Empathy

Philosophical and Psychological Perspectives

EDITED BY

Amy Coplan and Peter Goldie

OXFORD
UNIVERSITY PRESS

:. The first
1orance. If
is pregnant
method is
nd type of
idence that
quarantine'
ocess when
is results in
ubstantially
vith Vigne-
of the two
lem placed

1g route and
e route also
1coboni, this
discovery of
1us, parahip-
1eurons may
rt (Goldman,
7.[3] Thus it is
irror proper-
:ion between
processes are
with mirror
gination (see
pathy. These

1e same as midline
, specifically, the

4

Within Each Other:

Neural Mechanisms for Empathy
in the Primate Brain

Marco Iacoboni

4.1 Introduction

Empathy is commonly defined as the ability to understand and share the feelings of another. It is obviously a very complex ability. What are the neurophysiological mechanisms that underlie empathy? For years, nobody dared to investigate this issue. The main reasons were two. First, the study of the brain mechanisms associated with emotion and emotional understanding is relatively recent. Until approximately 20 years ago, the study of the neural systems associated with higher functions was focused exclusively on 'cold' cognitive processes. The dominant metaphor was 'the mind as a computer.' The study of emotions—especially complex social emotions— clearly did not fit in the prevalent paradigm. Second, even after emotions became a popular topic in cognitive neuroscience, mostly thanks to the influential work of Antonio Damasio, the neural mechanisms of empathy remained largely unexplored. This was likely due to the perceived complexity of empathy. Indeed, the complexity of a phenomenon is generally considered an obstacle for the study of its neural correlates, especially in single cell recordings. While neurophysiologists are able to study brain activity at its most exquisite spatial and temporal resolution, that is, the spiking activity of single cells, they also tend to study this activity in relation with relatively simple phenomena, such as the perception of individual sensory stimuli or the planning and execution of relatively simple actions. For this reason, neurophysiological data on empathy were virtually nonexistent until a few years ago. In recent years, however, a new wave of studies has investigated the links between empathic behavior and brain activity. The recent studies have been inspired by the discovery of mirror neurons in the macaque brain. These cells, which I describe in detail in the next section of the chapter, have *physiological properties* that are ideal to facilitate empathy. Indeed, the properties of mirror neurons seem to map extremely well onto emotional contagion, a phenomenon studied for decades by psychologists (Hatfield et al. (1994)). Most

scholars would probably agree that mirror neurons are likely critical neural elements for the relatively simple forms of empathic resonance that are observed in emotional contagion. However, most scholars would also argue that mirror neurons cannot account for more cognitively complex forms of empathy. In this chapter, I will argue instead that there are many different kinds of mirror neurons, and that they are also much more widely distributed in the primate brain than previously thought. The rich variety of mirroring responses and their diffuse anatomical localization suggest that neural mirroring may be a fundamental building block of empathy, even in its more complex forms.

In the next three sections of this chapter, I will review the single cell recordings on mirror neurons in macaques, the brain imaging data in humans that suggest links between activity in the human mirror neuron system and empathic behavior, and finally a set of unique data on single cell recordings in the human brain that demonstrate mirroring response in individual human neurons. In the final section of the chapter, I will discuss the theoretical implications of these empirical findings.

4.2 Mirror Neurons in the Macaque Brain

The first peer-reviewed scientific report on mirror neurons was published in 1992 (di Pellegrino et al. (1992)). This very short paper was followed four years later by a much more detailed report in which the term *mirror neurons* was used for the first time (Gallese et al. (1996)). The cells described in these two papers were recorded from the anterior sector of the ventral premotor cortex of *Macaca Nemestrina*. The anterior sector of the ventral premotor cortex in macaques is called area F5, and contains neurons with quite interesting motor properties (Rizzolatti et al. (1988)). The F5 motor neurons fire in relation with specific goal-oriented actions, rather than with individual movements. For instance, several F5 neurons fire during grasping actions, others during holding actions, and others during tearing actions. Interestingly, the same neuron may fire for a grasping action with the left hand *and* for a grasping action with the right hand. This firing pattern demonstrates that the firing of the cell does not occur in relation with the contraction of a specific set of muscles. Indeed, the same neuron will not fire for a different kind of action, say, scratching the head, which also involves the contraction of finger muscles used during grasping. Motor neurons in F5 show specificity of responses in relation with the *type* of grasp. Some cells discharge only during *precision grip* (opposition of thumb and index to grasp very small objects), others only during *finger prehension* (all fingers grasp a relatively small object), and others only during *whole hand prehension* (the whole hand grasps a big object). Taken together, these properties suggest that F5 motor neurons form a vocabulary of goal-oriented actions (Rizzolatti et al. (1988)), that is, as the vocabulary is the body of words of a language, the actions coded by F5 neurons seem to represent the body of actions that can achieve specific goals.

Approximately 20 to 25% of F5 motor neurons have also amazing *visual* properties. These cells fire also when the monkey is completely still and is just watching somebody

else making a goal-oriented action. It is as if the monkey is watching her own actions reflected by a mirror. This is why these cells are called mirror neurons. Mirror neurons are defined exclusively on the basis of their physiological properties. They are cells that are specialized for actions, and that discharge in association with motor and perceptual aspects of actions. There are two main categories of mirror neurons: *strictly congruent* mirror neurons and *broadly congruent* mirror neurons. Strictly congruent mirror neurons discharge for the same action both when the monkey is performing it and when the monkey is simply observing it performed by somebody else. Broadly congruent mirror neurons, in contrast, fire not only for the same action, but also for different actions that achieve the same goal both when the monkey is performing the action and when the monkey is simply observing it performed by somebody else (Gallese et al. (1996)). The properties of mirror neurons suggest that these cells map the actions of others onto the motor repertoires of the observer, thereby providing an internal simulation of the actions of other individuals in the observer.

Further studies demonstrated that mirror neurons discharge also when the observed action is partially occluded (Umiltà et al.(2001)) and when the monkey does not see the action at all, but simply listens to sounds typically associated with the action (for instance, the sound of breaking a peanut) (Kohler et al. (2002)). These data suggest that mirror neurons are multimodal cells that can provide a fairly abstract representation of the actions of other individuals. How abstract? A recent experiment has addressed this question (Fogassi et al. (2005)). The single cell recordings of this recent study were performed in area PF/PFG, an area in the anterior sector of the inferior parietal cortex. This sector of the inferior parietal cortex is anatomically connected with area F5 in the ventral premotor cortex (Rizzolatti & Luppino (2001)). Parietal motor neurons in PF/PFG were recorded when the monkeys were performing grasping actions associated with different outcomes, eating or placing. In some trials the monkey was allowed to grasp food and eat it, while in others the monkey was rewarded with food only after the animal had successfully completed the trial by grasping the food and placing it in a container. While of parietal motor neurons fired equivalently for grasping to eat and for grasping to place, the remaining demonstrated differential discharges during the grasping action associated with different outcomes. Approximately ¾ of these cells discharged more vigorously for grasping to eat, while ¼ discharged more vigorously for grasping to place. Note that the neuronal discharges were measured during grasping, that is, *before* the monkey would eat or place the food in the container.

After this necessary testing of the motor properties of the parietal cells, the experimenters tested the mirror properties of the same pool of cells. In some trials, the monkey simply watched an experimenter grasping a piece of food and eating it. In some other trials, the monkey watched the experimenter grasping the food and placing it in a container. The visual cue that signaled the outcome of the grasping action was the presence of the container. If the container was present, the experimenter placed the food in it. If the container was not present, the experimenter ate the food. A subset of motor parietal neurons displayed mirror properties, that is, they discharged when the

monkey simply observed the experimenter's grasping action. The discharge of these parietal mirror neurons during grasping observation mirrored the discharge previously measured during grasping execution. Cells that discharged equivalently for executed grasping actions associated with different outcomes, also discharged equivalently when the monkey simply observed grasping actions associated with different outcomes. For these mirror parietal cells, a grasp is a grasp, regardless of the outcome associated with it. Cells that discharged more when the monkey grasped to eat, also discharged more when the monkey watched the experimenter grasping to eat (the outcome was cued by the absence of the container). Also, cells that discharged more when the monkey grasped to place, also discharged more when the monkey watched the experimenter grasping to place (here the outcome was cued by the presence of the container). The preferential discharge for the same grasping action associated with different outcomes suggest that these mirror neurons do not simply code the action, but also the intention associated with it. In less mentalistic terms one could say that these cells predict the outcome of the observed grasping action, the action or actions that follow the grasp. This is evidence for a rather abstract and sophisticated coding of the observed grasp. The grasp is coded by these cells as embedded in a chain of concatenated and coordinated actions.

Mirror neurons do not simply code hand and hand-to-mouth actions. Both the ventral premotor and the inferior parietal cortex contain mirror neurons that code for mouth actions only. Two main types of actions are mirrored by these cells: ingestive and communicative actions. A recent depth electrode study demonstrated the mirroring properties of single cells in ventral premotor cortex for biting and sucking, and for lip-smacking, a communicative facial gesture with positive social valence (Ferrari et al. (2003)). The evidence that communicative facial gestures can be mirrored at the level of individual premotor neurons is theoretically important. It suggests that the evolutionary antecedents of empathy are based on relatively simple mechanisms of contagion and motor resonance.

4.3 Neural Systems with Mirroring Properties in the Human Brain

The single unit recordings in macaques described in the previous section have inspired a series of studies on humans. A common aspect of all these studies is that they cannot measure brain activity at the exquisite resolution of a single cell, as in the depth electrode studies in macaques. The neuroscience methods applied to the study of the human brain typically provide measures of *ensemble* neural activity, the activity of a large number of brain cells working together. Obviously, none of these techniques can definitely prove that the ensemble neural activity with mirroring properties truly represents the activity in concert of many mirror neurons. In principle, a neuronal ensemble may be activated during both action execution and action observation even

though its individual cells do not. Let me give you an example. Suppose I am measuring activity from an ensemble of ten neurons (keep in mind that the techniques that will be discussed in this section of the chapter actually measure ensemble activity of *millions* of neurons). Four of these neurons fire during execution of a grasping action but not during observation of the same grasping action performed by somebody else. They seem classical motor neurons. Four other neurons do not fire during grasping executions but do fire during grasping observation. They are visual neurons. The remaining two neurons do not fire during both conditions. It is unclear what their properties are. In this hypothetical scenario, a neuroscience method that measures ensemble neural activity would likely show some level of activation during both action observation and action execution, even though the neuronal ensemble does not contain any mirror neuron.

In spite of these interpretational limitations, it is reasonable to hypothesize that human brain areas contain mirror neurons if at least two conditions are met: first, the brain areas demonstrate ensemble activity compatible with mirror neuron activity; second, the anatomical location of the human brain areas is consistent with the anatomical location of mirror neurons recorded in the monkey brain. This second condition, however, has generated a lot of confusion among scholars. Indeed, many scholars now conflate anatomy and physiology in their definition of mirror neurons or mirror neuron areas. This is a mistake that should be avoided. As pointed out earlier, the definition of mirror neurons in monkeys is based only on their physiological properties. While such cells have been recorded—at least so far—only in areas F5 and PF/PFG, a cell with identical physiological responses in primary motor cortex or SMA would also be called mirror neuron. In other words, the anatomical location of the recorded cell is absolutely irrelevant to the physiological characteristics of the cell. Brain imaging studies conservatively used also the anatomical location of the activation only because the brain imaging signal is, as discussed above, inherently ambiguous.

Given that mirror neurons fire during both execution and observation of the similar actions, it makes sense to hypothesize that these cells may be important neural elements for imitative abilities. Imitation is pervasive in human behavior, and it is thought to play a major role in skill learning, transmission of local cultures, and in a variety of social and cognitive domains (Hurley & Chater (2005)). The first brain imaging study that linked imitation to human brain areas presumably containing mirror neurons (Iacoboni et al. (1999)) used the pattern of firing rate changes observed in the single unit studies in monkeys to predict the pattern of brain activity measured with functional magnetic resonance imaging (fMRI). This technique measures brain activity indirectly, by looking at the level of blood oxygenation in the brain. In the healthy brain, blood oxygenation and neural activity are fairly well correlated, making it possible to monitor brain activity with techniques that respond to changes in blood oxygenation rather than neuronal firing.

The firing rate changes of mirror neurons during action observation is approximately half the firing rate changes during action execution, according to the monkey studies

(Gallese et al. (1996)). This predicts that human brain areas with mirror neurons should have increases in fMRI activity during action observation that are also approximately half the increases measured during action execution. Furthermore, given that during imitation subjects both observe and execute the imitated action, it was predicted that imitation should yield—in mirror neuron areas—fMRI signal increases that are approximately the sum of the signal increases measured during action observation and action execution. Using this relatively simple model, an fMRI study on imitation of finger movements revealed two human brain areas with a pattern of activity consistent with mirror neuron activity. The two areas were located in the posterior part of the inferior frontal cortex and in the anterior part of the inferior parietal cortex (Iacoboni et al. (1999)). These anatomical locations are consistent with the single unit recordings in monkeys demonstrating that area F5 in the inferior frontal cortex and area PF/PFG in the inferior parietal cortex contain mirror neurons (Gallese et al. (1996); (Fogassi et al. (2005)). Thus, the posterior part of the inferior frontal cortex and the anterior part of the inferior parietal cortex meet the two necessary conditions to be considered human brain areas that presumably have mirror neurons: activity profile compatible with mirror neuron activity and anatomical location consistent with the anatomical location of mirror neurons in monkeys.

Recently, a series of fMRI studies have applied the adaptation paradigm to the investigation of the human mirror neuron system. The results obtained from these studies are sometimes consistent (Chong et al. (2008); Hamilton & Grafton (2008)) and sometimes inconsistent (Dinstein et al. (2007)) with the more classical 'subtraction' studies discussed above (in which some experimental conditions are contrasted with other experimental conditions). These mixed results are not surprising, because the neural correlates of adaptation paradigms are unclear (Tolias et al. (2005); Wilke et al. (2006)) and because there is no evidence that mirror neurons adapt at all. Indeed, adaptation is not a ubiquitous property of all cortical neurons and until adaptation has been unequivocally demonstrated in mirror neurons, fMRI adaptation studies of the mirror neuron system in humans have a questionable rationale.

Other fMRI studies using the classical 'subtraction' method have linked more directly the activity of human mirror neuron areas with empathy. The first study that investigated this issue (Carr et al. (2003)) was inspired by social psychology research that demonstrated that being imitated increases liking and that more empathic individuals tend to imitate other people more than less empathic individuals (Chartrand & Bargh (1999)). These behavioral data suggested functional links between the human mirror neuron system and neural systems more traditionally associated with emotional processing. The fMRI study tested the hypothesis that empathy requires the simulation (or inner imitation) of the facial emotional expressions of other people. Mirror neurons would provide such simulation process. Their connections with the limbic system via the insula would allow mirror neurons to send signals to limbic areas, such that the observer can feel what others are feeling. This model makes two predictions: first, there should be activation of mirror neuron areas, insula, and amygdala during both

observation and imitation of facial emotional expressions; second, in this network of areas the activity during imitation should be higher than during observation, as the previous study on imitation of finger movements had shown. Importantly, the higher activity during imitation should not be restricted to mirror neuron areas only. Indeed, if empathy requires the simulation of others' actions and functional links between mirror neurons and limbic areas, one would expect that the higher activity during imitation in mirror neuron areas would also spread to the insula and limbic areas. The empirical data confirmed both predictions (Carr et al. (2003)).

While this study was compelling in describing a large scale neural network supporting empathy via a simulative process implemented by mirror neurons, it did not provide any evidence linking the activity in this neural network and individual differences in empathy. Three recent fMRI studies have addressed this question. In one study, subjects listened to action sounds. As we have seen in the previous section of this chapter, action sounds trigger a discharge in mirror neurons. This predicts that human brain areas with mirror neurons should also become activated while listening to action sounds. Indeed, the study demonstrated that action sounds increased the activity of inferior frontal cortex (Gazzola et al. (2006)). This area has both activity profile and anatomical localization compatible with mirror neuron activity. Importantly, there was higher activity in subjects with high empathy scores.

Another fMRI study measured brain activity while subjects observed grasping actions (Kaplan & Iacoboni (2006)). As expected, grasping observation activated the inferior frontal cortex, where presumably human mirror neurons are located. The inferior frontal activity was correlated with empathy scores. The main difference between this study and the previous one is that while this study found a correlation *visual* between scores at scales measuring *emotional* empathy (scales that measure concern for the emotions of others and the emotional responses one experiences when watching someone else experiencing strong emotions) and activity in the mirror neuron system, the previous study found that subjects with high scores at scales measuring *cognitive* empathy (scales that measure the ability to imagine another person's perspective and the tendency to imagine oneself in the place of fictional characters in books or movies) had higher mirror neuron activity. Perhaps this difference is due to the different kind of stimuli used in the two studies. Action simulation while listening to action sounds may be mediated more by cognitive mechanisms, whereas action simulation during action *yes* observation may be based more on emotional resonance.

A more recent study has investigated the relationships between activity in mirror neuron areas and empathy in children (Pfeifer et al. (2008)). The children were asked to imitate and to simply observe facial emotional expressions displaying basic emotions. As in the previous study on adults (Carr et al. (2003)), mirror neuron areas, the insula, and the amygdala activated for both observation and imitation of facial emotional expressions, with higher activity during imitation. Correlation analyses were performed between brain activity and two types of scores: empathy scores and interpersonal competence scores. Emotional empathy scores correlated with activity in mirror

neuron areas during *observation* of facial emotional expressions, in line with the study on grasping observation (Kaplan & Iacoboni (2006)). Interestingly, mirror neuron activity during *imitation* of facial emotional expressions correlated with interpersonal competence scores. Indeed, overtly mirroring the emotions of others plays an important role in social interactions. It is through this mirroring that we communicate to other people that we understand what they are feeling. The fact that activity in mirror neuron areas maps well onto interpersonal competence during emotion imitation suggests that the mirror neuron system is a fairly nuanced bio-marker of sociality.

Further evidence in support of this hypothesis comes from imaging studies of autism. A recent fMRI study of observation and imitation of facial emotional expressions has revealed not only reduced activity in mirror neuron areas in children with Autism Spectrum Disorder compared to typically developing children, but also a correlation of the severity of the disorder with mirror neuron activity (Dapretto et al. (2005)). The more severe the disorder, the more reduced the activity in mirror neuron areas.

All these studies provide compelling evidence in support of the hypothesis that the mirror neuron system is a critical neural system for empathy. However, none of these studies provided evidence of the existence of mirror neurons in humans. Obviously, it would be very puzzling if humans had no mirror neurons. However, there are many evolutionary steps between macaques and humans. It is reasonable to assume that mirror neurons have evolved too. Thus, depth electrode recordings in the human brain would be very valuable not only to provide empirical evidence for the existence of mirror neurons in humans, but also because it would allow us to compare the properties of mirror neurons in macaques and humans. The next section of the chapter describes preliminary observations on individual human neurons with mirroring properties.

4.4 Mirror Neurons in the Human Brain

In some neurological patients, epilepsy cannot be controlled efficiently by anti-epileptic drugs. In these cases, neurosurgery is necessary. It is imperative to localize the epileptogenic brain tissue with precision, and in some patients it is necessary to implant depth electrodes, stop the administration of anti-epileptic drugs, and wait until the patient seizes to find exactly where is the brain locus of epileptogenic activity. This procedure typically takes few days up to two weeks. During this time, the patient has implanted electrodes in the depth of the brain, which makes it possible to measure brain activity. Typically, these electrodes can only record EEG signal, the slow waves of electrical brain activity that represent neuronal firing of many brain cells working together. However, the group of the neurosurgeon Itzhak Fried at UCLA has modified the electrodes typically used in these patients, such that it is possible to measure activity from individual neurons (Fried et al. (1999)).

Obviously, the brain location of the implanted electrodes is exclusively determined by medical considerations and electrodes are quickly de-planted when the patient

seizes. Furthermore, the modification of the electrodes that allows individual neuron recordings is located only at the tip of the electrode. These electrodes are relatively large electrodes. The way they are typically implanted determines that the tip of the electrode is always located in medial brain structures, that is, in brain areas quite different from the human brain areas that are widely presumed to contain mirror neurons. In spite of all these limitations, it makes sense to investigate mirror properties in these patients. As already discussed, there are many evolutionary steps from macaques to humans. It is reasonable to assume that the mirror neuron system evolved too, from the macaque brain to the human brain. This evolution may have taken two forms: on the one hand, an anatomical expansion of the system, such that other brain areas may contain mirror neurons (we should also keep in mind that it is possible that other brain areas in the macaque brain actually contain mirror neurons—neurophysiologists have not mapped out the properties of all neurons in the macaque brain); on the other hand, the physiological properties of the cells may have been at least in part modified by the evolutionary process.

Indeed, after recording neural activity from a total of more than 1,000 neurons in more than twenty patients, we now have evidence for the existence of mirror neurons in multiple areas of the frontal and temporal lobe of the human brain. A preliminary description of these data has been published in abstract form (Mukamel et al. (2007)). The patients were tested under a variety of experimental conditions: grasping execution (the patients grasped a mug either with precision grip or with whole hand prehension, following computer generated instructions), grasping observation (the patients watched videoclips of grasping actions, including both precision grips and whole hand prehension), execution of facial emotional expressions (smiling and frowning), observation of facial emotional expressions, and some control experimental conditions. In the frontal lobe, we found mirror neurons in all four medial frontal areas we recorded from: supplementary motor area (SMA) proper, pre-SMA, dorsal anterior cingulate cortex (ACC) and ventral ACC. The monkey homologues of these areas are anatomically connected with area F5, either directly, as is the case for area F6 (which is the homologue of pre-SMA), or indirectly, through F6 (Rizzolatti & Luppino (2001)). Thus, in principle this could be a system of mirror neurons that exerts some form of 'control' over mirror neurons located in the inferior frontal cortex (the human homologue of area F5).

Why do mirror neurons need such a control mechanisms? Given their physiological properties, the answer to this question seems obvious. If mirror neurons—which are neurons with motor properties located in premotor cortex, a brain region important for planning an action—fire uncontrollably during action observation, we may find ourselves imitating each other all the time. This is less than ideal. Indeed, even though imitation is quite pervasive in human behavior, we do not imitate all the time. It is necessary to have a neural mechanism of control to inhibit unwanted imitation. The neurons in these medial frontal areas seem to be doing exactly that. In contrast to the recordings performed in macaques, in which the large majority of neurons show

increase in firing rate, we found that only one-third of mirror neurons show firing rate increase in the human medial frontal cortex. Another third of these cells demonstrates decrease of firing rate, whereas the remaining third exhibited properties never before observed in the macaque brain. These cells—which I called super mirror neurons (Iacoboni (2008))—have opposite firing rate changes for action execution and action observation. The majority of super mirror neurons increase their firing rate during action execution and decrease their firing rate during action observation, while a minority shows the opposite pattern. This pattern of firing rate change suggests that the overall neuronal population activity of these medial frontal mirror neurons shows a majority of excitatory responses during action execution and a majority of inhibitory responses during action observation. Why? While increase and decrease in firing rate do not automatically translate in excitation and inhibition (it depends on the kind of neuron firing up or down), in the cortex 85% of the units are excitatory and only 15% inhibitory (Braitenberg & Schuz (1991)). Furthermore, single unit recordings tend to measure action potentials from relatively large units, that is, pyramidal cells that are typically excitatory units. Thus, it is likely that the firing rate changes recorded in our study represent a relatively simple mechanism used by the mirror neural system in the medial frontal cortex to control unwanted imitation.

We found relatively similar proportions of excitatory, inhibitory, and super mirror neurons in all medial frontal areas we recorded from. This suggests that this system is quite widespread in the medial frontal cortex. This makes sense if this system has the important role of inhibiting unwanted imitation. One does not want to confine such an important neural system in a restricted brain area. A focal lesion might determine the loss of this important function. In keeping with this idea, the rare neurological patients with uncontrolled imitative behavior have very large lesions in the frontal lobe (Lhermitte et al. (1986)); De Renzi et al. (1996)).

The role of super mirror neurons may be relevant not only to the control of unwanted imitation. The role of these complex mirroring responses may be relevant to the differentiation between self and other. While mirroring seems a clever solution that evolution has devised to facilitate learning and social interactions, mirroring also generates the problem of differentiating between actions of the self and of other people. The opposing pattern of excitation and inhibition for actions of self and others seems an elegantly simple neuronal mechanism that helps keeping this differentiation. Indeed, the medial frontal areas we recorded from have been often associated with self-processing and the self-other distinction in the imaging literature (Gusnard et al. (2001)) (Uddin et al. (2007)).

We also recorded from the amygdala, hippocampus, parahippocampal gyrus, and entorhinal cortex. Amazingly, we found mirror neurons in these areas too. The majority of the temporal lobe mirror cells were found in the hippocampal formation (hippocampus, parahippocampal gyrus, and entorhinal cortex), which is connected with the frontal lobe through the uncinate fasciculus and other cortico-cortico projections (Blatt et al. (2003)); Kondo et al. (2005); Lavenex et al. (2002); Mohedano-Moriano et al. (2007)). Mirror neurons have been originally interpreted as neurons

firing rate
nonstrates
ver before
r neurons
ind action
ite during
ι while a
gests that
ns shows a
inhibitory
firing rate
ie kind of
only 15%
gs tend to
ls that are
led in our
tem in the

per mirror
s system is
m has the
nfine such
ermine the
al patients
ontal lobe

unwanted
the differ
t evolution
nerates the
e opposing
ntly simple
dial frontal
self–other
7)).
gyrus, and
too. The
formation
connected
ico projec-
Aohedano-
as neurons

with motor properties (coding goal-oriented actions) that also have specific sensory properties, that is, they respond to the sight of somebody else's action. The main idea being that these cells help understanding the actions of others by mapping them onto the observer's motor repertoire. However, medial temporal lobe neurons are typically associated with higher-order visual properties and memory properties, but not with motor properties. Thus, the existence of medial temporal neurons with mirror properties requires a revision of the original notion of mirror neurons. The mirror neurons we recorded in medial temporal lobe are likely higher-order visual neurons that also discharge during goal-oriented actions (in this particular case, either grasping a mug or communicating an internal state with a facial expression). Thus, their functional role seems to be the mapping of our own actions onto our perception of the actions of others. Given the widespread presence of mimicry during social interactions, the mapping of our own actions onto our perception of others' actions may represent some form of 'neural expectation,' (Arbib & Rizzolatti (1997)) for instance the anticipation of seeing somebody else smiling in response to our own smile. Another possible interpretation of these medial temporal mirror neurons is that in order to have a *perceptual* (or maybe even conceptual) experience of our own actions, we need to invoke the neural activity typically associated with the perception of the actions of other people. This 'broader,' 'non-motor,' notion of mirror neurons also fits well with the only human neuron displaying mirroring properties that has been recorded in previous studies. This neuron in the anterior cingulate cortex seemed to be mirroring pain, rather than actions (Hutchison et al. (1999)).

4.5 Pervasive Mirroring

The empirical findings reported in the studies on mirror neurons in macaques have inspired a large literature on the theoretical implications of these cells. Obviously, mirror neurons have been invoked in support of theories of mindreading and empathy that put simulation processes at center stage (Gordon (1986); Gallese & Goldman (1998); Gordon (2005); Goldman & Sripada (2005); Goldman (2006a)). The human data demonstrating a correlation between activity in mirror neuron areas and the tendency to empathize with others have certainly reinforced this hypothesis. A classical objection to this view is that mirror neurons have properties that seem 'too simple' to account for complex forms of mindreading and empathy, since these cells seem really 'monkey see, monkey do' cells. This objection seems to ignore that two-thirds of mirror neurons in macaques are 'broadly congruent,' that is, they code for actions that achieve the same goal, but are not necessarily identical. Furthermore, the new data in humans show that mirror neurons have more complex and more flexible properties than previously thought. Mirror neurons in humans (and maybe in monkeys too) are also anatomically much more widespread than previously thought. All these combined physiological and anatomical features suggest a very sophisticated neural system that may support complex forms of mindreading and empathizing. Thus, while it has been

proposed that mirror neurons and mirroring may be critical for low-level forms of mindreading and empathy, but not for high level forms (Goldman (2006a)), the new data make it entirely plausible that high-level mindreading may also be based on neural mirroring.

Another theoretically important implication of the new findings on mirror neurons in humans is that for the first time we have convincing evidence that mirroring is not a property *exclusive of motor neurons*. Because mirror neurons were discovered in the motor systems, they have been widely interpreted as reflecting the mapping of the actions of others onto the motor repertoire of the self. This interpretation, combined with the dominant tendency to put the self at the center of pretty much everything (at least in the Western world) (Kagitcibasi (1996); Triandis (1995)), led to conceptualize mirroring as a way of connecting with other people by simulating that they are 'another self' (Gallese (2006)), that is, by assimilating the other to the self. However, if neural mirroring occurs in higher-order visual areas (typically activated when we see or remember *other people* involved in their own activities, for instance when they are smiling) while we perform actions we cannot see (for instance when we smile), we must conclude that mirroring is not only simulating others as self. What is this form of mirroring supported by higher-order perceptual areas as the medial temporal cortex? In order to better understand the nature of these temporal mirror neurons we compared the timing of the activation in frontal and temporal mirror neurons. We found that frontal and temporal mirror neurons activate at the same time. This simultaneous co-activation between distant units suggests some form of Hebbian learning (Iacoboni (2009a)). Indeed, the pervasiveness of imitation and other factors (for instance, the pervasive presence of mirrors) may facilitate the coupling of activation in motor cells in the frontal cortex and in visual cells in the medial temporal cortex. When I smile, you smile back to me. The repeated co-occurrence of these two events may shape mirroring properties in both frontal (motor) and temporal (perceptual) units, such that even when I smile all by myself for whatever reason, I evoke the sight of your smile through the firing of neurons in my medial temporal cortex. When the self acts, the self also perceives the other. Self and other become two sides of the same coin. As the two sides of a coin are worthless pieces of metal when separated, self and other also make little sense when separated. Maybe this is why empathy *feels* so powerful. As Montagne wrote: 'Everyone feels its impact, but some are knocked over by it. On me it makes such an intense impression, my practice is rather to avoid it than to resist it... the sight of another's anguish gives me real pain; and my body has often taken over the sensations of some person I am with. A persistent cougher tickles my lungs and my throat.'

If self and other become deeply united in empathy and mindreading, one would predict that people that do not empathize and do not read others' minds, also refuse or resist to read their own mind. Anthropological studies seem to confirm this prediction. A number of Pacific societies behave as if it is impossible to know what goes on in another person's mind. This is called the 'doctrine of the opacity of other minds.' This

attitude is amazingly extended to the self too. Indeed, when natives are asked to provide motivations for their own actions, they tend to resist any kind of intentional reading of what they have just done (Duranti).

We empathize effortlessly and automatically with each other because evolution has selected neural systems that blend self and other's actions, intentions, and emotions. The more we learn about neural mechanisms of mirroring, the more we realize that the distinction between self and other may be almost fictitious in many cases. We have created the self–other distinction in our explicit discourse, along with many other constructs that divide us. Our neurobiology, in contrast, puts us 'within each other.' This is a major revision of the long held tradition according to which our biology would lead us only to self preservation but we are able to rise above our biological make up to become 'social.' It is our biology that makes us social and empathic, even though only at implicit, non-propositional level, while our ideas often divide us and sometimes lead us to commit atrocities. Hopefully, the awareness of our neurobiological mechanisms for empathy that is finally reaching the explicit level of understanding ourselves will also change the deliberate, reflective discourse in our society (Iacoboni (2008); (Iacoboni (2007)).[1]