VITTORIO GALLESE
Università degli Studi di Parma
vittorio.gallese@unipr.it

NEUROSCIENCE AND PHENOMENOLOGY

abstract

This text contributes to a necessary dialogue, and possibly a translation of the different notions employed by neuroscience and phenomenology. This effort is particularly significant for cognitive neuroscientists whose main topic is social cognition and the related notion of intersubjectivity. What I qualify as "embodied simulation" (which exploits, not only but mainly, the intrinsic functional organization of the motor system) is a crucial functional mechanism in social cognition, not confined to the domain of action, but encompassing other aspects of intersubjectivity such as emotion and sensation. It is "embodied" because it uses a pre-existing body model in the brain: all the brain areas showing mirror mechanisms model our interaction with the world. This model of interaction, this praktognosia, turns out to be highly relevant not only when the task is to guide our own behavior, but also to understand the behavior of others. The very last part of the text is devoted to reply to some arguments against embodied simulation coming from phenomenologists themselves.

keywords

Cognitive neuroscience; embodied simulation; mirror neurons; peri-personal space; phenomenology; praktognosia

NEUROSCIENCE AND PHENOMENOLOGY

VITTORIO GALLESE *Università degli Studi di Parma*

I'd like to start by thanking Roberta De Monticelli and the Center of Research in Phenomenology and Sciences of the person of Università Vita-Salute San Raffaele – Milan for giving me the opportunity to be here and take part to this very interesting – and innovative, for Italy – enterprise. We all believe that this is just the beginning of a future series of meetings where people from different disciplines but nevertheless sharing the same questions will have the opportunity, like we are doing during these three days, to meet and discuss.

I'll start with few opening remarks to make it clear where I am positioned with respect to this specific topic: the relationship between Phenomenology and Neuroscience. As already kindly quoted by Roberta De Monticelli, few years ago I wrote this statement: "we should phenomenologize Cognitive Neuroscience rather than naturalizing phenomenology" (Gallese 2006, p. 294). Both enterprises at first sight might look a bit spurious at least to the phenomenological tradition, which during the early phases of its development was very critical towards an approach meant to psychologize the content of our phenomenal world. I must say that I am a bit sceptical too about the possibility to naturalize Phenomenology, especially if one aims to do that by translating Phenomenology in the quantitative language of Mathematics, which is one of the possible ways of accomplishing such naturalization, fully endorsed by some of the people who were at the origin of the now famous book "Naturalizing Phenomenology" (Petitot *et al.* 1999).

Why should we try to phenomenologize Neuroscience? Because if one of the aims of Cognitive Neuroscience is to shed light on the human condition, we certainly cannot but start from how the world is constituted within our own phenomenal appreciation.

Further, I strongly believe that a dialogue, an exchange of views, possibly an attempt to translate the different notions employed by both disciplines, Neuroscience and Phenomenology, should not only be hoped for, but it is necessary. Certainly it is necessary for the cognitive neuroscientists whose main topic of investigation is the notion of intersubjectivity. And I think that the research agenda of Cognitive Neuroscience in the near future should certainly encompass the first-person aspect of human experience, but also the personal characteristic of the individual subject of that experience. We shouldn't forget that much of what we know about brain function, when we are dealing with "garden variety" participants, most volunteers will come from a not better specified cohort of Psychology students – this is the truth.

Thus, I think we could do a lot more than that by trying to correlate the brain function with the personal life history of the individuals we are employing in the vest of volunteer to enquire, to understand more about the relation between behaviour, between the life of the mind and the contribution of the brain-body system.
What has to be done is to much better specify who are the volunteers we are convincing to enter the fMRI. We must correlate how the specific being-in-the-world of the individual can be translated into a specific way of functioning of her/his brain-body system.
The larger field I would like to enter now is that of the project of naturalizing social cognition and we learned from Michele Di Francesco how we should envisage such enterprise, what do we mean by naturalizing.
Let's start from the mainstream view: according to the mainstream classic cognitive view, action and intention understanding, which constitute a very important part of our intersubjectivity, consist in interpreting and explaining in mental terms the behaviours of others. These behaviours, according to this perspective, are intentionally opaque, because they consist of biological motion. So, the way someone moves or behaves or acts is intrinsically intentionally opaque unless I can identify a hidden internal mental state that most likely caused that behaviour, and it is through this sort of ascending routine starting from behaviour, but aiming to reach the hidden internal mental state, that I can possibly, and only in this way, make sense of the behaviour of the other. This explanatory process is referred to as "mind reading", that is, the attribution to others of internal mental states mapped in the mind of the observer as internal representations. For most people this representation is conceived as being implemented in a propositional format. Here is a quote from Alvin Goldman's (Goldman 2008, p. 3): "In other words, to mind read is to form a judgement, belief or representation".
The problem for people like me is to translate this perspective into the working of the tiny little things we have in our head – neurons – which collectively we name the brain. And here we should acknowledge that we know very little of what is going on, in spite of the fact that many colleagues of mine believe they really know what's going on. Because the problem is this: we do not have a clear neuroscientific model of how humans can understand the intentions and other mental states promoting the behaviour of others: what we have is a series of brain imaging studies showing the activation of a set of cortical regions, mesial frontal areas, the temporal parietal junction etc., during explicit mentalizing tasks. No one to date was able to provide a convincing explanation about why those specific areas do activate during mentalization, beside – I want to be very outspoken – the tautological

statement that mind reading is implemented in those brain areas. What we are left with is just a mere correlation which in science is not the ultimate goal, it is the beginning of the story, it is not the end. TQZhis is one out of many examples which really epitomizes this huge problem we are facing in Cognitive Neurosciences.

These tiny little squares shown on the mesial aspect of the frontal cortex portray activation foci activated in healthy volunteers while engaged in explicit mentalizing tasks, like for example attributing false beliefs to a fictitious character. The systematic activation of these parts of the mesial aspect of the frontal cortex let many scholars to make the statement: bingo, we found the home in the brain for the theory of mind module since what these people are doing is to mind read someone else. This part of the brain is systematically activated during explicit mind reading, hence it follows that this is where mind reading occurs at the level of the brain.

However, this is not the brain of a healthy individual, but it's the brain of a lady who suffered a bilateral damage to the supposed theory of mind module and the problem lies in the fact that this lady, in spite of having faced the complete destruction of the supposed theory of mind module, is fully competent in making sense of the behaviour of others (Bird *et al.* 2004). So that the authors of this paper conclude that their findings urge caution against using functional imaging as the sole method of establishing cognitive neuroanatomy. There is another problem: we are told that certain sectors in the brain are specifically active with a unique series of tasks, namely tasks that ask people to explicitly mentalize, mind read, attribute propositional attitudes to others. But we are not even sure of the specificity of this activation. One crucial area in the domain of the neural correlates of mind reading is the temporal parietal junction (TPJ), the *carrefour* at which the temporal and the parietal lobe merge. For many years authors like – I quote one for many – Rebecca Saxe, made the claim that this area is solely and uniquely active in mind reading tasks. We learn – from the work of other colleagues of Saxe, for example, Jason Mitchell who is at Harvard – that this is not true, because the very same TPJ which is active when I attribute false beliefs to others is also active when I am engaged in an attentional task which has nothing to do with explicit mind reading. Even sexual arousal can lead to its activation. Well, we can always say that there are highly sophisticated persons who can get sexually aroused only through mind reading but – I mean – I would rate this as a very weak argument.

So, beside the reification of mentalistic notions like belief and intention into things to be found at specific locations in the brain, the main problem with this approach consists in the fact that the mind reading specificity of

this activation is at best highly debatable. Then I ask myself and ourselves a rhetoric question: do we really believe there are mind reading specific neurons in the brain? We know what a neuron is all about: basically is a machine producing action potentials. Mentalizing though, whatever it is, is a personal level competence and therefore it cannot be fully reduced to the subpersonal activity of mind reading specific clusters of neurons wherever they might be located. We should remind ourselves that neurons are not epistemic agents, insofar as they only "know" about the ions passing through their membranes, giving them the property they have, being excitable cells. Neurons, and of course *a fortiori* mirror neurons or brain areas, are necessary but not sufficient conditions for mentalizing, because in order to mentalize we need an individual. Let me offer you my reductionist definition of what an individual can be reduced to: a properly wired brain-body system interacting with a specific environment populated by other brain-body systems. A further aspect I would like to emphasize is that when we refer to the brain we should always bear in mind that the brain is not a sort of magic box, the brain is a subpersonal constituent of our body which gets all the possible information about the so-called external world through an interface which is the body, which is a very peculiar interface which develops to acquire the given shape we now have because of the adaptation to a specific environment which obeys to specific physical laws etc. So it is a situated brain-body system, the one which is the target of the investigation of people like me.
I endorse a "bottom-up" approach to social cognition and I would like to start by quoting the authors of the book which is the main topic of this Winter School: "The other is given in its bodily presence as a lived body" as Gallagher and Zahavi write, "a body that is actively engaged in the world". "Empathy is defined as a form of intentionality in which one is directed towards the other lived experiences [...]". "In empathy we experience the other directly as a person, as an intentional being whose bodily gestures and actions are expressive of his or her experiences or states of mind" (Gallagher-Zahavi 2008, p. 183). Great!
My problem starts now: how do we accomplish this? My point: at the basis of the capacity to understand others' intentional behaviour in such a direct way, both from a phylogenetic and an ontogenetic point of view, there is a basic functional mechanism which I qualify as *embodied simulation* which exploits, not only but mainly, the intrinsic functional organization of the motor system. So the natural evidence of the world stems from our potentialities for action: this is a topic that I assume will be much more detailed and developed in the following talk given by Corrado Sinigaglia. Here is a quote from Aron Gurwitsch, in a paper which appeared in 1941: "The world in which we live and act is peopled with items endowed not only with colors, warmth, smells,

shapes etc., but also with qualities like attractive, repulsive, agreeable, [...], fit for some purpose or other, and so forth. In this world there are actions done or to be done, and these actions deposit themselves like qualities upon the things with which they are connected" (Gurwitsch 1941, p. 328). Merleau-Ponty: "my body appears to me as an attitude directed towards a certain existing or possible task. And indeed its spatiality is not, like that of the external objects or like that of *spatial sensations*, a *spatiality of position*", we are not dealing with geometrical space here, "but a *spatiality of situation*" (Merleau-Ponty 1962, p. 100). And a few pages later he develops this and clarifies what by spatiality of situation one has to mean. The body "provides us with a way of access to the world and the object, with" what he denotes, refers to, as "a *praktognosia*, which has to be recognized as original and perhaps as primary" (Merleau-Ponty 1962, p. 140).

And just to give you a flash of how true I take to be such a statement, an example comes from the investigation of one of the many parallel cortico-cortical networks that reciprocally connect distinct fields within the posterior parietal cortex and the ventral premotor cortex of the macaque monkey brain. We are dealing with one specific field within the premotor cortex of the macaque, area F4, which is immediately facing the primary motor cortex, so, if you recall the neural properties of the neurons sitting there, these are motor neurons and they control the execution of goal-related purposeful movements like stretching out the arm to reach something within the peripersonal space or turning the head to orient towards or move away from something which is approaching the macaque monkey's body. These neurons combine motor properties – they control the reaching of the arm or the orienting or avoidance movement of the head – with tactile properties on the same region whose movements the same neuron controls, and visual properties. But the most interesting feature of these visual properties is that the visual receptive field, the part of the visible space observed by the monkey that is effective in driving the discharge of the neuron, is not only bidimensional, but has also the dimension of depth. A visual stimulus is effective in driving this motor neuron only if it is presented within the peripersonal space. Peripersonal space is the outcome of the motor potentialities our body instantiates. Premotor neurons map objects within peripersonal space as potential targets and, if you allow me, we could define these objects in heideggerian terms as "*zu-handen* objects" of goal directed motor behaviour. Premotor neurons accomplish such mapping by means of simulation because when the neuron that normally controls the reaching of my arm is activated by a visual stimulus within my reaching space, within the potentiality of reaching of my arm, then, that object activates a motor program which nevertheless doesn't lead to an overt movement on the side

of the observer. The observer is merely watching that object approaching to a given body part. This motor activation devoid of any actual movement, at least in my opinion, can fully qualify as a motor simulation, because it is the activation of a motor neuron which nevertheless doesn't lead to produce overt movement. And here there is another quote of a phenomenologist – Jan Patocka – who moves one step further with respect to Heidegger, giving to the practical knowledge of the world envisaged by Heidegger, the flash and bones of the living body, so to speak. Patocka in *Body, Community, Language, World* writes: "Our primary experience of ourselves is an experience of primordial dynamism that manifests itself in our awareness of our existence as a moving, active being. This dynamism appears as distinctively linked to that which orients us in our movements [...] in such a way that our energy is always focused on something, on what we are doing" (Patocka 1998, p. 40).
So, if we go back to intersubjectivity, the conclusion we may provisionally draw is the following: we should abandon the Cartesian view of the primacy of the ego and adopt a perspective emphasizing the fact that the other is co-originally given as the self. Both self and other appear to be intertwined because of the intercorporeity linking them. Intercorporeity describes a crucial aspect of intersubjectivity not because the latter is phylogenetically and ontogenetically grounded on a mere similarity between our body and the body of others, because the pictorial similarity, as we will see in a minute, is not the crucial point here. Intercorporeity describes a crucial aspect of intersubjectivity, because we and others all share the same intentional object – obviously to a certain degree - and our situated motor systems are similarly wired to accomplish similar goals. It is the sharing of the same situatedness and the sharing of the same intentional goals that makes intercorporeity a privileged access to the world of the other.
What we learn from Cognitive Neuroscience is the following: the motor system of primates is primarily organized not in terms of movements, but rather in terms of goal-directed motor acts. Goal-directed motor acts are the nuclear building blocks around which action is produced, perceived and, to a certain degree, at least, understood. What makes of a movement a motor act? The presence of a goal. The same movement can accomplish different goals – a typical example I am always reiterating is that of flexing the finger of my hand. So, this is a movement. However, this movement can be put into service of completely different goals: scratch my hand, grab this mike, say hello: the movement is the same, the purposes are different. For many years we thought that the defining feature of the motor system is to enable the execution of movements; in order to attain to the teleological level we should put into action other parts of our brain. Most likely, the most anterior you go, the more cognitive you get, people claim, so, probably the pole of the

frontal lobe is where goals are living and where goals are communicated to the motor system whose main task is to set different muscles into action, thus enabling movement. I believe this characterization of the motor system is wrong or, at best, captures only one aspect of the functionality of the motor system. Why do I say that?

Because now we have the empirical evidence enabling us to justify the statement I've just made, namely that the motor system is best understood as enabling the accomplishment of motor goals. Neurons recorded in the ventral pre-motor cortex area F5 by Giacomo Rizzolatti and co-workers in the late eighties of last century (Rizzolatti *et al.* 1988), have the distinctive feature of being driven by goal-related motor acts like grasping a piece of food, with the mouth, with the right hand, or with the left hand. Different movements, different effectors. Nevertheless, these different movements lead all the same to the activation of the very same neuron. The great discovery of Giacomo Rizzolatti was to interpret such a firing as the outcome of goal coding: so these neurons do not code movement, they code the goal of taking possession of an object, no matter if with the mouth, right or left hand. We can resist, and indeed many people resist this hypothesis, many people still resist this interpretation, even after the data I am going to present.

You can reason in the following way: why do you need to invoke mentalistic notions like that of *télos*, of goal? These neurons could be easily interpreted in a much more parsimonious way by making the statement: what these neurons are doing is sending divergent inputs to motor neurons sitting in the primary motor cortex which in turn control the closure of different effectors, the mouth, the right hand or the left hand. So we wouldn't need to invoke the notion of a goal to make sense of these neurons. How can we falsify this parsimonious interpretation? By dissociating the movements from the goal, and this is exactly what we did.

In this experiment (Umiltà *et al.* 2008) we have a dissociation: in order to accomplish the goal of grasping something, the monkey is no longer required to close the finger, but to open them while using a reverse type of pliers. The same neuron that fires when the monkey grasps with the bare hand or with normal pliers, in which goal accomplishment always coincides with finger flexion, also fires when the goal is accomplished by performing exactly the opposite type of movement, namely extending the fingers instead of flexing them. These neurons implement a goal state motor representation whose content is both intentional and motor in nature. Why intentional? Because it is a goal-centered motor representation that although referring to movement, cannot, as we have seen, be reduced to a single sequence of movements. But it's also motor because the goal is mapped in motor terms. In other words, as the end point of a motor act, and although this motor

representation can differ with respect to single movements, nonetheless it must have a coherent motor content enabling it to determine a given behaviour and to control its execution, because we are in the motor system. Mirror neurons add to this property of goal relatedness the multimodal property of being active not only when the goal is accomplished by the owner of the mirror neuron, but also when a similar motor goal is accomplished by someone else.

So, the neuron fires both when the monkey grasp, but also when the monkey is witnessing the grasping being made in front of it by another individual, being a human being or a monkey. There is the coding of a similarity which is not a similarity of shape – curiously Edith Stein uses specifically the example of the hand – so a child hand, a monkey hand, a human adult hand is always a hand. What makes of this physical entity a hand? Well, among other things, the possibility, the intrinsic possibility of being instrumental to the accomplishment of a motor goal.

Here I want briefly to condense some novel findings on mirror neurons: the first is that the motor goal-relatedness of grasping motor neurons also applies to mirror neurons. As demonstrated by the recent study of Rochat (Rochat *et al.* 2010), the same mirror neurons that fire when the monkey grasps by hand, or with reverse pliers, also fire when observing a grasp being made by a hand or by means of reverse pliers. It's symmetric. The peak of the discharge is reached by mirror neurons at the accomplishment of the goal. However, if you confront the level of activation when the action is executed by the monkey, with respect to when the monkey is observing the action being performed by the other – although the pattern of discharge is identical, the intensity of the discharge in the latter case is significantly lower. This is crucial – I will come back to this point – because we have a neural mechanism that is mapping the same goal, no matter who is accomplishing it. However, mirror neurons react differently when the action is actively executed with respect to when it is only observed.

But there is more. The monkey, in order to perform this task, needs to be trained. So we trained these monkeys for five months, in order to teach them how to control these crazy instruments. However, we investigated the discharge of mirror neurons also during the observation of objects being "grasped" by means of sticking them, a motor strategy monkeys never learned to execute. If you compare the temporal development of the discharge of mirror neurons during action observation, you see that the discharge occurs much earlier on to the extent that the observed action belongs to the motor repertoire of the monkey. The more the observed action belongs to the motor repertoire of the observer, the earliest occurs the response of the mirror neuron. This finding suggests two things: 1) motor

expertise plays a role in modulating the response of the mirror mechanism; 2) the mirror mechanism has the potentiality to generalize, to map the very same goal also in observed behaviours that monkeys are not capable of performing.

Similar results were recently obtained in humans by Cattaneo and colleagues (Cattaneo *et al.* 2010) with transcranial magnetic stimulation (TMS). By means of TMS one can stimulate in a somatotopic way the motor system. According to where one places the stimulating coil, a very strong magnetic field excites a relatively limited population of neurons sitting below the location where the stimulator is placed. So, if you put the stimulator in correspondence with the region of the primary motor cortex that controls the movement of distal muscles, you can evoke a twitch in the muscles being controlled by the spinal neuron which in turn is controlled by the cortico-spinal neuron that you are stimulating. So, you are activating the piramidal tract and you lead to the activation of a muscle. Here, this kind of experiment, wants to learn how much the excitability of the motor system is influenced by viewing someone else doing something. In this particular case, people while being subjected to the stimulation were looking at an actor either playing with reverse pliers, similar to that employed by the monkey, or accomplishing a goal, taking possession of an object. To make a long story short, the motor facilitation has a completely different profile, according to the stimulus that was observed during the stimulation. If you stimulate the motor cortex while participants are watching a movement, opening and closing the gripper, you see a facilitation which is congruent with the movement: you facilitate the flexor muscles during the opening phase and you facilitate the extensor muscles during the closing phase. In this case motor facilitation is synchronous with the observed movement. However, if you are observing gripper being operated to accomplish a goal, e.g. to grasp a peanut, the facilitatory effect coincides with the accomplishment of the goal. The observation of a *tool movement* activates the cortical motor representation of the hand *movements* involved in the observed motor behaviour, but the observation of the *tool goal-related motor act*, activates a cortical representation of the observed *motor goal*, irrespective of the individual movements and the order required to accomplish the very same goal. In conclusion, we see in the human motor system the very same effect that we were able – with a much higher level of granularity – to measure in the macaque monkey brain.

Next is another recent experiment (Caggiano *et al.* 2009) that was done in joint collaboration between the University of Tubingen and the Università degli Studi di Parma. This study demonstrates that the firing of mirror neurons is modulated by the proxemic relationship between actor and observer. You already know what peripersonal space is, it's a motor space, the

space I can reach by outstretching my arm. When we first described mirror neurons, we didn't systematically test if peripersonal modulation could be uncovered, which is specifically what is at stake here. In this study the action observed by the monkey is executed by the agent either within or outside the peripersonal space of the monkey. These authors discovered that while 50% of mirror neurons are not modulated by the distance at which the observed action occurs, the remaining 50% is modulated by this manipulation. So there are mirror neurons that are driven by the observation of the grasping only if it occurs in the extrapersonal space of the monkey, and others showing the opposite modulation. Some neurons are even more interesting, because they don't fire if the experimenter grasps within the peripersonal space of the monkey. However, if the grasp is observed being performed at the very same spatial location but with the interposition of a transparent barrier which makes that object unreachable by the observing monkey, these neurons resume their activation. Although, geometrically speaking, the location is the same that before was totally uneffective in driving the cells, the potentiality for action is strongly modulated by this interposition of a transparent barrier and this manipulation of the potentiality for action of the observer modulates the discharge of the very same neuron. I think that this is very important and could lead to a new line of research on humans and there are people actively seeking to investigate this effect by means of virtual reality in humans.
To sum up: mirror neurons discharge when the action is executed or observed and when the consequences of the action can only be predicted. Even the noise produced by the action is sufficient to specifically trigger mirror neurons, so the neuron that fires when the monkey breaks the peanut also fires when the monkey sees someone else breaking the peanut but will also fire when the only sensory information the monkey receives is the noise of the peanut being broken by someone else. We believe that all these properties entitle us to interpret the mirror mechanism as a non-metarepresentational form of action understanding. Or better: this mechanism underpins, enables, makes it possible for us to directly understand the *what*, the motor goal, and – at least to a certain degree – also the *why*, the motor intention, of the observed behaviour of others. By motor intention I don't imply the reasons causing a given behaviour. This is a grasp, ok? So mirror neurons, in my opinion enable us to understand directly that such biological motion is a grasp. Why do I grasp this? To drink, and we believe that this is still within the potentialities of this mechanism. Why do I drink? This is beyond, at least so far we have no empirical evidence enabling us to claim that the why of this motor intention is within the coding capabilities of the mirror mechanism. Because I am thirsty, because I am anxious and my salivation is reduced to zero and then I need to restore the hydric equilibrium within my mouth. I mean, this is beyond, but the what and the motor why, we

believe, are within the mirror mechanism capabilities, is something that this mechanism can buy you.

Let's turn now to the mirror mechanism in humans. A meta-analysis of several brain imaging studies show that different regions within the premotor and posterior parietal cortices of the human brain are activated both during motor execution and action observation. Furthermore, there are other brain regions that show the same double pattern of activation. Not anymore in the domain of action, but in the domain of emotion or sensations. There are cortical sites that are activated during both the first-person experience and the observation of emotions or sensations. My model aims to provide a coherent framework and functional explanation of this variety of mirror mechanisms that nevertheless seem to share something: a functional mechanism that I characterize, unfortunately it turned out, as 'embodied simulation'.

What is embodied simulation in short? Is a crucial functional mechanism in social cognition, not confined to the domain of action, but encompassing other aspects of inter-subjectivity such as emotion and sensation. Why embodied? Because it uses a pre-existing body model in the brain. I mean: all of these brain areas showing mirror mechanisms, what they are doing? They model our interaction with the world. Well, this model of interaction, this *praktognosia*, turns out to be highly relevant also when the task is not to guide our own behaviour, but to interpret, to decode, to understand, the behaviour of others.

Why simulation? Because we have an isomorphic representational format – indeed we map the actions of others onto our own motor representation as well as other emotions and sensations onto our visuomotor and sensorimotor representation. My disgust: I activate my insula. Your disgust: I still activate my insula. The term "representation" of course is employed in a very different way from its standard meaning. I refer, when I speak of motor representation, to a type of content which is generated by our interactions with the world, that is pre-theoretical and pre-linguistic, but that nonetheless has attributes or some attributes normally referred to conceptual content (Gallese 2001, 2003).

The very last part of my talk is devoted to reply to some arguments against embodied simulation. Not those coming from the field of Classic Cognitive Neuroscience, but, unfortunately, from phenomenologists. Life is never easy. Okay, here at the end I quote Dan Zahavi and Shaun Gallagher. "The sub-personal simulation process", – they refer here to Alvin Goldman – "like its explicit cousin", embodied simulation, "involves a multi-step process. First, we perceive a certain behaviour; this is followed immediately by activation of shared representations – in neutral mode; and this is followed by a determination of agency" (Gallagher-Zahavi 2008, p. 178). My reply:

such steps are unwarranted both at the phenomenal level – and we all agree on that – but also at the sub-personal level. I dispute that a given behaviour, like my grasping the microphone, can be directly perceived as such unless by evoking the activation of its motor content. What makes of the physical displacement of this physical entity which you call "my hand" contacting this other physical entity, the microphone, a grasping hand, is motor simulation. You must activate the motor system in order to have a direct appreciation of this biological motion as a grasping hand. But since the activation of such motor content occurs in the brain of the observer without any explicit movement, such activation in my opinion qualifies as motor simulation.
But apparently there are more simulation troubles. *Simulation trouble* (2007) is the title of one paper from Gallagher, but the same argument I think is employed in the joint co-authored book we are discussing today. According to Gallagher and Zahavi the phenomena that occur during the activation of mirror neurons should not be understood as a simulation for multiple reasons. These reasons include the "as if" quality of mirror mechanism, reportedly at odds with the fact that mirror neurons map intentional relations in a fashion that is neutral about the identity of the agentive/subjective parameter.
My reply: it is certainly true that mirror neurons fire no matter whether the action is executed or perceived. However, it is also true that the intensity of their response is not the same in these two different situations. More generally, embodied simulation doesn't imply that we experience others the way we experience ourselves. The I-Thou identity relation constitutes only one side of the inter-subjectivity coin. The cortical circuits at work when we act, neither completely overlap nor show the same activation intensity as when others are the agents and we are the witnesses of their action. The same logic applies to sensations and emotions. There is a very recent paper by Christian Keysers and co-workers (Jabbi *et al.* 2008), whose main target is the appreciation of disgust in others. So, again, like in our original experiment, voluntary participants in the fMRI are either being subjectively disgusted – this time by tasting some disgusting liquid – or view the expression of disgust of someone else or – third condition – read a narrative about somebody else's disgust. The results of the experiment are the following: no matter whose disgust is at stake, the anterior insula always activates. At the very same location that we discovered in our original disgust experiment (Wicker *et al.* 2003). However, this is only one part of the story. Together with this common, shared activation focus, there are other brain regions which uniquely activate during my disgust but not during your disgust or the disgust of a fictitious character in the narrative, and the other way around.
Second, I don't share Gallagher and Zahavi's view that embodied simulation

– this is a more philosophical argument – must necessarily be characterized as simulation exclusively based upon the resemblance between target and simulator. As argued by the late Susan Hurley simulation can be more plausibly characterized in terms of *reuse* (Hurley 2005). According to the reuse notion of simulation, which I advocate, what distinguishes embodied simulation from theorizing is the reuse of a process for generating knowledge about that process. Neuroscientific evidence shows that motor neurons reuse motor processes enabling direct understanding of the actions of others and, similarly, viscero-motor or somatosensory neurons reuse emotion- and somatosensory-related processes enabling direct understanding of others' emotions and sensations.

What qualifies simulation as embodied is specifically this notion of reuse describable as an isomorphic type of mapping between target and simulator. What makes the activation of mirror neurons or mirroring mechanisms in the human brain during the observation of the actions of other an "as if" process, is not its resemblance aspect. This cannot be the case because we have seen at the level of the single neuron that there is nothing resembling the movement of the hand operating with the reverse pliers with the way macaque monkeys' hand appears when accomplishing the same goal. What makes the activation of mirror neurons during the observation of the actions of others an "as if" process is not the resemblance aspect, but the fact that in spite of an activation of the motor system in the observer's brain, the action is not executed. Thus, I cannot find a better word than qualifying it as a "motor simulation". This is the reason why I disagree with the claim that in order to invoke simulation mirror neurons "must generate an extra copy of the actions as they would be if they were the perceiver's own actions" (Gallagher 2001, p. 102). Mirror neurons' activation is the activation of the motor representation of that action, so there is no extra step to be invoked in order to explain it.

That said, I think that our perspectives share a lot more than what transpires from Gallagher and Zahavi's critique of embodied simulation. We all think that mind reading should not be identified with the mostly theoretical enterprise usually defined as Theory of Mind. We all think that the primary way of understanding others is direct in nature.

However, I do believe, pace Gallagher and Zahavi, that such directedness is completely compatible with the reuse notion of simulation I am advocating. Claiming that the understanding of others is mediated by mirror-based embodied simulation is not tantamount to saying that a sort of pretence mediates the perception of others' behaviour. All of these considerations make it difficult to account for mirroring phenomena as form of "direct perception". Social cognition is not only explicitly thinking about the

contents' of someone else's mind. Our brain has developed a basic functional mechanism which I qualify as embodied simulation which gives us an experiential insight of other minds. The specific nature of such experiential insight is still very loosely defined, and there is a lot of very meticulous philosophical work in parallel with neuroscientific work to much better specify what we qualify as "experiential". The shared "we-centric" space mapped by mirror neuron mechanisms generates the implicit certainties – which I think somehow overlaps with the notion of background as spelled out by John Searle – we entertain about others.

Before and below mind reading is intercorporeity as the main source of knowledge we directly gather about others (Gallese 2007). Embodied simulation is one crucial ingredient of intercorporeity. The social cognitive endowements of our species could be the evolutionary outcome, the exaptation of mechanisms that are not mind reading specific. The I-Thou relation provides the basic ground for our cognitive/affective development, hence for our intimate being social individuals capable of mutual recognition and understanding. I would like to end with a quote from a book that should perhaps be retranslated in Italian, *Ich und Du* by Martin Buber (1923): "In the beginning is relation". By empirically investigating the ontogeny of action you don't know how true this statement is. But this is a different story. Thank you.

REFERENCES

Bird C. M., Castelli F., Malik O., Frith U. and Husain M. (2004), *The impact of extensive medial frontal lobe damage on "Theory of Mind" and cognition.* "Brain", 127 (Pt. 4), pp. 914-28.

Caggiano, V. *et al.* (2009), *Mirror Neurons Differentially Encode the Peripersonal and Extrapersonal Space of Monkeys*, "Science" 324, p. 403.

Cattaneo, L. *et al.* (2010), *State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal and premotor cortices. "Cereb. Cortex*"; 20 (9), pp. 2252-8.

Gallese, V. (2001), *The "Shared Manifold" Hypothesis: from mirror neurons to empathy. "Journal of Consciousness Studies":* 8, N° 5-7, pp. 33-50.

Gallese, V. (2003), *The manifold nature of interpersonal relations: The quest for a common mechanism.* "Phil. Trans. Royal Soc. London B", 358, pp. 517-528.

Gallese V. (2006), *Corpo vivo, simulazione incarnata e intersoggettività: Una prospettiva neurofenomenologica*, in Cappuccio, M. (2006), Neurofenomenologia, Bruno Mondadori Milano, pp. 293-326.

Gallese V. (2007), *Before and below "theory of mind": embodied simulation and the neural correlates of social cognition.* "Philos Trans R Soc Lond B Biol Sci.", 362, pp. 659-669.

Gallagher S. and Zahavi D. (2008), *The Phenomenological Mind*, London.
Gallagher, S. (2001), *The practice of mind: Theory, simulation, or interaction? in "Journal of Consciousness Studies"*, 8 (5-7), pp. 83-107.
Goldman, A.I. (2008), *Mirroring, Mindreading and Simulation, in J. Pineda (ed.), Mirror Neuron Systems: The Role of Mirroring Processes In Social Cognition, Humana Press,* 2008. Online at: http://fas-philosophy.rutgers.edu/goldman/Mirroring,%20Mindreading,%20and%20Simulation.pdf
Gurwitsch, A. (1941), *A Non-Egological Conception of Consciousness*, in "Philosophy and Phenomenological Research", Vol. 1 (1941), pp. 325-38.
Hurley, S. (2005), *Perspectives on imitation: from neuroscience to social science*, Cambridge.
Jabbi M, *et al.* (2008), *A Common Anterior Insula Representation of Disgust Observation, Experience and Imagination Shows Divergent Functional Connectivity Pathways.* "PLoS ONE" 3(8): e2939.
Merleau-Ponty, M. (1962), *Phenomenology of Perception*, London.
Patocka, J. (1998), *Body, Community, Language*, World, Chicago.
Petitot J. *et al.* (eds.) (1999), *Naturalizing Phenomenology*, Stanford.
Rizzolatti G. *et al.* (1988), *Functional organization of inferior area 6 in the Macaque monkey. II. Area F5 and the control of distal movements*, in "Exp. Brain Res.", 71, pp. 491-507.
Rochat, M.J., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G. and Umiltà, M.A. (2010), *Responses of mirror neurons in area F5 to hand and tool grasping observation.* "Experimental Brain Research", 204(4), pp. 605-616.
Umiltá, M.A. *et al.* (2008), *When pliers become fingers in the monkey motor system*, in "Proc. Natl. Acad. Sci", USA, 105, pp. 2209–2213.
Wicker, B. *et al.* (2003), *Both of Us Disgusted in My Insula: the Common Neural Basis of Seeing and Feeling Disgust.* "Neuron", 40, pp. 655-664.

VIDEO ON-LINE
Winter School Milano Vittorio Gallese part I
Winter School Milano Vittorio Gallese part II