The second-person approach:
Implications for a realistic phenomenology of social cognition

Michele, Giuli

University of Turin
michelegiuliwriter@protonmail.com
DOI: 10.22034/IYP.2021.245793

Abstract

This paper analyses social cognition by considering the analytic philosophy of mind, neurophenomenology, and social neuroscience. Many social neuroscientists rely unconsciously on different philosophical answers to the question, "How do we understand each other?". Consequently, we will compare the principal intellectual and experimental approaches to social cognition proposed so far and join them in an integrationist account by considering the direct embeddedness of social interactors.

First, the "theory theory" (T.T.) affirms that mindreading involves inferring the other's mental state by observing his behavior from a third-person perspective. A neural network called the "mentalizing

system" (M.E.N.S.) underlies mindreading activities.

Second, the Simulation Theory (S.T.) assumes that social cognition involves simulating the mental states of the other. The "mirror neurons system" (i.e., M.N.S.) is the neural substrate for the simulatory activities. T.T. and S.T. are fastened to the "observer paradigm" since the experimental set-ups detect a participant's brain's activity observing or simulating someone else's movement, and intersubjective dynamics are not at play.

Finally, the second-person approach invites us to consider the other as the one who directly intervenes in our perception and is responsible for the meaning we assign to his mental states (cf. Schillbach et al., 2013). Consequently, Schilbach et al. (2013) have established an experimental setting that is "minimalist and naturalistic" because it

focuses on fundamental embedded interactions such as mutual gaze.

This paper argues that the philosophical theories underlying those approaches do not conflict with each other, but they highlight different moments of social interaction in real life. Indeed, their neural substrates partially overlap. Hence, we want to establish in which order these three moments of social interaction occur. We hold that a realistic phenomenology must consider second-person interactions as the beginning of a realistic phenomenology.

KeyWords: mindreading, neuroscience, neural, substrate, embeddedness.

Introduction: The “veil of Maya” in the classic accounts of social cognition.

In Western intellectual tradition, the paradigms for inquiring social interaction have principally referred to a detached observer not actively engaging with other agents. The setup of recent experiments often consists of participants who observe others' behavior and try to infer their mental states (beliefs, desires, and intentions) (de Bruin, 2012). Hence, the notion of reciprocity is not at play because participants are not performing a "joint action"; social understanding then is a solipsistic activity conducted by a spectator that ponders the mental states of others.

Secondly, such accounts of social cognition imply the representational theory of mind.[1], which maintains the surroundings, and the others are never directly experienced. Contrarily, our perception of the observed interactors consists of mediating mental pictures. Moreover, the content of experience is conveyed to the subject by "intrinsic qualities" (Slors et al., 2015, p. 78), e.g., repugnance or beauty of people's facial expressions, which are subjective and contrast with "extrinsic properties," which are objective, physical and relational, e.g., Mark is bigger than Sara. If we never perceive others but only a mental representation of them, then it is possible that the intrinsic qualities of our experience deceive us. Indeed, the classical approaches to social understanding are intrinsically skeptical. Representation of others is limited and needs to be clarified; hence, there is a gulf between our impressions and their actual feeling or beliefs. Therefore, we must engage in an "intellectual detour" to bridge the gap between immediate experience and the other's psychological states (see Asch 1952:144–50).

At this point, an epistemic problem arises. How can we justify that, in ordinary cases, we can grasp others' feelings at a certain degree of immediacy? If we do not perceive the interactor directly, inferring her mental state through a sophisticated intellectual detour seems awkward. A scientific account of social interaction should address the direct perception of others that in the classical approaches needs to be included. The "observer-paradigm" entails a "veil of Maya," which separates the social actors and renders the more basilar interactions hardly intelligible and more mentalistic than they are.

Recently, Schilbach et al. (2013) have challenged the "observer paradigm," promoting a second-person approach (2nd p.a.) to social cognition and emphasizing the importance of dynamic, real-time interactions with others, e.g., eye-tracking. The affectivity springing by the mutual gaze responses between two interactors is primary to the conceptual comprehension of the other's mental state because it generates "common attentional patterns" toward the other (CFR. Elgin, 1999, pp. 146-69). The "attentional pattern" constitutes the direct entanglement between the interactors, which permits them to share common representations of the world, and, therefore, it is primary and constitutive of higher levels of mentalistic understanding of others' behavior. The 2nd p.a. then conceives the others like actual "You-person" directly influencing our social experience (Schilbach et al., 2013, p. 395).

This paper defends two ideas. First, the 2nd p.a. is necessary for a scientific and realistic account of social understanding. Accounting for the direct entanglement of social interactors is essential for overcoming the skeptical "spectatorial gap" (*Ibid.*:397-8). Second, the 2nd p.a. is not mutually exclusive of more mentalistic paradigms of social interactions, but, on the contrary, it is constitutive of them. So, we will propose an "integrative account" of social cognition. In the first and second sections, we will briefly expose the classical theories of social understanding ("theory theory" and "simulation") and their supposed neural correlates. Furthermore, we will uphold their inconsistency with a realistic account of the social experience. Finally, we will succinctly display the 2nd p.a. and its influence on social neuroscience, and we will advocate the phenomenological privilege of the second p.a. over the other accounts.

The “Theory theory” (or mind reading) and the third person approach (3 p.a.)

How is our understanding of others' minds realized? According to the report of the "theory theory" (T.T.), our knowledge of others requires the attribution of mental states[2] Through inferences from our perception of the observed behavior (Meltzoff, Gopnik, 2013). T.T. relies on "folk psychology" (F.P.), a term introduced by Wilfrid Sellars (1956) to refer to the system of psychological concepts (e.g., beliefs, feelings) used in everyday practice for ascribing mental states

to people. Besides, Sellars defines F.P. as a theory that postulates some unobservable entities- i.e., mental states- which are the occupants of specific causal roles for explaining behavior: e.g., "He kissed his daughter because he was cheerful (mental state: happiness)." Analogously to the scientific method, the application of F.P. is an inferential process from an observation to a hypothesis and a conceptual revision in the case of a mistaken conjecture.

Manifestly, F.P. theory is continually at play in social interactions. However, it is still being determined what kinds of cognitive mechanisms would allow persons to infer about complex mental states so swiftly and successfully. Hence, the term "theory theory" refers to a metatheory of F.P., which aims to explain the connection of mental states with related perceptual inputs and scientific evidence. If F.P. is merely a theory about the contents of the mind and their causal relations, T.T. should also try to demonstrate how we mentalize. Following the internalist TT[3] F.P.'s application consists of representations of others' minds, which occur in the brain of the individual agent (Slors et al., 2015, p. 256). Indeed, Frith and Frith (2006), in their seminal paper, try to propose the neural basis of "mentalizing," that is, our ability to read the F.P. mental states of other agents. According to the authors, "mentalizing" about the others' behavior is a complex activity that requires comprehension of their emotional state, the intentions lying behind their actions, and their stable attitudes and predilections (*Idem.*:531). Grasping these properties of others' mental conditions requires taking their perspective and engaging in many neural processes. So, Frith and Frith enlist every neural correlate involved in "mentalizing," specifying their peculiar function and quoting much experimental evidence. A critical bulletin about the authors' results is unnecessary because this chapter's chief point is the conceptual bias behind their approach. Here is a list of the proposed neural correlates:

> The region of the brain at the posterior end of the superior temporal sulcus (pSTS) and the adjacent temporoparietal junction (TPJ): This neural area is involved in many functional tasks, such as the recognizing of others' face and the observation of the other's eye movement. The direction of the other's gaze constitutes a clue for representing her visual perspective and inferring about the cause of its emotion, e.g., He is scared because a tiger is pointing at him, or about his intention, e.g., He

is looking at the cigarette on the table (because he wants to smoke) (Ibid.:532).

The temporal poles (T.P.): Anatomo-clinical studies have suggested that they are related to autobiographical memory (Dupont, 2002), but they also have a role in social and emotional processes, including face recognition and theory of mind (Olson et al., 2007). Frith and Frith propose that the T.P., in virtue of their connectivity to dorsal (auditory), medial (olfactory), and ventral (visual) streams, binds highly processed perceptual inputs to visceral emotional responses. Indeed, the authors suggest that this neural area is involved in converging sensorial information for recognizing a recurrent environment or situation. The T.P.'s activity is necessary to answer questions like how Mark usually feels when crossing an unsafe street. So, it permits "contextual" social comprehension, that is, understanding how a person is likely to be and feel in a recurrent context or situation, e.g., Mark is scared whenever he crosses a road (Cf. *Ibid.*:532).

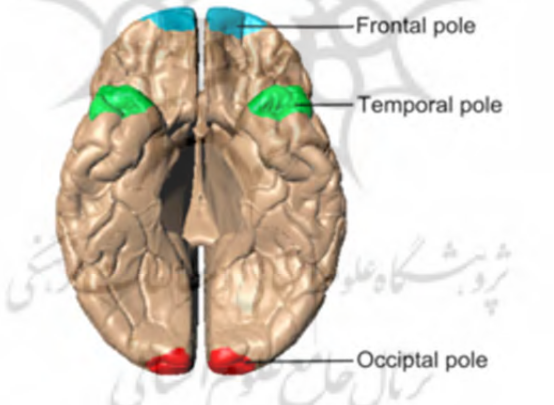


Fig. 1: The poles of the cerebral hemispheres

The medial prefrontal cortex and the adjacent paracingulate cortex: Following Stuss et al. (2001), lesions to the frontal lobes comport impairments in inferring the others' visual perspective, which implies difficulties in representing another's perceptions according to one's own past (*Ivi.*:282-283). Thus, there is a failure to recognize somatic markers, so emotional

experience helps guide response options. Consequently, the frontal lobes' activity entails the F.P. theory of mind. As claimed by Frith and Frith (2006), in general, the prefrontal cortex is concerned with planning for the future and representing anticipated states of the world (*Ivi.*:532). Thinking about the possible reactions of a person dissimilar from us, e.g., with diverse political ideas, involves dorsal regions of mPFC; meanwhile, for people similar to us, it requires activity in ventral mPFC (Amodio & Frith, 2006).

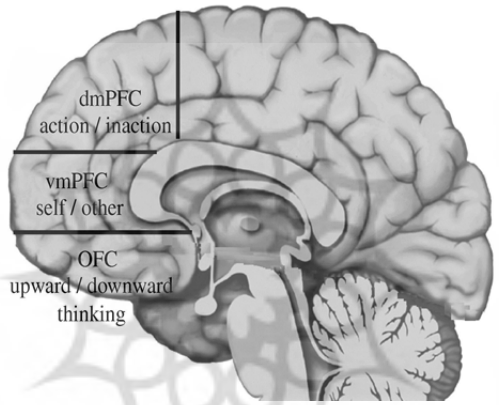


Fig. 2 The prefrontal cortex and other "mind-reading" functions.

These neural areas and mutual connections compose the "mentalizing system" (MENT). By convergence of perceptual and mnemonic information, the MENT gives the evidence to infer about others' mental states. Moreover, the MENT system has the advantage of associating the identification of abstract thoughts with a physiological make-up, connecting their cognition to an actual corporeal process. This proposal remains inevitably fastened to the "observer paradigm" because these capacities for social understanding imply a complex inferential process that starts with evidence about the other's behavior and the situational clues and concludes by inferring the other's mental states. Never is there a direct comprehension that does not involve a double-step procedure from observing to insinuating the other's feelings. Metaphorically, the other is a 3d-person, "He-She-person." Similarly to storytelling, making sense of

the other's mental state involves a narration from a detached viewpoint like an omniscient narrator, e.g., He was scared because he saw a tiger attacking him (Cf. Hutto, 2008). Although eye contact is present in the account of Frith and Frith, there is only one evidence for conjecturing over the intention and emotional state of the other, who is never directly engaged in a mutual gaze.

Hence, some problematic points arise. First, in the third-person perspective of "mindreading," the observer is an ideal subject. Indeed, the mental states of the person engaged in social understanding of the other are not in consideration. Consequently, her mind seems to be a "tabula rasa," which merely estimates the other's mental state without emotional engagement. The other's mental states should be self-experienced to some degree; otherwise, their common representation and comprehension are hardly conceivable (Cf. Northoff and Heinzel, 2006).

Relevantly for our concern, some issues for the realism of social cognition follow. The proposal of Frith and Frith (2006) remains fastened to a representational account of mind-reading. Namely, in the "perspective-taking" task, the other's glance is detected by the observer, but it does not tell anything by itself. Through her mind-reading capabilities, the spectator fulfills the other's silent representation. In truth, the expression "representing the mental state" repeats six times in the brief paper of Frith and Frith.

Consequently, if no direct self-experience and neither unmediated other's influence on our perception is at play, the justification and guarantee for the final mental inference is absent, and only an oral confirmation could assure its correctness. Again, the skeptical doubt is behind the corner. With a phenomenological account of our affectivity and the other's influence on that, the representations of the other's mental states arise from mere detached observation, and their phenomenology needs to be clarified. In conclusion, our private affectivity and direct and physical contact with the other are the only assurance of the reality of our "mind-reading," or we would remain with a very sophisticated and mentalistic theory that loses touch with existence. For this reason, in the next chapter, an analysis of the first-person approach will show some more primitive capabilities for social understanding, which consider self-affectivity and are necessary but insufficient for a realistic account of mindreading.

The "Simulation theory" (S.T.) and the 1st-person approach (1st p.a.)

The "Simulation Theory" (S.T.) claims that social cognition involves "putting ourselves in the shoes of others" by simulating the mental states we would have in their situation. Hume (2000) stated that studying one's conscious states leads to discovering general principles applicable to others. Thus, the introspection has a dominant role in the understanding of others. The insights from Hume have influenced Goldman's version of S.T. (2006). The core assumption of S.T. is the homogeneity of people's mental state.[4] Hence, the observer's cognitive mechanisms are similar to the person whose behavior he is trying to understand. If compared to T.T., S.T. is a deflationary account because there is no need for multiple steps inferential processes. However, only for one analogical inference, that is to say, the observer tries to imagine the other in her situation by simulating to occupy her place (Slors et al., 2015, p. 258). The S.T. relies on the first-person perspective, which bases the social understanding upon one's self-perception. Although S.T. can explain more basic forms of social comprehension, it could only reduce some of the explanatory power of T.T., such as understanding the other's political opinions.

For clarifying S.T., the history of discovering its neural correlates- i.e., mirror neurons- is helpful. In the 80s, mirror neurons were discovered in the brains of macaque monkeys by Rizzolatti and his colleagues from the University of Parma (1992). In the first moment, they detected that an area of the premotor cortex called F5 fired whenever the monkeys reached for a peanut. Successively, they surprisingly noticed that when the researchers grasped an object, such as a peanut, to hand it to the monkey, the same monkey's motor neurons would also fire when the monkey itself grasped the peanut. Further, they detected that individual neurons would only respond to specific actions, such as one Neuron for simulating the grasping of a peanut and a different one for putting a peanut in the mouth. Accordingly to the findings, many have speculated that the comprehension of the other's intentionality- i.e., what is he looking for?- depends upon the inner mimicry and resonating with the other's action.[5]

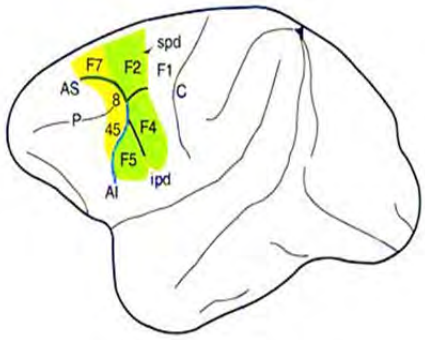


Fig. 3: Mirror neuron system (MNS) in macaque monkeys.

Recently, many researchers have committed to determining the correlates of MNS in the human brain. Most studies on the human mirror-neuron system have used neuroimaging, generally functional magnetic resonance imaging (fMRI). However, when we record signals from the human brain containing millions of neurons, the task is complicated, and only vast neural areas, instead of single neurons, are identifiable. The current study of Watanabe et al. (2017) is a prototype of the human brain mapping of MNS. Their experiment was based on an fMRI scan of participants performing imitative tasks of the other's finger movement. The authors identify the following neural areas:

> The right ventral premotor area (PMv) and the inferior parietal lobule (IPL) are associated with synthesizing visual and kinesthetic information from observed limb movements. PMv further contributes to visuomotor transformations required for correcting the hand posture configuration (*Ivi.* 6225).
>
> The Inferior frontal gyrus (IFG) is involved in recognizing the intentionality or the goal of the observed action (*Ivi.*:6226).
>
> The insula is engaged in the sense of self-awareness and body ownership (*Ibid.*). So, it permits one to distinguish the self and the simulated other when performing the same motor task, e.g., imitating the rapid finger movements.

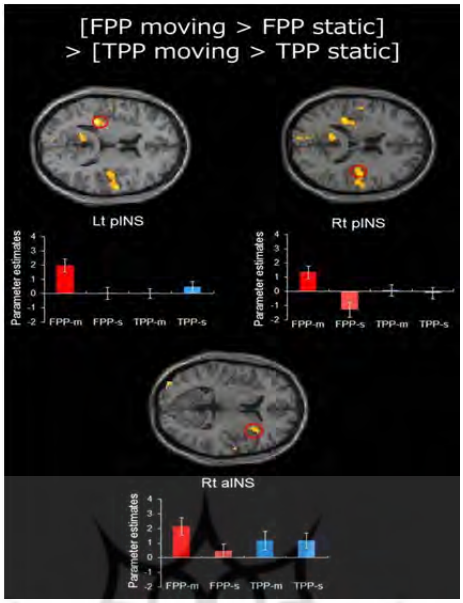


Fig 4: The activity of the insula in the Watanabe et al. experiment.

As implied above, the research in the human brain MNS has moved beyond matching the motor system for explaining more sophisticated mental tasks, such as understanding the other intentionality and the self-other distinction. Vittorio Gallese and Alvin Goldman (1998) suggested that the discovery of MNS provided additional support for the simulation theory. They also proposed that S.T. is not incompatible with mindreading but is constitutive of it. Indeed, the mirror neurons' activity largely depends on unconscious motor processes rather than consciously imagining the beliefs and desires of the other.[6] From the evolutionary standpoint, the discovery in macaques' brain also suggests that this basilar kind of social understanding could have developed earlier than more cogitative ways of comprehending. So, the inferential processes in the third-person approach should be constituted by an "analogical inference" in the first-person approach. Further studies have suggested that the role of the insula in simulating the other's disgust (Wicker et al., 2003) or the role of the somatosensory cortex in "tactile empathy" how we experience the sight of others being touched, simulating the same sensorial stimuli (Keysers et al., 2004).

A complete account of social cognition requires addressing the first-person approach and S.T. for linking the inferential processes (3d p.a.) to the egocentric sensorimotor activity (1st p.a.). The connection explains the development from involuntary motor simulation to highly evolved "mind-reading" capabilities in children. Moreover, first, p.a. could account for the self-experience in the phenomenology of social cognition. Consequently, it resolves at least a problem for the realism of social understanding. The observer and simulator is not a "tabula rasa"; instead, he "lives the other's feeling." At least one element- i.e., one's own affectivity- constitutes the justification of the representation of the other's mental states, which are not only inferred by the observed evidence but the automatic self-experience forms them.

Another clue suggesting the necessity of integrating first p.a. and third p.a. is that MNS and MENS partly overlap because the mirror neurons are found in a vast network of neural areas according to their specific function. For instance, if we look at the MNS map of the Indian Association of Psychiatry (Rajmohan & Mohandas, 2007), the mPFC and the superior temporal sulcus are present like in the MENS (Frith & Frith, 2006). So, the same neural areas are activated for simulation and Folk-psychology tasks. Hypothetically, the two processes occur at a different time scale - i.e., simulation activity would be faster on a millisecond measure-or different neural plasticity is at play for diverse functions, or the same brain networks have different functional connectivity, both internal and external. Further experimental data suggests adequate ways for integrating the MNS (S.T.) and the MENT (T.T.) for a complete account of social cognition.

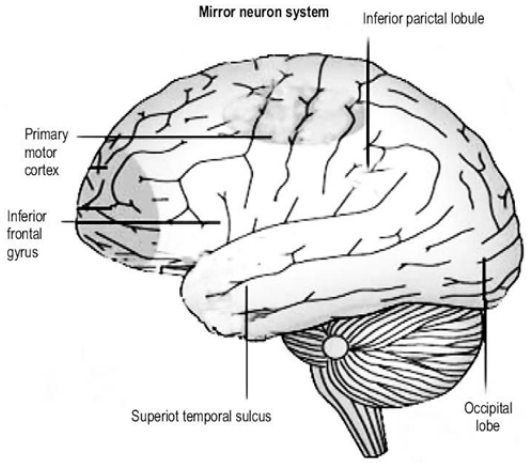


Fig. 5: Mirror neuron system in the human brain (Indian Association of Psychiatry)

However, the S.T. and first p.a. are insufficient to overcome the "realistic gap" mentioned above. Also, the 1st p.a. is fastened to the "observer paradigm" (De Bruin et al., 2012) because the experimental setup usually involves detecting the brain's activity of a participant observing or simulating someone else's movement without addressing intersubjective "joint attention." Furthermore, the simulative social understanding depends, in a particular way, on a representational theory of mind. Even if the other is not depicted by imagining, the internal activity of the insula and primary motor cortex is, to a certain degree, representative of the other, who never is directly perceived. Consequently, the observer's bodily feeling depicts the interactor by mediating analogical inference. A realist phenomenology of social cognition should address the active role of the other, or it would seem unintelligible how he affects and directs our comprehension in the right direction. Moreover, the simulative activity likely deceives us if the other is not directly involved. Hence, the skeptical doubt pops up again.

The next chapter will analyze the second-person approach to social cognition and the sensorimotor theory, which are strictly intertwined. This approach explains the more basilar characteristics of social experience in which the interactors mutually affect each other before applying for an intellectual or simulative detour to understand the other.

3- The second person approach (2nd p.a.) and the sensorimotor theory (SMT).

A non-representational theory of mind is indispensable for assessing a direct and realistic account of social cognition. The "sensorimotor" (S.M.) theory of perception offers the necessary conceptual background. Accordingly, cognition is an action-oriented exploratory activity more than a passive representation; that is to say, sensorimotor contingencies determine our vision (CFR. O'Regan and Noe, 2001). Following Husserl (1973), the perception of worldly objects relies on “horizons" of perspectives. For instance, the mere representation of a door, which is closed, does not imply that there is a posterior surface on the other side. The perceiver has the tacit knowledge that he could pass to the back surface, opening the door; therefore, only by understanding his motor possibilities could he mentally grasp a complete picture of the object. Moreover, the implicit knowledge of sensorimotor capabilities influences our perception of the objects. However, it does not require further representations- i.e., no one needs to imagine the door's back to determine the possible action on the door's handle or vice-versa.

Moving to our topic, the other is an active interactor, which offers "social affordances," which is a subcategory of affordances (Rietveld et al.). From phenomenology, we learn that we usually engage skillfully with our environments under the unreflective actions "solicited" by the situation. "Social affordances" are sensorimotor loops for interaction driven by others, allowing for interpersonal behavior coordination (Schilbach et al., 2013, p. 401). For specimen, the ability of "taking perspective" (see above, Chp.1) of the other does not solely depend on observing the other's glance to understand what he is glancing or staring at. On the contrary, in real-time social interaction, the other is an initiator or a responder (Schilbach et al., 2013, p. 5); that is to say, he or she invites us to look in a determinate direction, or vice-versa, he or she is responsive to the shift of our eye-movement. Thus, the interactor immediately intervenes in our perception in a way that an intellectual or simulative detour cannot address. In this intersubjective account, the other is thought to be a second-person. He or she is concretely existent, and her influence or responsiveness to our presence provides a "perceptual common ground" (Cf. De Bruin et al. 2012), which represents the backbone of high-level cognition- i.e., "mindreading" mental states (T.T.) or S.T.

Following the second-person approach (2 p.a.), Schilbach et al. (2013) have created an experimental setup worth determining the neural correlates of essential mutual interaction. Respectively, the setup must be "minimalist and yet naturalistic" (*Ibid.:*404), which means that the focus is on elementary forms of interactions that do not require highly abstract thoughts-e.g., mutual gaze or correlated hand gestures. The emphasis on automatic tasks is needed to avoid confusion with first p.a. and third p.a. that require the ability to read the other's mental states. Thus, their experiment establishes a participant looking at a virtual character, which, in virtue of developed algorithms, is responsive to the direction of her gaze. Some objects-i.e., three grey squares, are placed at the sides of the screen (see Fig. 6), and the virtual interactor can also ignore the participant's glance and point his eyes at a different object. Last, the participant should be convinced to interact with an actual human "behind the screen" to be motivated in the performance and recreate a likely accurate situation. The obtained data through fMRI scanning and eye-tracking want to determine which neural areas are functionally active in specific roles of the 2nd-person interaction, e.g., the participant invites the virtual character to watch at a particular object (initiator) or vice-versa (responder), or they are both neglecting each other (non-joint attention).

As per Schilbach et al., a significant result is that the neural correlates for non-joint attention and joint attention are different. This factor implies a neural discrepancy between mere observation and intersubjective experience (*Ibid.*:407). The lateralized frontoparietal network is involved in non-joint attention; that is to say, the interactors look at different objects. Generally, the medial prefrontal cortex (mPFC) and the posterior cingulate cortex (PCC) are involved in "joint attention." Specifically, the ventral striatum is active for self-initiated joint-attention, which means that the participant invites the virtual other's gaze toward an object, and the anterior mPFC is active when following someone else's glance (*Ibid.*:403).

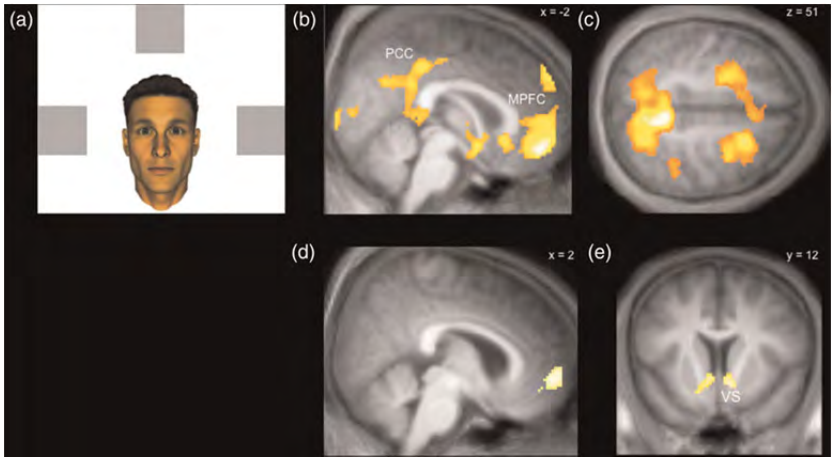


Fig.6.: a- the virtual character, b- neural correlates of joint-attention, c- non-joint attention, d- other initiated joint-attention, e- self-initiated joint-attention. (Schilbach et al. 2013).

Surprisingly, the authors themselves state that these brain regions are overlapping with the MENT system (*Ibid.*:404). According to Amodio and Frith (2006), the PFC's activity facilitates "contextual knowledge" about others, like knowing that such a disgusted expression occurs whenever my friend sees the face of a political opponent on the T.V. screen. This sophisticated form of social comprehension requires the ability to think about the other's conceptual mental states, like political beliefs. Nevertheless, demanding that a rudimentary level of cognition foreruns such a compound form of social understanding is a logical assumption. Following the second p.a., the activity of responding to social affordances- i.e., eye tracking- is not a radical alternative to "mindreading" (T.T.) or simulating (S.T.), but rather it is constitutive of them. We propose that the "2nd-person cognition" represents and explains our direct entanglement with others, which cannot be exhaustively described by a conceptual (T.T.) or physical deduction (S.T.). Therefore, it is the backbone of S.T. and T.T.

Let us consider the case of our friend's hateful expression when looking at a political rival on the T.V. screen. At the first moment, the anterior portion of mPFC actively follows the fellow's "social affordance"- i.e., his invite to look in the direction of the television. Hence, he exhorts us to understand his state of mind in a purely

physical way (SMT) without needing meta-representations or inferential processes. In the second step, the MNS system can account for some elementary forms of empathy mirroring the other's facial mimicry. During facial emotion processing, mirror neurons in the primary motor cortex provide an internal simulation of the observed facial expression that elicits a similar emotion in the observer, thus aiding the identification of that sentiment (Enticott et al., 2008). Given our attunement to others, to account for the inner affectivity provides the physiological make-up underlying our perception of the other's feeling-i.e., "He is expressing disgust or hate in the same way as I do." Finally, the activity of the MENT system, in virtue of the convergence of perceptual and mnemonic information, can address the "contextual" knowledge. Namely, the association of the occurring situation- e.g., what is going on television- with the other's bodily "affordances"- e.g., disgusted facial expression. To resume:

> Lower-level: second p.a. The other invites me to a "common attentional pattern," which is constitutive of conceptual knowledge and guarantees the direct entanglement with the other's existence →Medium-level: first p.a. Through the inner mimicry; I understand the other's corporal expression. At this level, the self-affectivity becomes part of the realistic phenomenology of intersubjective cognition → Higher-level: 3d p.a. The sensorimotor information and the surrounding factors integrate for inferring the other's complex mental states, which depend upon socio-cultural knowledge or recurrent situations, e.g., political beliefs or being afraid whenever he crosses the road

4- Concluding remarks and prospects.

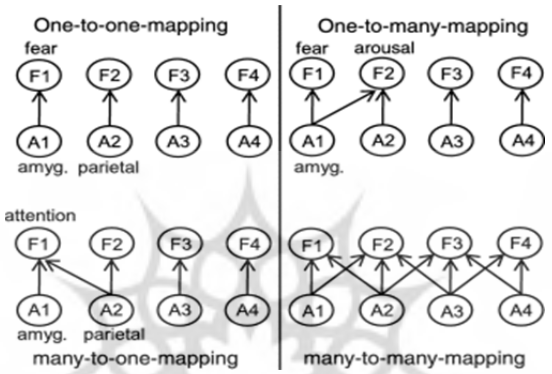
First, none of the approaches and related social cognition theories (S.M., S.T., and T.T.) are mutually exclusive. An "integrative account" is the best way of addressing a complete and realistic phenomenology of intersubjective understanding. The challenge at the philosophical level is consistently joining the theories of mind. The S.M. theory and the two p.a. should be considered primary to diverse approaches because of their directedness. This allows us to address the other's unmediated influence on our experience as in real-time interactions. Here, an objection arises. The 2nd p.a. is not explicative of the other's mental states. The research of Schilbach et al. only

permits the existence of a mutual "attentional pattern." However, it cannot explain intentionality (Moore & Paulus, 2013) or cultural factors that require conceptual and contextual knowledge and mental representation.

Nonetheless, I do not conceive it as a real issue because the three approaches and theories of mind are at different levels and irreducible to each other. So, the 2nd p.a. and S.M. theory can account for the common "attentional pattern," which is the basement of every social encounter. However, although it necessarily constitutes the higher mind-reading capabilities, they are not reducible to it. Mind-reading skills (e.g., empathy and contextual knowledge) duly belong to different "realms of explanation," which are S.T. (1st p.a.) and T.T. (3d p.a.). For instance, new research could focus on how the different roles in the "joint-attention task"- i.e., initiator or responder- could affect the reading of the other's mental state in the same complex real-like situation. Lastly, an over-comprehensive account would need to be more accurate in explaining the stratified and multi-step process of social understanding. Every different theory has the privilege of its specific level of explanation.

Consequently, the real challenge for social neuroscience is providing evidence to justify the conceptual connection between the different theories. For instance, the fact that Schilbach et al. determined correlates for second p.a. overlap with the MENT system, according to the brain mapping of other authors (Frith & Frith, 2006), provides additional support to my proposal. The ideal proof would be reproducing a complete social interaction by experiment and detecting at which time-scale the different neural areas are functionally active. Such an accurate record is currently non-testable because it is arduous to discern different processes on a millisecond scale. Besides, if the regions overlap, distinguishing which kind of function (e.g., eye-tracking, second p.a. or intentionality-grasping, first p.a. or 3d p.a.) they are currently performing is almost impossible. So, the same brain regions are active for a multitude of functions (one-to-many mapping). The task is discovering the physical connectivity among brain areas, which is representative of the conceptual connectivity among social cognition theories. This is a two-pronged process, which requires a deconstructive phase and a reconstructive one. The first step involves elucidating how the same neural area could perform different teleology and complexity tasks at distinct times. Schilbach et al.

(2006:407) propose that neural plasticity differences could exist in the same neural networks based on one's own involvement- i.e., the mPFC in direct (2nd p.a.) or indirect interaction (3d person p.a.). The further step involves determining the functional connectivity for the high-level social cognition, which involves a plurality of regions for the same complex "mind-reading" task (many-to-many mappings)- e.g., mPFC (2nd p.a.), inferior frontal gyrus (MNS, first p.a., S.T.) and MENT system (3d p.a., T.T.).



Stands for neural area; F stands for function.

In conclusion, we wish to avoid misunderstandings on the link between 2nd p.a. and sensorimotor theory with the realism of social cognition. The sensorimotor theory accounts for the direct entanglement of the interactors, justifying their actual role in social understanding. From this, it does not follow that our perception of the other is not partial, and it could not be mistaken. In contrast, our cognition is always and necessarily perspective and incomplete. The relevant implication of the 2nd p.a. is that the other is a concrete element of the social cognition loop. So, the phenomenon of mutual comprehension and communication becomes more intelligible than through a mysterious "intellectual detour." The history of the philosophical theories of social understanding goes from extremely intellectual accounts, which cannot explain the primary interactive phases, to deflationary and physical proposals, which cannot make sense of the more abstract "mental states." Now, it is time to take all the different insights and join them in a complete and realistic integrative account.

Notes

1. For an overview of the topic, see Siegel 2016.
2. For a complete overview of the nature of mental states, See Putnam (1967).
3. A different version of T.T. is the externalist; see Slors et al. (2015:256).
4. Gallese: "It seems we're wired to see other people as similar to us, rather than different (...) At the root, as humans, we identify the person we're facing as someone like ourselves".
5. For objections to this conclusion, see Slors et al. (2015:266-267).
6. Gallese: "This neural mechanism is involuntary and automatic," he says. With it, we don't have to *think* about what other people are doing or feeling; we know.

Bibliography

Amodio, D.M., & Frith, C.D. (2006). "Meeting of mind: the medial prefrontal cortex and social cognition," in *Nat. Rev. Neurosci.* 7, 268–277.

Asch, S. (1952). *Social psychology*, Prentice-Hall.

Chisholm JD, Chapman CS, Amm M, Bischof WF, Smilek D, Kingstone A (2014). A Cognitive Ethology Study of First- and Third-Person Perspectives, in *PLoS ONE* 9(3).

de Bruin, L., van Elk, M., Newen, A. (2012). "Reconceptualizing second-person interaction," in *Front Hum Neurosci. 2012; 6*: 151, Published online 2012 Jun 6.

Dupont, S., "Investigating temporal pole function by functional imaging," in *Epileptic Disord*. 2002 Sep;4 Suppl 1:S17-22.

Elgin, C.Z., *Considered judgment* (1996), Princeton, Princeton University Press.

Enticott, P.J., Johnston, P.J., Herring, S.E., Hoy, K.E., Fitzgerald, P.B. (2008). "Mirror neuron activation is associated with facial emotion processing," in *Neuropsychologia 46 (2008)*, 2851–2854,

Frith, C.D., Frith, U. (2006). "Neural basis of mentalizing," in *Neuron 50*, 531–534, May 18, 2006, a2006.

Gallese, V., Goldman, A. (1998). "Mirror neuron and the simulation theory of mindreading" in *Trends in Cognitive Sciences*, 2, 1998, pp.493–501.

Goldman, A. (2006). *Simulating minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: University Press.

Hume, D. (2000). A Treatise of Human Nature, Clarendon Press, Oxford.

Husserl, E. (1939). *Experience and Judgement*, trans. J. S. Churchill and K. Ameriks, London: Routledge 1973.

Hutto, D. (2008). *Folk psychology narratives: the socio-cultural basis of understanding reasons*, Cambridge MA: MIT Press.

Hutto, D., Myin, E. (2013). *Radicalizing Enactivism: Basic minds without content*, Cambridge MA, MIT Press.

Jenkinson, J. (2017). "Enactive subjectivity as flesh," in *Phenomenology and the Cognitive Science 2017 16*: 931–951, published online, Berlin: Springer.

Keysers, C., Wicker, B., Gazzola, V., Anton, J.L., Fogassi, L., Gallese, V. (2004). " A Touching Sight: SII/PV Activation during the Observation and Experience of Touch," in *Neuron Volume 42, Issue 2*, Apr 22, 2004, Pages 335-346

Meltzoff, A.N., Gopnik, A. (2013). "Learning about the mind from evidence: Children's development of intuitive theories of perception and personality," in *Understanding other minds: perspectives from developmental social neuroscience* (Ed. Baron-Cohen, Lombardo and Tager Flusberg), Published to Oxford Scholarship Online: September 2013.

Moore, C., Paulus, M. (2013). " A second-person approach cannot explain intentionality in social understanding," Comment in *Authors' response: A second-person neuroscience in interaction*, in *BEHAVIORAL AND BRAIN SCIENCES (2013) 36*, 393–462.

Northoff, G., Heinzel, A. (2006). "First-Person Neuroscience: A new methodological approach for linking mental and neuronal states," in *Philos Ethics Humanit Med. 2006; 1*: 3.

Olson, I.R., Plotzker, A., Ezzyat, Y. (2007). "The Enigmatic temporal pole: a review of findings on social and emotional processing," in *Brain.* Jul;130(Pt 7):1718-31. Epub 2007 Mar 28.

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G. (1992). "Understanding motor events: a neurophysiological study," in *Exp Brain Res. 1992;91(1)*:176-80.

Pessoa, L. (2014). "Understanding brain network and brain organization," in *Physics of Life Reviews* 11 (2014): 400–435.

Putnam, H. (1967). The nature of mental states. In W.H. Capitan & D.D. Merrill (eds.), *Art, Mind, and Religion*. Pittsburgh University Press: pp. 1–223.

Rajmohan, V., and Mohandas, E. (2007), “Mirror neuron system”, in *Indian J Psychiatry. 2007 Jan-Mar; 49(1)*: 66–69.

Rietveld, E., De Haan, S. & Denys, D. (forthcoming). Social affordances in context: What is it that we are bodily responsive to?

Invited commentary article on Schilbach et al., BBS, Behavioral and Brain Sciences.

Schilbach, L. et al., (2013), "Toward a second person neuroscience," in *BEHAVIORAL AND BRAIN SCIENCES (2013) 36:* 393–462.

Sellars, W. (1956). "Empiricism and the Philosophy of Mind," in *Minnesota Studies in Philosophy of Science*, 1: 253–329.

Siegel, S. (2016). "The contents of perception," in Stanford Encyclopedia of Philosophy, published online Tue Oct 4, 2016: https://plato.stanford.edu/entries/perception-contents/.

Slors, M., de Bruin, L., Strijbos, D., (2015), *Philosophy of mind, brain, and behavior*, Amsterdam: Boom.

Stuss, D.T., Gallup G.G. Jr., A. M.P. (2001). "The frontal lobes are necessary for "theory of mind," in *Brain. 2001 Feb;124(Pt 2):*279-86.

Van Fraassen, B. (1980). *The scientific image*. New York: Oxford University Press.

Watanabe, R., Higuchi, T., Kikuchi, Y., Masato, T. (2017). "Visuomotor Effects of Body Part Movements Presented in the First-Person Perspective on Imitative Behavior," in *Human Brain Mapping 38*:6218–6229.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., Rizzolatti, G. (2003). "Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust," in *Neuron. 2003 Oct 30;40(3)*:655-64.

Wilson, R.A. (2015). "Embodied cognition," in Stanford Encyclopedia of Philosophy, published online Tue Dec 8, 2015: https://plato.stanford.edu/entries/embodied-cognition/.