Embodied simulation: from mirror neuron systems to interpersonal relations

Vittorio Gallese

Dipartimento di Neuroscienze, Università di Parma, Via Volturno 39, 43100 Parma, Italy

Abstract. A direct form of 'experiential understanding' of others is achieved by modelling their behaviours as intentional experiences on the basis of the equivalence between what the others do and feel and what we do and feel. This modelling mechanism is embodied simulation. By means of embodied simulation we do not just 'see' an action, an emotion, or a sensation. Side by side with the sensory description of the observed social stimuli, internal representations of the body states associated with actions, emotions, and sensations are evoked in the observer, as if he/she would be doing a similar action or experiencing a similar emotion or sensation. Mirror neurons are likely the neural correlate of this mechanism. The mirror neuron matching systems map the different intentional relations in a compressed fashion, which is neutral about the specific quality or identity of the agentive/subjective parameter. By means of a shared neural state realized in two different bodies that nevertheless obey to the same functional rules, the 'objectual other' becomes 'another self'.

During the last decades, developmental psychology research has provided one of the major contributions to a new understanding of human social cognition. In the course of infancy and childhood, we all heavily rely on interactions with our caregivers and with other individuals to learn how to cope with the world. Developmental psychology, by providing an enormous amount of data, has literally revolutionized our way of looking at newborns and infants as cognitive agents. These results have shown, among other things, that at the very beginning of our life we almost immediately interact with others by *reproducing* some of their behaviours.

Several studies have shown that the capacity of infants to establish relations with 'others' is accompanied by the registration of behavioural invariance. As pointed out by Stern (1985), this invariance encompasses unity of locus, coherence of motion and coherence of temporal structure. This experience-driven process of constant remodelling is one of the building blocks of cognitive development, and it capitalizes upon coherence, regularity and predictability. Social identity

guarantees all these features, henceforth its high social adaptive value. The experience of identity between infant and caregiver is *the* starting point for the development of social cognition.

The seminal study of Meltzoff & Moore (1977) and the subsequent research field it opened showed that newborns as young as 18 hours are capable of reproducing mouth and face movements displayed by the adult they are facing. That particular part of their body replies, though not in a reflex way, to movements displayed by the equivalent body part of someone else. More precisely, this means that newborns set into motion a part of their body they have no visual access to, but which nevertheless matches an observed behaviour. To put it crudely, visual information is transformed into motor information. The issue then consists in clarifying the nature of this peculiar feature and the possible underlying mechanisms. The relational character intrinsic to the interaction between any biological system and its environment appears to be a good candidate. Our environment is composed of a variety of lifeless forms of matter, and of a variety of 'alive stuff', whose peculiar character is more and more focused by the infant's immature eye. Individuals confront themselves with all possible kinds of 'external' objects, in virtue of their peculiar status of biological systems, thus by definition constrained in their peculiar 'modes of interaction' (see Gallese 2003).

Interpersonal relations are established at the very onset of our life, when a full-blown self-conscious subject of experience is not yet constituted. Yet, the absence of a subject doesn't preclude the presence of a primitive 'we-centric space', a paradoxical form of intersubjectivity without subject. The infant shares this space with others. The physical space occupied by the bodies of the adult-others is 'hooked up' to the body of the infant to compose a blended shared space. In a way, it is as if the mother, who creates and holds the fetus within her body during pregnancy, continues to hold and create the child in his/her first months and years of life, being both biologically and culturally connected in fundamental ways. This intersubjective process continues for the entire lifespan, becoming much richer and multifaceted, due to the wider range and meaning of interpersonal relations in the course of development.

The shared we-centric space enables the social bootstrapping of cognitive and affective development because it provides a powerful tool to detect and incorporate coherence, regularity and predictability in the course of the interactions of the individual with the environment. The we-centric space is paralleled by the development of perspectival spaces defined by the establishment of the capacity to distinguish self from other, as long as self-control develops. Within each of these newly acquired perspectival spaces information can be better segregated in discrete channels (visual, somatosensory, etc.) making the perception of the world more finely grained. The concurrent development of language contributes to further segregate from the original multimodal perceptive world, single characters or modalities of

experience. Yet, the more mature capacity to segregate the modes of interaction, together with the capacity of carving out the subject and the object of the interaction, do not annihilate the shared we-centric space.

The shared intersubjective we-centric space progressively acquires a different role. It provides the self with the capacity to simultaneously entertain self-other identity and difference. Once the crucial bonds with the world of others are established, this space carries over to the adult conceptual faculty of socially mapping sameness and difference ('I am a different subject'). Within intersubjective relations, the other is a living oxymore, being just a different self. Social identity, the 'selfness' we readily attribute to others, the inner feeling of 'being-like-you' triggered by our encounter with others, are the result of the preserved shared we-centric space. Self-other physical and epistemic interactions are shaped and conditioned by the same body and environmental constraints. This common relational character is underpinned, at the level of the brain, by neural networks that compress the 'who-done-it', 'who-is-it' specifications, and realize a narrower content state, a content that specifies what kinds of interaction or state are at stake. This narrower content is shared just not only because the shareable character of experience and action is the earliest constituent of our social life, but also because it is underpinned by shared neural mechanisms.

The posited important role of identity relations in constraining the cognitive development of our mind provides a strong motivation to investigate from a neuroscientific perspective the functional mechanisms, and their neural underpinnings, at the basis of the self-other identity. This will be the focus of the next sections.

The mirror neuron system for actions in monkeys and humans: empirical evidence

About 10 years ago a new class of premotor neurons was discovered in the ventral premotor cortex of the macaque monkey brain. These neurons discharge not only when the monkey executes goal-related hand actions like grasping objects, but also when observing other individuals (monkeys or humans) executing similar actions. They were called 'mirror neurons'[1] (Gallese et al 1996, Rizzolatti et al

[1]This paper is exclusively focused on the relationships among the mirror neuron system, embodied simulation and the experiential aspects of intersubjectivity For sake of concision, many other issues related to mirror neurons and simulation will not be addressed here. The vast literature on the mirror neuron system in humans and its relevance for theory of mind, imitation and the evolution of language is reviewed and discussed in several papers (Gallese & Goldman 1998, Rizzolatti & Arbib 1998, Gallese 2003, Rizzolatti & Craighero 2004, Gallese et al 2004). For the analysis of the role played by embodied simulation in conceptual structure and content, see Gallese & Lakoff (2005).

1996). Neurons with similar properties were later discovered in a sector of the posterior parietal cortex reciprocally connected with area F5 (PF/PG mirror neurons; see Gallese et al 2002, Rizzolatti & Craighero 2004, Fogassi et al 2005).

Action observation causes in the observer the automatic activation of the same neural mechanism triggered by action execution. It has been proposed that this mechanism could be at the basis of a direct form of action understanding (Gallese et al 1996, 2004, Rizzolatti et al 2001).

Further studies carried out by our research group at the Department of Neuroscience of the University of Parma corroborated and extended the original hypothesis. It was shown that F5 mirror neurons are also activated when the final critical part of the observed action, that is, the hand–object interaction, is hidden (Umiltà et al 2001). A second study showed that a particular class of F5 mirror neurons, 'audiovisual mirror neurons', can be driven not only by action execution and observation, but also by the sound produced by the same action (Kohler et al 2002).

More recently, the most lateral part of area F5 was explored where a population of mirror neurons related to the execution/observation of mouth actions was described (Ferrari et al 2003). The majority of these neurons discharge when the monkey executes and observes transitive, object-related ingestive actions, such as grasping, biting or licking. However, a small percentage of mouth-related mirror neurons discharge during the observation of intransitive, communicative facial actions performed by the experimenter in front of the monkey ('communicative mirror neurons'; Ferrari et al 2003). Thus, mirror neurons seem also to underpin aspects of monkeys' social facial communication.

Several studies using different experimental methodologies and techniques have demonstrated also in the human brain the existence of a mirror neuron system matching action perception and execution. During action observation there is a strong activation of premotor and parietal areas, the likely human homologue of the monkey areas in which mirror neurons were originally described (for review, see Rizzolatti et al 2001, Rizzolatti & Craighero 2004, Gallese et al 2004). Furthermore, the mirror neuron matching system for actions in humans is somatotopically organized, with distinct cortical regions within the premotor and posterior parietal cortices being activated by the observation/execution of mouth-, hand- and foot-related actions (Buccino et al 2001).

The involvement of the motor system during observation of communicative mouth actions is also testified by the results of recent functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS) studies (Buccino et al 2004, Watkins et al 2003). The observation of communicative, or speech-related mouth actions, facilitate the excitability of the motor system involved in the production of the same actions.

D2

Mirror neurons and the understanding of intentions

When an individual starts a movement aimed to attain a goal, such as picking up a pen, he/she has clear in mind what he/she is going to do, for example writing a note on a piece of paper. In this simple sequence of motor acts the final goal of the whole action is present in the agent's mind and is somehow reflected in each motor act of the sequence. The action intention, therefore, is set before the beginning of the movements. This also means that when we are going to execute a given action we can also predict its consequences.

Monkeys may exploit the mirror neuron system to optimize their social interactions. My hypothesis is that monkeys might entertain a rudimentary form of 'teleological stance', a likely precursor of a full-blown intentional stance. This hypothesis extends to the phylogenetic domain the ontogenetic scenario proposed by Gergely & Csibra (2003) for human infants. New experiments are being designed in my lab to test this hypothesis.

But monkeys certainly do not entertain full-blown mentalization. Thus, what makes humans different? At present we can only make hypotheses about the relevant neural mechanisms underpinning the mentalizing abilities of humans, still poorly understood from a functional point of view. In particular, we do not have a clear neuroscientific model of how humans can understand the intentions promoting the actions of others they observe.

A given action can be originated by very different intentions. Suppose one sees someone else grasping a cup. Mirror neurons for grasping will most likely be activated in the observer's brain. A simple motor equivalence between the observed action and its motor representation in the observer's brain, however, can only tell us what the action is (it's a grasp) and not why the action occurred. Determining why action A (grasping the cup) was executed, that is, determining its intention, can be equivalent to detecting the goal of the still not executed and impending subsequent action (say, drink from the cup).

In an fMRI study we recently published (Iacoboni et al 2005), subjects watched three kinds of stimuli: grasping hand actions without a context, context only (a scene containing objects), and grasping hand actions embedded in contexts. In the latter condition the context suggested the intention associated with the grasping action (either drinking or cleaning up). Actions embedded in contexts, compared with the other two conditions, yielded a significant signal increase in the posterior part of the inferior frontal gyrus and the adjacent sector of the ventral premotor cortex where hand actions are represented. Thus, premotor mirror areas—areas active during the execution and the observation of an action—previously thought to be involved only in action recognition are actually also involved in understanding the 'why' of action, that is, the intention promoting it. Detecting the intention of Action A is

D2

equivalent to predict its distal goal, that is, the goal of the subsequent Action B.

Similar findings were recently obtained in monkeys. Fogassi et al (2005) described a class of parietal mirror neurons whose discharge during the observation of an act (e.g. grasping an object), is conditioned by the type of not yet observed subsequent act (e.g. bringing the object to the mouth) specifying the overall action intention. Thus, these neurons not only code the observed motor act but also seem to allow the observing monkey to predict the agent's next action, henceforth his/her overall intention. It is possible to interpret this mechanism as the neural correlate of the dawning of more sophisticated mentalizing abilities, as those characterizing our species.

The statistical detection of what actions most frequently follows other actions, as they are habitually performed or observed in the social environment, can constrain preferential paths of inferences/predictions. It can be hypothesized that this can be accomplished by chaining different populations of mirror neurons coding not only the observed motor act, but also those that in a given context would normally follow. Ascribing intentions would therefore consist in predicting a forthcoming new goal. If this is true, it follows that one important difference between humans and monkeys could be the level of recursivity attained by the mirror neuron system in our species. According to this perspective, action prediction and the ascription of intentions are related phenomena, underpinned by the same functional mechanism. In contrast with what mainstream cognitive science would maintain, action prediction and the ascription of intentions—at least of simple intentions—do not appear to belong to different cognitive realms, but both pertain to embodied simulation mechanisms underpinned by the activation of chains of logically related mirror neurons (see Iacoboni et al 2005, Fogassi et al 2005).

Mirroring emotions and sensations

Emotions constitute one of the earliest ways available to the individual to acquire knowledge about its situation, thus enabling a reorganization of this knowledge on the basis of the outcome of the relations entertained with others. The coordinated activity of sensory–motor and affective neural systems results in the simplification and automatization of the behavioural responses that living organisms are supposed to produce in order to survive. The integrity of the sensory–motor system indeed appears to be critical for the recognition of emotions displayed by others (see Adolphs 2003), because the sensory–motor system appears to support the reconstruction of what it would feel like to be in a particular emotion, by means of simulation of the related body state. The implication of this process for empathy should be obvious.

D2

A recently published fMRI study showed that experiencing disgust and witness-ing the same emotion expressed by the facial mimicry of someone else, both acti-vate the same neural structure—the anterior insula—at the same overlapping location (Wicker et al 2003). This shows that when we see the facial expression of someone else, and this perception leads us to experience a particular affective state, the other's emotion is constituted, experienced and therefore directly understood by means of an embodied simulation producing a shared body state. It is the acti-vation of a neural mechanism shared by the observer and the observed to enable direct experiential understanding. A similar simulation-based mechanism has been proposed by Goldman & Sripada (2005) as 'unmediated resonance'.

Let us now examine somatic sensations as the target of our social perception. As repeatedly emphasized by phenomenology, touch has a privileged status in making possible the social attribution of lived personhood to others. 'Let's be in touch' is a common clause in everyday language, which metaphorically describes the wish of being related, being in contact with someone else. Such examples show how the tactile dimension can be intimately related to the interpersonal dimension.

New empirical evidence suggests that the first-person experience of being touched on one's body activates the same neural networks activated by observing the body of someone else being touched (Keysers et al 2004, Blakemore et al 2005). This double pattern of activation of the same somatosensory-related brain regions suggests that our capacity to experience and directly understand the tactile experi-ence of others could be mediated by embodied simulation, that is, by the externally triggered activation of *some* of the same neural networks underpinning our own tactile sensations. A similar mechanism likely underpins our experience of the painful sensations of others (see Hutchison et al 1999, Singer et al 2004, Avenanti et al 2005).

Intentional attunement, embodied simulation and empathy

Various mirror neurons matching systems mediate between the multimodal expe-riential knowledge we hold of our lived body, and the experience we make of others. Such body-related experiential knowledge enables a direct grasping of the sense of the actions performed by others, and of the emotions and sensations they experience. Our capacity to conceive of the acting bodies of others as *persons* like us depends on the constitution of a shared meaningful interpersonal space. This 'shared manifold' (see Gallese 2001, 2003, 2005) can be characterized at the func-tional level as embodied simulation, a specific mechanism constituting a basic functional feature by means of which our brain/body system models its interac-tions with the world. Embodied simulation constitutes a crucial functional mecha-nism in social cognition, and it can be neurobiologically characterized. The different

D2

mirror neuron systems represent the sub-personal instantiation of embodied simulation.

When we confront the intentional behaviour of others, embodied simulation generates a specific phenomenal state of 'intentional attunement'. This phenomenal state in turn generates a peculiar quality of familiarity with other individuals, produced by the collapse of the others' intentions into the observer's ones. By means of embodied simulation we do not just 'see' an action, an emotion, or a sensation. Side by side with the sensory description of the observed social stimuli, internal representations of the body states associated with these actions, emotions, and sensations are evoked in the observer, 'as if' he/she would be doing a similar action or experiencing a similar emotion or sensation.

Any intentional relation can be mapped as a relation between a subject and an object. The mirror neuron matching systems described in this paper map the different intentional relations in a fashion that is neutral about the specific quality or identity of the agentive/subjective parameter. By means of a shared functional state realized in two different bodies that nevertheless obey to same functional rules, the 'objectual other' becomes 'another self'.

Of course, embodied simulation is not the only functional mechanism underpinning social cognition. The same actions performed by others in different contexts can lead the observer to radically different interpretations. Social stimuli can also be understood on the basis of the explicit cognitive elaboration of their contextual perceptual features, by exploiting previously acquired knowledge about relevant aspects of the situation to be analysed. Our capacity of attributing false beliefs to others, our most sophisticated mind reading abilities, likely involve the activation of large regions of our brain, certainly larger than a putative and domain-specific theory of mind module. Embodied simulation and the still poorly understood more sophisticated mentalizing cognitive skills, however, are not mutually exclusive. Embodied simulation, probably the most ancient mechanism from an evolutionary point of view, is experience-based, while the second mechanism can be characterized as a 'detached' cognitive description of an external state of affairs. It might well be the case that embodied simulation scaffolds the propositional, language-mediated mechanism. When the former mechanism is not present or malfunctioning, as perhaps in the autistic spectrum disorder (ASD), the latter can provide only a pale, detached account of the social experiences of others (see Gallese et al 2004). Recent evidence seems to support the hypothesis of ASD as at least in part due to a defective intentional attunement (see Oberman et al 2005, Theoret et al 2005).

Conclusions

D2

Social cognition is not only thinking about the contents of someone else's mind. Our brains, and those of other primates, have developed a basic functional mecha-

nism, embodied simulation, which gives us an experiential insight of other minds. The neuroscientific evidence here reviewed suggests that social cognition is tractable at the neural level of description. This level is implicit, though, when the organism is confronting the intentional behaviour of others, it produces a specific phenomenal state of 'intentional attunement'. This phenomenal state generates a peculiar quality of familiarity with other individuals, produced by the collapse of the others' intentions into the observer's ones. This seems to be one important component of what being empathic is about.

However, self-other identity is not all there is in empathy. Empathy, at difference with emotional contagion, entails the capacity to experience what others do experience, while being able to attribute these shared experiences to *others* and not to the self. The quality of our *erlebnis* of the external world and its content are constrained by the presence of other subjects that are intelligible, while preserving their alterity character. An alterity that is present also at the sub-personal level, instantiated by the different neural networks coming into play and/or by their different degree of activation when *I* act with respect to when others act, or when *I* experience an emotion or a sensation with respect when others do the same.

References

Adolphs R 2003 Cognitive neuroscience of human social behaviour. Nat Rev Neurosci 4: 165–178

Avenanti A, Bueti D, Galati G, Aglioti SM 2005 Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. Nat Neurosci 8:955–960

Blakemore S-J, Bristow D, Bird G, Frith C, Ward J 2005 Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. Brain 128:1571–1583

Buccino G, Binkofski F, Fink GR et al 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur J Neurosci 13:400–404

Buccino G, Lui F, Canessa N et al 2004 Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. J Cogn Neurosci 16:114–126

Ferrari PF, Gallese V, Rizzolatti G, Fogassi L 2003 Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. Eur J Neurosci 17:1703–1714

Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G 2005 Parietal lobe: From action organization to intention understanding. Science 302:662–667

Gallese V 2001 The 'shared manifold' hypothesis: from mirror neurons to empathy. J Consciousness Stud 8:33–50

Gallese V 2003 The manifold nature of interpersonal relations: The quest for a common mechanism. Phil Trans Royal Soc London B 358:517–528

Gallese V 2005 Embodied simulation: from neurons to phenomenal experience. Phenomenology Cog Sci 4:23–48

Gallese V, Goldman A 1998 Mirror neurons and the simulation theory of mind-reading. Trends Cogn Sci 12:493–501

Gallese V, Lakoff G 2005 The brain's concepts: The role of the sensory-motor system in reason and language. Cogn Neuropsychol 22:455–479

Gallese V, Fadiga L, Fogassi L, Rizzolatti G 1996 Action recognition in the premotor cortex. Brain 119:593–609

D2

Gallese V, Fogassi L, Fadiga L, Rizzolatti G 2002 Action representation and the inferior parietal lobule. In Prinz W, Hommel B (eds) Attention and performance XIX. Oxford University Press, Oxford, p 247–266

Gallese V, Keysers C, Rizzolatti G 2004 A unifying view of the basis of social cognition. Trends Cogn Sci 8:396–403

Gergely G, Csibra G 2003 Teleological reasoning in infancy: the naive theory of rational action. TICS 7:287–292

Goldman A, Sripada CS 2005 Simulationist models of face-based emotion recognition. Cognition 94:193–213

Hutchison WD, Davis KD, Lozano AM, Tasker RR, Dostrovsky JO 1999 Pain related neurons in the human cingulate cortex. Nat Neurosci 2:403–405

Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta J, Rizzolatti G 2005 Grasping the intentions of others with one's owns mirror neuron system. PLOS Biology 3:529–535

Keysers C, Wickers B, Gazzola V, Anton J-L, Fogassi L, Gallese V 2004 A Touching sight: SII/PV activation during the observation and experience of touch. Neuron 42: 335–346

Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G 2002 Hearing sounds, understanding actions: Action representation in mirror neurons. Science 297:846–848

Meltzoff AN, Moore MK 1977 Imitation of facial and manual gestures by human neonates. Science 198:75–78

Oberman LM, Hubbard EH, McCleery JP, Altschuler E, Ramachandran VS, Pineda JA 2005 EEG evidence for mirror neuron dysfunction in autism spectrum disorders. Cog Brain Res 24:190–198

Rizzolatti G, Arbib MA 1998 Language within our grasp. Trends Neurosci 21:188–194

Rizzolatti G, Craighero L 2004 The mirror neuron system. Ann Rev Neurosci 27: 169–192

Rizzolatti G, Fadiga L, Gallese V, Fogassi L 1996 Premotor cortex and the recognition of motor actions. Cog Brain Res 3:131–141

Rizzolatti G, Fogassi L, Gallese V 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. Nat Rev Neurosci 2:661–670

Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, Frith CF 2004 Empathy for pain involves the affective but not the sensory components of pain. Science 303:1157–1162

Stern DN 1985 The interpersonal world of the infant. Karnac Books, London

Theoret H, Halligan E, Kobayashi M, Fregni F, Tager-Flusberg H, Pascual-Leone A 2005 Impaired motor facilitation during action observation in individuals with autism spectrum disorder. Curr Biology 15:84–85

Umiltà MA, Kohler E, Gallese V et al 2001 'I know what you are doing': a neurophysiological study. Neuron: 32:91–101

Watkins KE, Strafella AP, Paus T 2003 Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41:989–994

Wicker B, Keysers C, Plailly J, Royet J-P, Gallese V, Rizzolatti G 2003 Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. Neuron 40:655–664

DISCUSSION

C Frith: The key theme so far is the idea that we can actually share the experiences of other people because there are built-in brain mechanisms that somehow interpret what we see and recreate the experience in ourselves. Do people feel that this is sufficient to explain concepts such as empathy?

Singer: Clearly, this affect sharing mechanism is just one mechanism underlying what is broadly referred to as empathy. In addition to the ability to share other peoples' feelings, we also have to distinguish between self and other, and be able to modulate and control our empathic abilities. Thus, we do not always engage in empathy. Most of the time we actually do not empathize with others. Coming back to the shared affect mechanism or research on mirror neurons, I asked myself whether the emergence of such shared representations between self and other could not simply be accounted by associative learning mechanism? Associative learning mechanisms could easily explain empathic responses in the action and the emotional domain. Do we need to implicate mirror neurons as a specific neuronal mechanism to account for the brain imaging data acquired? Could we not just assume that these are parts of extended associative network connecting perceptual inputs to action or emotional outputs?

Gallese: Your question is about how this matching has been formed ontogenetically. This is an interesting point. Unfortunately, the truth is that we know very little about the ontogenetic aspects of mirror neurons. Colleagues of mine have started studying imitation in newborn monkeys, in collaboration with Steve Suomi. The results seem to suggest that macaque monkeys have early imitation, just as humans and chimps do. Some of these mechanisms could therefore be hard-wired. Nevertheless, there are other sets of data showing that the mirror matching system is highly plastic. For example, we now have data on two monkeys trained to grasp objects using a tool. In these monkeys mirror neurons also respond to the observation of tool use. This shows that by means of association the system can learn to respond to different stimuli. Why do you put associative and mirroring mechanisms as possibly conflicting? I don't understand this.

Singer: You are right. These accounts are not contradictory. The one is the mechanism allowing these mirror neurons to emerge. The question is probably more related to how specific you assume mirror neurons to be. Do you assume that the mirror neurons you have measured in the monkey brain are very highly-specialized neurons, or just part of a huge network coding less specifically for all types of actions? I mean these neurons could be of a large associative network that has been formed by learned association by, for example, seeing yourself doing the hand action and thus associating the side of a hand action to the motor performance of this action. The side of a similar action in the other is then a cue to activate the network also containing the motor program for this action.

Gallese: About one third of mirror neurons are highly specific. The remaining two thirds show a broader congruence between the executed and observed action. Self-observation coupled to action execution could indeed provide the starting association to be used also to map the actions of others. The properties of mirror neurons are the outcome of the integrative work pooling together different input information.

D2

Blair: Isn't the real issue the difference between the way that you describe the functional properties of mirror neurons for motor movements and the way that the concept is then translated with reference to empathy? Mirror neurons for motor movements cannot be established directly through simple association. To form an association between when you see someone else doing a movement and when you are seeing yourself doing the movement, a degree of translation is necessary. Whatever is going on, it has to be more complicated than what is going on in the pain studies. In a classic conditioning study you will see that some neurons in the amygdala will fire to pain and then come to fire to stimuli that actually anticipate the pain. You wouldn't want to call these neurons mirror neurons. There is a straight association process. If you have seen that a stimulus approach your own hand anticipates pain, it is unsurprising for an association point of view that the same stimulus approaching another hand might lead to pain associated activity. This would occur on basic association grounds. This has to be a different computational process from that seen with mirror neurons.

Gallese: I'm happy to confine the tag 'mirror neuron' to motor-related aspects of inter-subjectivity. Nevertheless, I think I'm right in pooling together these different sets of empirical evidence, to the extent that they all point to the same direction. In order to make the content of my social perception meaningful this has to go through an activation of similar embodied mechanisms in my brain. If I want to understand how it feels to be disgusted or how does it feel to be touched, this involves an activation of part of the brain that is actually activated when I am disgusted or touched. What binds together all the results I presented today is the underlying functional mechanism, what I qualify as embodied stimulation. It is a radically knew perspective. Knowing 'how does it feel' is not the result of a hermeneutical process applied to sense data. This is certainly possible, but it is not what is likely going on in most of our daily social interactions.

Gergely: How do you account for the perception of Heiderian types of stimuli (Heider & Simmel 1944), and their intentional interpretation? These animated events involve 2D abstract figures such as circles and rectangles moving in relation to each other in ways that evoke strong intentional interpretations as goal-directed actions of interacting agents not only by adult perceivers but importantly by one-year-old infants as well (Gergely et al 1995, Gergely & Csibra 2003). However, the figures and their movements have no easy way of being directly mapped onto already existing motor representations of actions within the repertoire of the perceiver. This seems especially problematic when such events lack any movement cues suggesting animacy or agency (such as self-propulsion) but are still interpreted as goal-directed actions by 9- and 12-month-olds (Csibra et al 1999). To me this suggests that understanding and attributing goals to such perceived actions must be accomplished by some entirely different mechanism than the activation of corresponding motor action representations through some process of 'direct matching' or 'motor resonance' (see Csibra's recent arguments on this, Csibra, 2005.

D2

Gallese: This may be true for the Heider and Simmel stimuli (Heider & Simmel 1944), but other abstract sequential stimuli have been used by R. Schubotz, and in that fMRI study they contrasted sequential biological actions and symbolic sequences of geometric shapes changing position on the screen. Subjects were required to predict whether the biological action or abstract sequence was goal directed or not. They had to anticipate the consequences of both abstract symbolic geometrical shape motion and biological motion. In both cases this led to strong activation of the ventral premotor cortex. Thus the abstract nature of stimuli doesn't prevent the involvement of the motor system. My bet is that we are going to learn more and more about the involvement of the sensory–motor system in the domain of syntax, for example. Embodied mechanisms may have something to say in this domain.

Gergely: If you get activation of the motor or mirror neuron system in such cases of non-biological motion of abstract figures that have no obvious similarity mapping on to the biomechanical motion properties of existing action schemes then I think you are postulating a rather mysterious mechanism of 'direct mapping' or 'motor resonance'. Without spelling out how you get from the perception of such abstract motion events to the activation of the premotor system, you have no viable model to account for the phenomena you are referring to.

Gallese: This is not necessarily due to the mirror system. When I am talking of premotor cortex, I am talking about of the neural correlates of different motor schemata.

Gergely: Doesn't this imply a kind of top–down route to activating the motor system? There has to be another system that infers and attributes the goal to the perceived action, which perhaps has a route of activating the motor system as a kind of action prediction or simulation mechanism.

Gallese: I know this line of argument. The problem is that no one knows this mysterious area where it is encoded. We stick to the extant empirical evidence and our claim is that we don't need to suppose an overarching top–down influence in order to have a neural mechanism that maps the goal. We already have it in the premotor system. We don't need to imply a further mechanism that maps the goal.

Gergely: I don't understand how the motor system becomes activated. What is the input that activates the motor system?

Gallese: This is what the motor system is there for: to guide actions by setting goals and end-states to be attained. The motor system is a lot more than a mere muscle controller! I should add that something we haven't looked for, but which must also play a key role, is the interaction between the reward system and the action system. Most likely we learn to code the fulfilment of a specific motor act as successfully leading to the acquisition of a target by means of a gating signal coming from reward-related brain areas. The interplay between the premotor cortex and reward-related areas is an interesting subject for future research.

D2

Montague: I am missing something about where goals come from. Humans can and do establish top down goals. People routinely kill themselves for political protests. People can hold goals in mind for long periods. Surely you would say that there are goal-forming systems in the brain that would have access to these mirror systems.

Gallese: I am not denying this. My point is that you don't need to imply a top–down mechanism to explain *these* data. Certainly, we entertain the capacity to have a distal goal and pursue it. But in principle this does not necessarily imply that you need a radically different mechanism. It could only be a matter of adding power to the same basic architecture we have uncovered in the monkey.

Montague: It operates on low level things. When I come home at night and I am starving and I jam my hand into a bag of potato chips, right before it goes into my mouth my expanding waistline and declining dating life flash into my mind and make me stop. So you are restricting it to these classes of data, such as the impact of watching other people in pain. That is a complicated representation, to think of someone else having pain. I'd be hard-pressed to give a simple associative learning account of this. Even if a simple associative learning account could explain the data you presented, there is still one variable that is missing, which is that you have to assign it to someone else.

Singer: Exactly.

Montague: That itself is an abstract entity that is forming.

Gallese: My point is that having a mechanism that enables the sharing of a given content with someone else is the most critical aspect of the story. If you don't have this mechanism, you are not going anywhere. The self/other distinction in my opinion is not the most difficult problem in social cognition, neither from a theoretical, nor from an empirical point of view. The 'hard problem' in social cognition is to understand how the epistemic gulf separating single individuals can be overcome. The solipsistic attitude, inspired by folk psychology and purported by the approach of classic cognitive science, leaves this hard problem unsolved. The discovery of mirror neurons and related mirroring phenomena for the first time provides a neurophysiological mechanism that explains how the intersubjective epistemic gap can be filled.

Montague: Would it be fair to say that you see this as us coming pre-equipped with these rich processes of what it feels like for our own bodies to have experiences? I look and see you doing something, and the most efficient way for me to process this is to plug it back into the way I do the same thing. But not a lot has been done for super-ordinate goals with respect to the mirror systems.

Gallese: That is correct. We are at the beginning of this new research. Just give it time, I am confident that very soon we'll know a lot more also about super-ordinate goals.

Call: How do you go from here to prior goals, or predicting new instances of some behaviour?

Gallese: I find it more interesting to pursue a line of research that is trying to emphasize cognitive continuity rather than sudden jumps. To oversimplify the issue, a quantitative leap forward can buy you a qualitative leap forward. We don't need to think about new areas or new magic cells we have and monkeys don't. The level of recursivity attained by the human brain is one possible explanation for humans' much more sophisticated social skills. We can run all these simulations without being driven by the local context. I can close my eyes and think what I will be doing in two weeks' time. From a qualitative point of view, this doesn't seem to be dramatically different from what a monkey can do. Perhaps it is just the way our brains are wired that enables us to have this greater predictive capacity or ability to entertain distal goals well before their execution.

Call: In the experiment you mentioned with the monkeys, where grasping to eat or grasping to play took place, in that case the monkey has experienced both and eventually discriminates both.

Gallese: Yes, there is also contextual information that helps the monkey.

Hauser: I want to go back to your sense of continuity. I can see the excitement surrounding these imitation results, but of course they stand in contrast to 50 years of failure to show imitation in monkeys. The story that has come out from your group is that there is an almost seamless connection between the physiological recordings and what humans seem to do. Is this really how you see it? That there is no difference in the capacity to form intentions, create goals and experience empathy?

Gallese: There is a huge difference between the animal and human data.

Hauser: If you run a cognitive subtraction, what is different? The way you argued today, I don't hear a difference.

Gallese: The paper was meant to highlight the similarities, not focus on the differences.

Hauser: What is left? What gives us as humans the particular signature?

Gallese: I don't know. One possibility would be that these mechanisms can use much more computational power, plus the development of language which gives an incredible leap forward socially.

Hauser: For me, language is too much of a throwaway. Saying that language is involved doesn't explain what's going, how language is involved or what aspect of language is doing the work. Let's go back to empathy. There is nothing in the animal literature that you have shown that has to do with emotion. You have the human studies which are claiming to be correlates of emotional experience, but there is nothing on the animal side. This is a big gap. Is there a case where an animal watches someone experience pain, for example?

Gallese: Colleagues of mine have started doing these sorts of experiments in the rat. The plan is to go soon into the monkey's insula.

Hauser: The simple experiment you could do, which would be ethical in monkeys, would be to use their vocalizations which are already coding information about emotion. The prediction is that when they produce a vocalization, hearing that call would be a trigger, and this would get around the association problem. It looks like the morphology of the signal is sometimes coded innately. I was intrigued that you were able to run such a natural experiment without any training at all. Now, if you can link the vocalization up to the emotions you have a natural experiment.

Gallese: We used chimpanzee vocalization as a stimulus to show the specificity of activation of audio–visual mirror neurons to the sounds produced by hand actions and it didn't work. But you are right, although it is really hard to induce a monkey to vocalize. You can record vocalizations and play them back to the monkey, but if you want to correlate the coding of the vocalization with the production, you also need to record the neuron when the monkey is actively producing the vocalization, which is a hell of a job.

Warneken: You said that this system creates an interpersonal space. If it is the case that an action is easier to understand when it is part of one's behavioural repertoire, it could also be that this goes beyond species barriers. Wouldn't this mandatory pre-rational process lead to false positives?

Gallese: This is what happens. I have friends literally in love with their pet boa constrictor!

Warneken: This means that there has to be another system coming in. What would that other system be? Is it something like face recognition?

Gallese: One thing that has been neglected so far is the specific quality of the observer. When we put people in the scanner we presume that our brains are wired up the same way. Personality traits can make a big difference. Our own social and cognitive history can make us react to the same stimulus in a different way. An encouraging line of research will be to show different patterns of activation induced by the same stimuli in subjects who have been screened before in a double blind way according to different personality trait ratings. I didn't include this in my paper. The take-home message of my short paper was that in order to start talking about empathy, we need a neural mechanism that enables us to bridge the gap.

De Vignemont: For philosophers mirror neurons are of great interest, because they could give a direct grasp on other people's feeling or thinking. However, in your account of intentions, you suggest that we have to *infer* the goal. Inferring is the opposite of a direct grasp. If we have indeed to infer the intention, then the mirror neurons account loses part of its interest. More specifically, in Fogassi et al (2005), monkeys have to put an apple into their mouth or on their shoulder. It

would have been interesting to see what would happen if they had to put an inedible, neutral object in their mouth? Then one would really see whether they can detect intention, because the action would be exactly the same.

Gallese: This is what they did: they described this in the paper.

De Vignemont: The object was placed into a container located near the monkey's mouth, not in the mouth itself.

Gallese: It is a kind of statistical evaluation of the situation. Context, stimulus and action. Some kind of stimuli can make a given intention more predictable than others. Indeed, if the quality of the object that should induce that intention to be activated is patently falsified because the action is different, they saw some of the neurons decrease the discharge rate. In a sense it is a probabilistic mechanism. The mechanism couldn't possibly work without other brain regions that carry out this type of analysis of the quality of the object.

De Vignemont: Thus, we go back to the question whether mirror neurons by themselves suffice to provide a direct grasp of intentions. It rather seems that intention understanding relies partly on mirror neurons, partly on other brain areas. Mirror neurons are not sufficient. Understanding intentions requires inferring from the goal, from the context and from the movement. If this is really the case, then one cannot claim that we know the intentions of others in the same way that we know our own intentions through the mirror system. We have a direct knowledge of our intentions, while we have only an inferential indirect access to someone else's intentions. Mirror neurons cannot solve by themselves the problem of other minds.

Gallese: The interest of this approach is that it reduces the space to be investigated related to the non-direct or top–down mechanisms. It may enable us to focus more specifically on the highly relevant top–down mechanism in social cognition by showing that a large part of the job is done at a lower level. This doesn't exclude higher-level mechanisms, but it enables us to focus our investigation.

References

Csibra G 2005 Mirror neurons or emulator neurons? http://mirrorneurons.free.fr/Cisbramirrorfunction.pdf

Csibra G, Gergely G, Biró S, Koós O, Brockbank M 1999 Goal-attribution and without agency cues: The perception of 'pure reason' in infancy. Cognition 72:237–267

Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G 2005 Parietal lobe: from action organization to intention understanding. Science 308:662–667

Gergely G, Csibra G 2003 Teleological reasoning about actions: the naïve theory of rational action. Trends Cog Sci 7:287–292

Gergely G, Nádasdy Y, Csibra G, Bíró S 1995 Taking the intentional stance at 12 months of age. Cognition 56:165–193

Heider F, Simmel M 1944 An experimental study of apparent behavior. Am J Psychol 57: 243–259

D2