimicked hand actions without the target object. However, the unique feature of F5 MNs resides in the fact that they also discharge during active movements of the observer. An observed action produces the same neural pattern of activation as does the action actively made by the observer.

The presence of two brain regions with neurons endowed with similar complex visual properties raises the question of their possible relationship. Two possibilities might be suggested. One is that F5 MNs and STS neurons have different functional roles: STS neurons would code the semantic properties, the meaning, of hand-object interactions, while F5 MNs would be engaged in the pragmatic coding of the same actions. A second

possibility, that we favor, is that these two 'action detector' systems could represent distinct stages of the same analysis. The STS neurons would provide an initial 'pictorial' description of actions that would be then fed (most likely through an intermediate step in the posterior parietal cortex) to the F5 motor vocabulary where it would acquire a meaning for the individual. The latter hypothesis stresses the role of action in providing meaning to what is perceived.

References

- a Gross, C.G. *et al.* (1972) Visual properties of neurons in inferotemporal cortex of the monkey *J. Neurophysiol.* 35, 96–111
- b Perrett, D.I., Rolls, E.T. and Caan, W. (1982) Visual neurones responsive to faces in the monkey temporal cortex *Exp. Brain Res.* 47, 329–342
- c Gross, C.G. *et al.* (1985) Inferior temporal cortex and pattern recognition, in *Pattern Recognition Mechanisms* (Chagas, C., Gattass, R. and Gross, C., eds), Springer-Verlag
- d Tanaka, K. *et al.* (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey *J. Neurophysiol.* 66, 170–189
- e Brothers, L., Ring, B. and Kling, A. (1990) Response of neurons in the macaque amygdala to complex social stimuli *Behav. Brain Res.* 41, 199–213
- f Perrett, D.I. (1989) Frameworks of analysis for the neural representation of animate objects and actions *J. Exp. Biol.* 146, 87–113
- g Perrett, D.I. (1990) Understanding the visual appearance and consequence of hand actions, in *Vision and Action: the Control of Grasping* (Goodale, M.A., ed.), pp. 163–180, Ablex

prefrontal lesions compulsively imitate gestures or even complex actions performed in front of them by an experimenter. This behavior is explained as an impairment of the inhibitory control normally governing motor schemas, or plans. It may be inferred from this that normal humans, when observing someone else perform an action, generate a plan to do the same action, or an image of doing it, themselves. Normally this plan is inhibited so that it does not yield motor output, but such inhibition is impaired in the patient population in question⁴⁸.

Non-human primates: behaviorists or mind-readers?

A mind-reading capacity for non-human primates is a hotly debated issue among primatologists and behavioral scientists. In a recent paper Heyes⁴⁹ argued that a survey of empirical studies of imitation, self-recognition, social relationships, deception, role-taking and perspective-taking fails to support the theory of mind hypothesis over non-mentalist alternatives. Although, for sake of concision, it is not possible here to address this issue thoroughly (for reviews see Refs 49 and 50), a few points are worth making.

Let us consider first the social nature of non-human primates. Social organization is by no means a distinctive feature of primates: within the realm of insects several species (ants are one example) are endowed with a clear social structure. The distinctive hallmark of the social organization of non-human primates is its sophisticated complexity. Non-human primates live in groups that can comprise as many as 100 individuals. These groups are characterized by intense

and diversified types of social interactions⁵¹. Within such a complex and hierarchically organized social structure, individuals are able to recognize kinship, hierarchical ranks, to discriminate allies from enemies. Stammbach⁵² showed that dominant macaque monkeys modified their social relationships with lower-ranking individuals who had previously learned how to retrieve food by pressing a lever. Dominant individuals started grooming the low-ranking ones more often than before, once they 'understood' that the newly acquired skills of the low-ranking individuals could be more easily triggered, and therefore exploited, by using this sort of social upgrading. All these examples, although not providing conclusive evidence of mind-reading abilities, nevertheless, in our view, provide a strong argument supporting the hypothesis that non-human primates are endowed with cognitive abilities that cannot be easily dismissed as the result of simple stimulus–response operant conditioning.

Being a 'cognizer', nevertheless, does not necessarily imply being a mind-reader, or a possessor of the ability to detect intentional states in others. The argument that seems to suggest the presence, in non-human primates, of elementary forms of mind-reading abilities comes from the discovery of deceptive behavior. In a series of field experiments, Hauser^{53,54} showed that rhesus monkeys can withhold information about food location in order to deceive conspecifics and obtain more food for themselves. Deception is particularly relevant here, since deceptive behavior calls for the existence of second-order intentionality, and therefore for the capability to attribute mental states to conspecifics.

Outstanding questions

- Is mirror-neuron activity innate or learned, and what is the relevance of this to the simulation-theory interpretation of mirror-neuron activity?
- Is the motor system involved in the semantic mode of internally coding actions?
- Can any evidence be found of 'matching' events for observed agent's beliefs (as well as their plans or intentions) in non-human primates?

The relevance of the data on deceptive behavior has been questioned on the basis of two main arguments. First, field reports of ethologists are anecdotal and therefore intrinsically ambiguous. Second, alternative non-mentalistic explanations, such as chance behavior, associative learning, and inferences about observable features of the situation have been proposed as more parsimonious explanations of deceptive behavior (see Heyes⁴⁹).

However, according to Byrne⁵⁵, who surveyed the literature thoroughly, there are at least 18 independent reports of intentional deception in non-human primates supporting the notion that they can represent the mental states of other conspecifics. On the basis of this evidence, Byrne and Whiten⁵⁶ suggested that primates act according to a manipulative strategy very similar to that put forward in the sixteenth century by Niccolò Machiavelli in his masterpiece *Il Principe*⁵⁷.

Our speculative suggestion is that a 'cognitive continuity' exists within the domain of intentional-state attribution from non-human primates to humans, and that MNs represent its neural correlate (see also Box 2). This continuity is grounded in the ability of both human and non-human primates to detect goals in the observed behavior of conspecifics. The capacity to understand action goals, already present in non-human primates, relies on a process that matches the observed behavior to the action plans of the observer. It is true, as pointed out by Meltzoff and Moore⁵⁸, that the understanding of action goals does not imply a full grasp of mental states such as beliefs or desires. Action-goal understanding nevertheless constitutes a necessary phylogenetical stage within the evolutionary path leading to the fully developed mind-reading abilities of human beings.

Acknowledgements

This work was supported by a Human Frontier Scientific Program grant to V.G. The authors wish to thank Giacomo Rizzolatti, Elisabeth Pacherie and the anonymous referees for their valuable comments and criticisms.

References

- 1 Matelli, M. and Luppino, G. (1997) Functional anatomy of human motor cortical areas, in *Handbook of Neuropsychology* (Vol. 11) (Boller, F. and Grafman, J., eds), pp. 9–26, Elsevier
- 2 Matelli, M., Luppino, G. and Rizzolatti, G. (1985) Patterns of cytochrome oxidase activity in the frontal agranular cortex of macaque monkey *Behav. Brain Res.* 18, 125–137
- 3 Matsumura, M. and Kubota, K. (1979) Cortical projection of hand-arm motor area from postarcuate area in macaque monkey: a histological study of retrograde transport of horseradish peroxidase *Neurosci. Lett.* 11, 241–246
- 4 Muakkassa, K.F. and Strick, P.L. (1979) Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized 'premotor' areas *Brain Res.* 177, 176–182
- 5 Matelli, M. et al. (1986) Afferent and efferent projections of the

- inferior area 6 in the macaque monkey *J. Comp. Neurol.* 251, 281–298
- 6 He, S.Q., Dum, R.P. and Strick, P.L. (1993) Topographic organization of corticospinal projections from the frontal lobe: motor areas on the lateral surface of the hemisphere *J. Neurosci.* 13, 952–980
- 7 Gentilucci, M. et al. (1988) Functional organization of inferior area 6 in the macaque monkey: I. Somatotopy and the control of proximal movements *Exp. Brain Res.* 71, 475–490
- 8 Hepp-Reymond, M.-C. et al. (1994) Force-related neuronal activity in two regions of the primate ventral premotor cortex *Can. J. Physiol. Pharmacol.* 72, 571–579
- 9 Rizzolatti, G. et al. (1981) Afferent properties of periacuate neurons in macaque monkey. II. Visual responses *Behav. Brain Res.* 2, 147–163
- 10 Okano, K. and Tanji, J. (1987) Neuronal activities in the primate motor fields of the agranular frontal cortex preceding visually triggered and self-paced movement *Exp. Brain Res.* 66, 155–166
- 11 Rizzolatti, G. et al. (1988) Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements *Exp. Brain Res.* 71, 491–507
- 12 Rizzolatti, G. and Gentilucci, M. (1988) Motor and visual-motor functions of the premotor cortex, in *Neurobiology of Neocortex* (Rakic, P. and Singer, W., eds), pp. 269–284, John Wiley & Sons
- 13 Murata, A. et al. (1997) Object representation in the ventral premotor cortex (Area F5) of the monkey *J. Neurophysiol.* 78, 2226–2230
- 14 Gallese, V. et al. (1996) Action recognition in the premotor cortex *Brain* 119, 593–609
- 15 Rizzolatti, G. et al. (1996) Premotor cortex and the recognition of motor actions *Cognit. Brain Res.* 3, 131–141
- 16 Fadiga, L. et al. (1995) Motor facilitation during action observation: a magnetic stimulation study *J. Neurophysiol.* 73, 2608–2611
- 17 Rizzolatti, G. et al. (1996) Localization of grasp representations in humans by PET: 1. Observation versus execution *Exp. Brain Res.* 111, 246–252
- 18 Grafton, S.T. et al. (1996) Localization of grasp representations in humans by PET: II. Observation compared with imagination *Exp. Brain Res.* 112, 103–111
- 19 Rizzolatti, G. and Arbib, M.A. (1998) Language within our grasp *Trends Neurosci.* 21, 188–194
- 20 Davies, M. and Stone, T., eds (1995) *Folk Psychology*, Blackwell
- 21 Davies, M. and Stone, T., eds (1995) *Mental Simulation*, Blackwell
- 22 Carruthers, P. and Smith, P., eds (1996) *Theories of Theories of Mind*, Cambridge University Press
- 23 Sellars, W. (1963) Empiricism and the philosophy of mind, in *Science, Perception and Reality*, pp. 127–194, Routledge
- 24 Lewis, D. (1972) Psychophysical and theoretical identifications *Australas. J. Philos.* 50, 249–258
- 25 Fodor, J.A. (1987) *Psychosemantics*, MIT Press
- 26 Churchland, P.N. (1988) *Matter and Consciousness*, MIT Press
- 27 Wellman, H. (1990) *The Child's Theory of Mind*, MIT Press
- 28 Gopnik, A. (1993) How we know our minds: the illusion of first-person knowledge of intentionality *Behav. Brain Sci.* 16, 1–14
- 29 Astington, J., Harris, P. and Olson, D., eds (1988) *Developing Theories of Mind*, Cambridge University Press
- 30 Perner, J. (1991) *Understanding the Representational Mind*, MIT Press
- 31 Gopnik, A. and Wellman, H. (1992) Why the child's theory of mind really is a theory *Mind Lang.* 7, 145–171
- 32 Gopnik, A. and Meltzoff, A. (1997) *Words, Thoughts, and Theories*, MIT Press
- 33 Leslie, A. (1994) Pretending and believing: issues in the theory of TOMM *Cognition* 50, 211–238
- 34 Baron-Cohen, S. (1995) *Mindblindness*, MIT Press
- 35 Kahneman, D. and Tversky, A. (1982) The simulation heuristic, in *Judgment Under Uncertainty* (Kahneman, D., Slovic, P. and Tversky, A., eds), pp. 201–208, Cambridge University Press
- 36 Goldman, A. (1989) Interpretation psychologized *Mind Lang.* 4, 161–185
- 37 Gordon, R. (1986) Folk psychology as simulation *Mind Lang.* 1, 158–171
- 38 Heal, J. (1986) Replication and functionalism, in *Language, Mind and Logic* (Butterfield, J., ed.), pp. 135–150, Cambridge University Press
- 39 Kosslyn, S. (1978) Measuring the visual angle of the mind's eye *Cognit. Psychol.* 10, 356–389
- 40 Decety, J. et al. (1991) Vegetative response during imagined movement is proportional to mental effort *Behav. Brain Res.* 42, 1–5
- 41 Farah, M., Soso, M. and Dasheiff, R. (1992) Visual angle of the mind's

- eye before and after unilateral/occipital lobectomy *J. Exp. Psychol. Hum. Percept. Perform.* 18, 241–246
- 42 Yue, G. and Cole, K. (1992) Strength increases from the motor program: comparison of training with maximal voluntary and imagined muscle contractions *J. Neurophysiol.* 67, 1114–1123
- 43 Jeannerod, M. (1994) The representing brain: neural correlates of motor intention and imagery *Behav. Brain Sci.* 17, 187–202
- 44 Currie, G. and Ravenscroft, I. (1997) Mental simulation and motor imagery *Philos. Sci.* 64, 161–180
- 45 Goldman, A. (1992) In defense of the simulation theory *Mind Lang.* 7, 104–119
- 46 Goldman, A. The mentalizing folk, in *Metarepresentation* (Sperber, D., ed.), Oxford University Press (in press)
- 47 Liberman, A. (1996) *Speech: A Special Code*, MIT Press
- 48 Lhermitte, F., Pillon, B. and Serdaru, M. (1986) Human autonomy and the frontal lobes: I. Imitation and utilization behavior: a neuropsychological study of 75 patients *Ann. Neurol.* 19, 326–334
- 49 Heyes, C.M. (1998) Theory of mind in nonhuman primates *Behav. Brain Sci.* 21, 101–148
- 50 Barresi, J. and Moore, C. (1996) Intentional relations and social understanding *Behav. Brain Sci.* 19, 107–154
- 51 Cheney, D.L. and Seyfarth, R.M., eds (1990) *How Monkeys See the World*, Chicago University Press
- 52 Stambach, E. (1988) Group responses to specially skilled individuals in *Macaca fascicularis* *Behaviour* 107, 241–266
- 53 Hauser, M.D. (1992) Costs of deception: cheaters are punished in rhesus monkeys *Proc. Natl. Acad. Sci. U. S. A.* 89, 12137–12139
- 54 Hauser, M.D. and Marler, P. (1993) Food-associated calls in rhesus macaques (*Macaca mulatta*): II. Costs and benefits of call production and suppression *Behav. Ecology* 4, 206–212
- 55 Byrne, R.W. (1995) *The Thinking Ape. Evolutionary Origins of Intelligence*, Oxford University Press
- 56 Byrne, R.W. and Whiten, A. (1988) *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*, Clarendon Press
- 57 Machiavelli, N. (1532) *Il Principe* (English transl., 1979, *The Prince*), Penguin Books
- 58 Meltzoff, A. and Moore, M.K. (1995) Infants' understanding of people and things: from body imitation to folk psychology, in *The Body and the Self* (Bermudez, J.L. et al., eds), pp. 43–69, MIT Press
- 59 Stich, S. and Nichols, S. (1992) Folk psychology: simulation versus tacit theory *Mind Lang.* 7, 29–65

Books Received

Review copies of the following books have been received. Books that have been reviewed in Trends in Cognitive Sciences are not included. The appearance of a book in the list does not preclude the possibility of it being reviewed in the future.

M.A. Arbib (ed.) *The Handbook of Brain Theory and Neural Networks* MIT Press, 1998. \$75.00/\$49.95 (xi + 1118 pages) ISBN 0 262 51102 9

P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis and D. Pesetsky (eds) *Is The Best Good Enough? Optimality and Competition in Syntax* MIT Press, 1998. £21.95 (450 pages) ISBN 0 262 52249 7

A.G. Cairns-Smith *Evolving the Mind on the Nature of Matter and the Origin of Consciousness* Cambridge

University Press, 1998. £9.95/\$15.95 (viii + 329 pages) ISBN 0 521 63755 4

S. Healy (ed.) *Spatial Representation in Animals* Oxford University Press, 1998. HB £45.00, PB £22.50 (x + 188 pages) ISBN HB 0 19 850007 6, PB 0 19 850006 8

F. Jongejan, W. Goff and E. Camus (eds) *Tropical Veterinary Medicine: Molecular Epidemiology, Hemoparasites and their Vectors, and General Topics* The New York Academy of Sciences, 1998. \$140.00 (xiv + 503 pages) ISBN 1 57331 142 1

D.A. Norman *The Design of Everyday Things* MIT Press, 1998. £9.95 (xiii + 257 pages) ISBN 0 262 64037 6

M. Paradis (ed.) *Pragmatics in Neurogenic Communication Disorders* Pergamon, 1998. \$122.50/Dfl. 213.00 (x + 257 pages) ISBN 0 08 043065 1

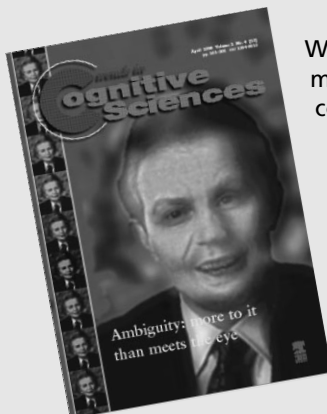
D. Scarborough and S. Sternberg (eds) *Methods, Models and Conceptual Issues* MIT Press, 1998. £24.95 (xiv + 950 pages) ISBN 0 262 65046 0

N.J. Wade *A Natural History of Vision* MIT Press, 1998. £34.95/\$55.00 (xvi + 466 pages) ISBN 0 262 23194 8

T. Watanabe (ed.) *High-level Motion Processing: Computational, Neurobiological, and Psychophysical Perspectives* MIT Press, 1998. £47.95 (x + 417 pages) ISBN 0 262 23195 6

And finally,

Trends in Cognitive Sciences relies on the goodwill and cooperation of all its contributors. We should like to acknowledge and thank the efforts of the authors, the referees and members of the Advisory Editorial Board in assisting the journal in all stages in its production.



We hope that readers have enjoyed 1998 as much as we have and look forward to your continued support in the coming year.

Peter Collins, Editor

Julian Ogilvie, Assistant Editor

Lucy Greenberry and Tracy-Ann Smith, Editorial Administrators

Margaret Newell, Production Coordinator

