

PART THREE: THE CONSTITUTION OF THE SOCIAL MIND

*He who understands baboon would do more towards metaphysics than
Locke. —Charles Darwin, (1838), Notebooks*

6 Neurons that Mirror



Fabrimago (2008) Make-up. <http://www.flickr.com/photos/fabrimago/2188512994/>

Creative Commons License

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 Unported License.



2008 Benoit Hardy-Vallée
www.hardyvallee.net/POSC

6.1 Explaining the social mind

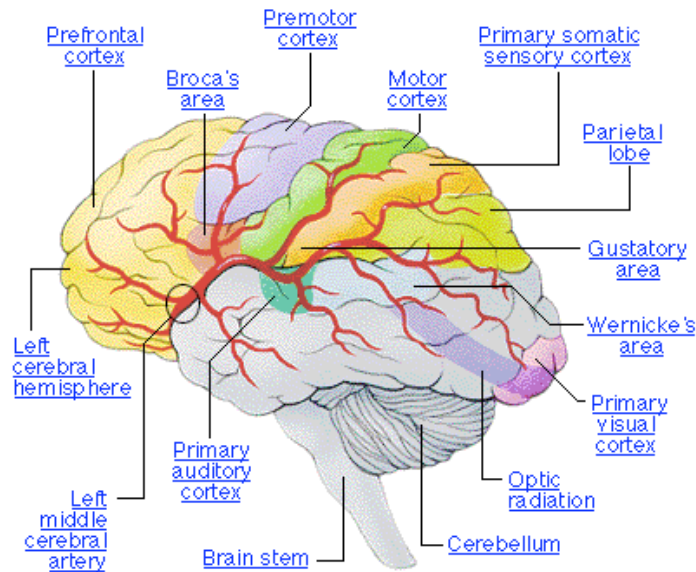
The past sections presented philosophical and cognitive perspectives on the social mind: the philosophical problems related to interpretation, and the different accounts of interpretation. In this section (chapter 6 and after), we will be mostly concerned by the biological foundations of our social-cognitive competence: the evolutionary path, the adaptive function and the neural mechanism that underlie our interpretive faculties.

Explaining biological phenomena, as Tinbergen argued, might be done by four means. For any phenomenon (a trait, a faculty, an organ, a behavior, etc.), one may lay out its:

1. **Mechanistic** component: how the phenomenon is causally generated
2. **Development**, or ontogeny: how the phenomenon unfolds in the individual's lifetime
3. **Evolution**, or phylogeny: how the trait evolved in the "Tree of Life"
4. **Adaptive function**: what the trait is for (if there is an adaptive function)

This chapter will focus on the neural mechanism of social cognition; however, since presenting these mechanisms would require an entire book, we will rather discuss a certain class of neural mechanisms that are particularly important for social cognition (and that generated many discussions in cognitive science and philosophy of mind): the so-called mirror neurons.

6.2 The brain



from http://www.medem.com/MedLB/article_detailb.cfm?article_ID=ZZZ0ZFP46JC&sub_cat=75

Before discussing mirror neurons, a little introduction to neuroscience and its object, the brain, may be relevant (a good introduction can be found online at <http://thebrain.mcgill.ca>).

The brain is the control center of the central nervous system. In human beings, it consumes 20-25% of the energy while it represents about 2% of the body mass (for a 150 pounds person). A 100 billion neurons are connected to between 1000 and 10,000 other neurons. This complex sensorimotor input/output device performs all psychological functions: sensation, pleasure, perception, emotions, reflex, motor control, motor planning, learning, representation, decision-making, imagination, reasoning, cognitive control, working memory, consciousness, etc. This mass of neurons shows a certain degree of specialization: brain areas tend to be involved in certain types of functions (e.g. the amygdala in fear, the prefrontal cortex in thought). However, as Anderson (2007) showed, the relationship between cognitive functions (e.g. perception, valuation, emotions, etc.) and neural structures is a many-to-many mapping:

1. **A cognitive function may recruit many brain areas.** For instance, social cognition recruits the anterior cingulate cortex, the amygdala, the prefrontal cortex, the medial temporal lobe, etc.
2. **A brain area may be involved in many cognitive functions.** The dorsolateral prefrontal cortex is involved in motor planning, organization, working memory, attention and many executive functions.

Thus the right attitude is not strictly *localizationist* (one-to-one mapping between function and structure) and not strictly *holistic* (functional equipotentiality in brain structures).

Neuroscience is the scientific study of the nervous system. Molecular, cellular, behavioral, and cognitive mechanisms are revealed through different means. Functional imaging technologies study degrees and locations of activation; single-cell recordings analyze the activity of individual cells, while lesion studies try to determine the function of certain brain areas by investigations of brain-impaired subjects. Computational neuroscience uses computer simulations to support or undermine hypotheses about brain mechanisms

According to the classic DN (deductive-nomological) model, one explains an empirical phenomenon by showing how it “fits into a nomic nexus” (Hempel 1965:488). If, from the statement of a scientific law or a universal generalization (called also a covering law) and certain conditions one can logically deduce an event, this event is then regarded as being explained:

(1) L (scientific laws or universal generalization)

(2) C (condition)

(3) E (event)

But this model is now considered to be an inaccurate picture of scientific explanations, for different reasons. Which law, for instance, explain why cats land on four feet? Or the effects of a drug? Explaining any phenomena, in most sciences, means showing how it fits in the causal, not the nomic, nexus (Salmon, 1998). We want to know why molecule X causes inhibition of molecule Y uptake in the rat’s brain. The explanation of the inhibition of molecule Y does not cite a law, but a causal mechanism. A mechanism is “a structure performing a function in virtue of its component parts, component operations, and their organization” (Bechtel and Abrahamsen 2005). Mechanisms are described by their initial, intermediary and final conditions. Initial conditions specify the parameters that make the unfolding of the operations possible; intermediary conditions are causal chains (cycles, branching, network); and final conditions are the state reached by the mechanism (production, elimination, equilibrium). A mechanistic explanation of a phenomenon is therefore a specification of how the components of the system produce the phenomenon: it reconstructs the pattern of causal interactions that lead to certain results that need to be explained. For instance, explaining how the pancreas regulates blood glucose levels involves explaining how alpha cells of the islets of Langerhans lower blood sugar by producing insulin while the beta cells raise blood sugar by producing glucagons. The production of these hormones can also be subject to a mechanistic explanation.

Explanations in neuroscience tend to be mechanistic, multi-level and causal. That is, to explain a phenomenon, at any level of explanation, requires usually a mereological (part-whole) decomposition in *entities* and *activities* (Craver) or *parts* and *operations* (Bechtel). Mechanistic explanations are not (completely) reductive: higher-level phenomena are not reduced, but explained by lower-level phenomena. Explaining consists not, as the logical positivists put it, in a logical subsumption of a phenomena under a law, but in constructing a model of the systems and sub-systems that lead to the causal production of a phenomena. Mechanistic explanations, also contrast with traditional functionalism or computationalist accounts (or “two-levelism” see Lycan, 1987). Instead of positing only two levels of explanations (mind and brain, related by implementation), mechanistic explanations put emphasis on multiple levels of structure and

functions, not just two. In order to explain spatial memory, one needs to decompose this phenomenon in systems, sub-systems, sub-sub-systems, etc. A rough sketch (from Craver 2002) shows how four levels of descriptions articulate entities and activities

1. Organism-behavioral: mice navigate through a maze to find food
2. Neuronal-computational: information processing, learning, connectivity of the hippocampus.
3. Synaptic-electric: neuron configurations (synaptic, dendritic) long-term potentiation
4. Molecular level-kinetic, attraction or phosphorylation by NMDA receptors

6.3 The Mirror Neurons System (MNS)

The **Mirror Neuron system (MNS)** is an action *execution/observation* matching system first discovered in the macaque *premotor cortex*, a brain area involved in motor planning. Visuomotor mirror neurons discharge both 1) when the monkey *executes* a motor act and 2) when it *observes* another individual performing the same or a similar motor act. The system does not react in presence of the object alone, the agent alone, the action without a target object, or the action performed by using a tool. Similar mirror phenomena have also been observed:

audio-visual mirror neurons react to sounds that indicate an action (e.g. nut cracking)

communicative mirror neurons: react to communicative facial expressions

phonological mirroring: hearing engages tongue muscles involved in speaking

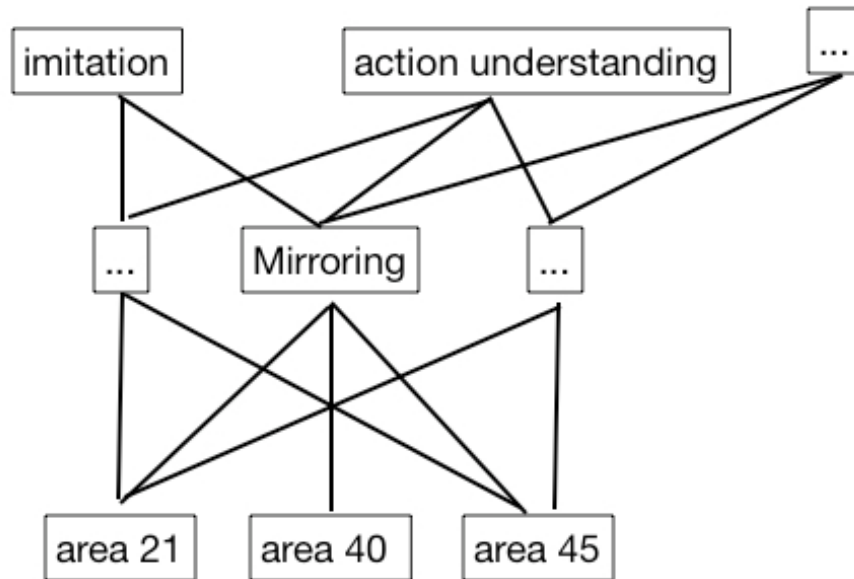
sympathy - emotional mirroring: experiencing and perceiving someone experiencing disgust involves the same areas, i.e., same sector of the anterior insula (Wicker et al. 2003).

Thus it has been suggested that the processes of production and understanding of at least simple actions share the same neural machinery. We use our neural resources involved in doing A to understand that someone else is doing A; we internally and effortlessly reconstruct the motor program that led to A:

Every time we are looking at someone performing an action, the same motor circuits that are recruited when we ourselves perform that action are concurrently activated.”
(Gallese & Goldman, 495)

Motor neurons are mirroring when they are activated by the perception of the action they typically perform.

Since the relationship between the mind and the brain is a hierarchical, multi-level nesting of systems and the relationship between cognitive function and brain function is a many-to-many mapping, we should understand mirroring as a “mid-level” mechanism: it is realized in many brain areas (in humans, in Brodmann areas 21, 40, 45) that also perform other functions, and the mirroring mechanisms participate in many higher-level functions.



6.4 Mirroring and interpreting

One of the primary functions of the MNS is production and understanding-through-reproduction of *actions* and emotional expressions. It is important to note that the MNS is sensible to actions, (movements that have a goal, performed with an intention) not mere movements. Mere visual perception does not provide information about the goal/intention behind the movement. Hence when one see a barking dog, imaging studies showed that it elicits only visual activity, not motor, because we cannot understand experientially (mirroring) the barking (see Gallese 2007 for a review). One might say that the category of actions is (almost) a natural kind for our brain. MNS systems are also necessary for **imitation**: reconstructing someone else's action requires a fast reverse-engineering of motor plan, a function subserved by the MNS.

According to Gallese, the MNS is also crucial for **language**. We mirror at the “vehicle” level (hearing a sentence activates our tongue muscle and phonological motor plans) and at the “content” level: our understanding of many concepts is based on sensorimotor simulation possibility (e.g. according to Lakoff and Johsson, we need the schema of force or the topological representation of a container to be able to grasp modal and set-theoretical concepts).

Finally, according to many researchers interested in the MNS, it is involved in **mindreading**. The fact that autistic persons have a deficient MNS supports this interpretation. More precisely, the MNS has been seen by many (such as Goldman) as a vindication of **simulation theory** (ST). Remember that, according to ST, mindreading is a method, a skill, rather than a theory; we “impersonate the mental life of the target” (Gallese and Goldman 1998) rather than theorize about it. We allegedly use “as-if”, pretend, fake desires, process them through our decision-making system, and generate hypotheses about the beliefs and desires of the target agent. In prediction, a forward simulation infers action from beliefs/desires; in retrodiction a backward simulation infers the beliefs/desires from the action.

The MNS, according to Gallese and Goldman, might be involved in retrodiction. The MNS first encodes plans or intentions and then tags them as “my” or “your” action, depending on whether the system was internally (when I perform the action) or externally (when I observe the action) activated. When we observe an action, motor areas are activated but the movement is not completed. Externally activated plans are largely inhibited, or taken ‘off-line’ (but not always, as people who talk a long time together and get along well often adopt the same postures without noticing it). Subjects with prefrontal lesions (and, to a certain extent, babies) compulsively imitate gestures, suggesting that imitation is a default mode. Thus mirror neurons would be neutral (regarding the authorship of the action): they would simulate first the intentional action (or “naked intentions”), and then the agent. The process allegedly goes through 1) observing behavior 2) registering naked intentions 3) attributing agency (“who” mechanism: me vs. the other).

The MNS seems to support ST because it shows that action understanding is arrived at through an “inner imitation”. The MNS, say Gallese and Goldman “seems to be nature’s way of getting the observer into the same ‘mental shoes’ as the target”. As Gallagher points out, however, intentions are always “already clothed in agency” (Gallagher, 2007, p.70). We never have to decide who performs the action: we just see who is. An action is always experienced as someone's action. Thus the mirror resonance+who mechanisms underlie inter-subjective *perception* (as a personal-level phenomenon) rather than *simulation*. We experiment a direct perception, not a step-wise process of mental replication:

The other person has an effect on us. The other elicits this [neural] activation. This is not a simulation, but a perceptual elicitation. It is not us (or our brain) doing it, but the other who does this to us (Gallagher, 2007, p. 73)

6.5 References

- Anderson, M. L. (2007a). Evolution of Cognitive Function Via Redeployment of Brain Areas. *Neuroscientist*, 13(1), 13-21.
- Anderson, M. L. (2007b). The Massive Redeployment Hypothesis and the Functional Topography of the Brain. *Philosophical Psychology*, 20(2), 143 - 174.
- Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., & Martin-Skurski, M. E. (2006). Neurobiological Substrates of Dread. *Science*, 312(5774), 754-758.
- Cartwright, N. (1989). *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.
- Craver, C. F. (2002). Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory. *Philosophy of Science*, 69, S83–S97.
- Gallagher, S. (2007). Logical and Phenomenological Arguments against Simulation Theory. In D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-Assessed* (pp. 63-78). Dordrecht: Springer.
- Gallese, V. (2007). Before and Below ‘Theory of Mind’: Embodied Simulation and the Neural Correlates of Social Cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 659-669.

Gallese, V., & Goldman, A. (1998). Mirror Neurons and the Simulation Theory of Mind-Reading. *Trends in Cognitive Sciences*, 2(12), 493-501.

Giere, R. N. (1999). *Science without Laws*. Chicago: University of Chicago Press.

Humphreys, P. (2004). *Extending Ourselves : Computational Science, Empiricism, and Scientific Method*. New York: Oxford University Press.

Lycan, W. G. (1987). *Consciousness*. MIT Press.

Ramachandran V.S., Oberman L.M. 2006 Broken mirrors: a theory of autism. *Scientific American*, 5: 62-9

Salmon, W. C. (1998). *Causality and Explanation*. New York: Oxford University Press.

Thagard, Paul, "Cognitive Science", *The Stanford Encyclopedia of Philosophy* (Summer 2007 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2007/entries/cognitive-science/>.

Tinbergen, N., On the aims and methods of ethology. *Zeitschrift fur Tierpsychologie*, 1963. 20:

Wicker B., Keysers C, Plailly J, Royet JP, 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