Husserlian Empathy and Simulationism

Zachariah A. Neemeh (zaneemeh@memphis.edu)
Department of Philosophy, University of Memphis
Memphis, TN 38122 USA

keywords: empathy; enactivism; Husserl; simulation theory; social cognition

Although “social cognition” is, in philosophy, typically circumscribed to discussions in the analytic philosophy of mind and the philosophy of cognitive science, the phenomenological tradition has engaged with similar problems since the beginning of the 20th century. There has been a growing realization that phenomenology intersects with these discourses in significant and surprising ways (e.g., Gallagher & Schmicking, 2010; Gallagher & Zahavi, 2012; Petitot et al., 1999). Indeed, as early as 1958, Gilbert Ryle and Merleau-Ponty met among other analytic and continental philosophers at Royaumont Abbey in France, where they noted important intersections in their work (Gallagher, 2017). Does phenomenology have anything to say about social cognition? In the domain of cognitive science, theories of social cognition are dominated by two paradigms: theory theory (TT) and simulation theory (ST). Theory theory (TT) hypothesizes that social cognition operates by the subject actively making inferences about other people in accordance with a folk psychology. Simulation theory (ST) comes in inferential and subpersonal forms. Inferential ST hypothesizes that social cognition operates on the model of the subject’s own mind. Subpersonal or neural ST hypothesizes that the mirror neuron system provides a rapid, unconscious mechanism for social cognition by activating regardless of the agent performing the action (whether it be oneself or another). Recently, an interaction theory (IT) theory of social cognition has been proposed as a challenge to these competitors. IT hypothesizes that the perception of other minds, intentions, and beliefs is direct, noninferential, and interactive. IT objects to TT’s and ST’s representationalism and detachment from embodied interaction. IT is influenced by developmental, phenomenological, enactive, and dynamical systems approaches to social cognition (Froese & Gallagher, 2012, p. 436). The phenomenological influence comes especially from Merleau-Ponty, whose concept of intercorporeity centers the role of embodied interaction in social cognition.

I will analyze Husserl’s theory of empathy and intersubjectivity as a theory of social cognition. While Gallagher (2013) has shown that Merleau-Ponty is significantly in step with IT (IT was partially influenced by Merleau-Ponty, after all), Husserl’s thought significantly diverges from Merleau-Ponty’s—especially in the domain of social cognition.¹ First, I will give a general outline of the TT, ST, and IT theories of social cognition. Then, I will attempt to explain Husserl’s empathic theory of social cognition. I will also discuss Heidegger’s (1926/2010) critique of Husserl’s theory in Being and Time as it relates to IT. I will argue that Husserl’s theory is closest to ST, while simultaneously highlighting his differences from ST. Ultimately, it is my hope that this work will contribute to the greater dialogue between phenomenology, the philosophy of mind, and the philosophy of cognitive science.

¹ While Froese and Gallagher (2012) refer to Husserl in the context of defending IT, I believe Husserl is closer to a simulationist position.
Theories of Social Cognition: Theory Theory, Simulation Theory, and Interaction Theory

TT posits that social cognition functions by reference to a folk psychological theory of mind. This theory may be innate, or it may be acquired through learning. Such a theory refers to beliefs, desires, hopes, and other folk psychological notions to explain other people’s mentalities and behavior. Inferential forms of ST argue that there is no need for such a theory, as our own mind already serves as a model of other minds. “Because they have one of those mechanisms themselves [viz., a mind], they can simply run their mechanism on the pretend input appropriate to the target’s initial input” (Goldman, 2006, p. 20). I understand your mind and behavior by simulating or pretending I am performing your actions. Subpersonal or neural forms of ST instead posit that such simulation is performed by the mirror neuron system. As such, it operates below the surface of awareness. Hybrid TT-ST theories have also been proposed, combining the folk psychological mentalizing with simulation.

IT is critical of TT and ST for logical, phenomenological, and developmental reasons. IT rejects subpersonal or neural forms of ST as incoherent. Gallagher and Hutto argue that in subpersonal processes, “(1) there is no first- or third-person…; (2) nothing (or no one) is using a model; and (3) neuronal processes cannot pretend” (2008, p. 19). Phenomenologically, TT and inferential ST are unsound, because our everyday experience of other people is not marked by continual reflective thoughts about their behavior or pretending as if we were performing their actions. Instead, much of our understanding of others is immediate, prereflective, and noninferential (Froese & Gallagher, 2012). I do not need to explicitly think about whether your tone of voice indicates that you are mad. Your tone of voice itself, along with your body posture—that is, my interaction with you—conveys as much to me. Finally, TT and ST are not developmentally sound. TT and inferential ST rely on the passing of the false belief test, which generally occurs around age four. Intersubjective interaction occurs long before age four, however (Gallagher & Hutto, 2008). IT rejects the methodological individualism of TT and ST and hypothesizes that social cognition is direct in the world (Froese & Gallagher, 2012). IT instead proposes an intersubjective interaction that is primarily second-person. Rather than being predicated on a theory of mind, “social understanding depends on, and is realized by, social interaction” (Froese & Gallagher, 2012). Instead of social cognition being something localized within the brain of an individual, it is instead (mostly) an emergent understanding that develops between two dynamically interacting embodied agents.

Notably, these theories are primarily concerned with the human social perception of other humans. For Husserl, the social cognition of other humans is a special case of the social cognition of lived bodies (Leiber) in general. Two avenues of research into the human social cognition of nonhumans are in the domains of human-robot interaction and in nonhuman animal stereotypes. Research in human-robot interaction shows that humans engage with robot agents differently depending on their physical and behavioral resemblance to humans. Although humans automatically track robot gaze regardless of degree of similarity, they only continue to track robot gaze when the robot is perceived to be in some sense minded. Attitudes towards robotic agents are likewise dependent upon their physical and behavioral similarity to humans (Abubshait & Wiese, 2017). Humans also tend to respond differently to robotic agents than to human agents. For example, human response times to robotic agents tend to be slower than to other humans. Many

---

2 This is possibly not the case in Cartesian Meditations. I will develop this claim in the following section.
subtle differences in bodily movement and bodily imitation are also manifest when humans interact with humanoid robot agents (Kim, Kwon, & Kim, 2018).

While IT draws on several fields, including developmental studies, enactivism, and dynamical systems accounts of cognition, my concern here is with phenomenology. IT’s phenomenological influences are primarily coming through Scheler and Merleau-Ponty. Scheler, in The Nature of Sympathy, notes that the perception of other people’s emotions is direct rather than inferential (Gallagher, 2008)—something also affirmed by Merleau-Ponty (Gallagher & Varga, 2014). Merleau-Ponty’s concept of intercorporeity, however, is a particularly important reference for IT. Intercorporeity refers to the mutual entanglement of my own body as I experience it with other agents in my environment. “[I]ntercorporeity involves an inter-enactive response to the other’s action; it is taking that action as an affordance for further coordinated action” (Froese & Gallagher, 2012, p. 448). Merleau-Ponty, however, conceives of intercorporeity as having more radical ontological consequences than they hint at. Intercorporeity is an aspect of the general chiastic structure of perception. This chiastic structure, already nascent in Phenomenology of Perception, is the reversibility of the seer and the visible, or the one who senses (le sentant) and that which is sensed (le sensible) (Dupond, 2001, p. 5). “There is no problem of the alter ego because it is not I who sees, not he who sees, because an anonymous visibility inhabits both of us, a vision in general, in virtue of that primordial property that belongs to the flesh [viz., chiastic structure]” (Merleau-Ponty, 1968, p. 142). Merleau-Ponty’s thought here can be usefully compared with what dynamical systems approaches to cognition refer to as the coupled brain-body-environment system, which Froese and Gallagher note “could include other agents” (2012, p. 451).

Froese and Gallagher also briefly refer to Husserl’s Zur Phänomenologie der Intersubjektivität, noting that he “prefigured the contemporary discussion of neural resonance, namely that our perception of the other person induces a sensory-motor process that reverberates kinetically and kinesthetically with their intentions” (Froese & Gallagher, 2012, p. 448). While Husserl’s Cartesian Meditations (1960) contains his most well-known and widely-discussed theory of intersubjectivity, he elsewhere discusses social cognition in broader terms that include nonhuman animals (Husserl, 1980; 1989; 2013). In the following section, I elucidate his theory of empathy as it is variously presented in Cartesian Meditations, Ideas II, Ideas III, and Addendum XXIII to the Crisis of European Sciences and Transcendental Phenomenology.

Empathy in Husserl

The Cartesian Meditations is Husserl’s most well-known development of his empathic theory of intersubjectivity and social cognition. In this text, “empathy” is narrowly referred to as a “transcendental theory of experiencing someone else” (Husserl, 1960, p. 92). Social cognition is narrower in scope here, whereas his other engagements with empathy extend to all lived bodies (Leiber). Empathy is the mechanism by which intersubjectivity develops—it is the bridge from subjectivity to the community of subjects. Husserl notes that some objects in the world (viz., human bodies) are perceived to be subjects or selves that are characterized by intentionality. The perception of other human bodies is different from the perception of rocks in important ways. These bodies are perceived as minded. The problem of other minds does not show up for Husserl, because he understands experience to already be replete with an experience of other minds.3 Our experience is not characterized by a strict, solipsistic interiority; the world is experienced as shared by multiple minds. This founds, for Husserl, the objectivity of the world. The world itself is given

3 Whether and to what degree the perception of other minds is accurate is another question (see Chudnoff, 2018).
as “experienceable by everyone,” “even if a universal plague had left only me” (p. 93). The empathic or intersubjective structure of consciousness is not, as Husserl highlights in the previous quote, a contingent feature.

Other minds are not given to my experience originally; only other people’s bodies (Körper) can be given to me originarily. In other words, others’ bodies alone are given as full, material presences in my perceptual field. As minds, however, they are “appresented” or co-presented with their bodies. Appresentation is a way that things are given without being present per se. Just as the hidden side of cube is appresented to me even though I can only “see” its front, other minds are given to me even though I can only “see” their bodies. But most appresentations can be verified by further experience. I can at least in principle verify that the cube actually has a far side. I can never, however, verify that the other person is actually minded. There is an epistemologically asymmetry between appresented objects and appresented subjects.

Husserl resolves this asymmetry by recourse to empathic transference or projection, an “aperceptive transfer from my animate organism” (1960, p. 110). Empathy is a kind of noninferential projection of my own embodied psyche (“animate organism” or Leib) into the other person. Empathic projection animates the Körper of the other into a lived body. Although Husserl emphasizes that others form a community or interconnection with me (which he calls “transcendental intersubjectivity”; 1960, p. 131), others are mostly considered from an epistemic point of view. Husserl is not concerned with bodily interaction so much as with the epistemological interaction of different minds constituting an objective world. Importantly, the pairing of my living body to the other person occurs under conditions of similarity. Furthermore, it must exhibit “harmonious behavior” (p. 114). If the behavior of the body before me is somehow discordant or does not follow typical human behaviors, I instead experience this body as a “pseudo-organism” (p. 114). These two elements, similarity of body and behavior, are precisely what Abubshait and Wiese (2017) highlight as what determines whether humans interact with robotic agents as agents or as mere machines.

In Ideas II, Husserl more expansively discusses empathy in terms of both human and nonhuman animal others. He uses the all-inclusive term “animalia” to designate the set including both, while leaving what particular species are included deliberately vague. Husserl notes that most material and extended bodies are fragmentable into smaller bodies that may differ in type, but not in kind. Half of a rock is two rocks, half of a tree is timber, and half of a car is a wreck. But, splitting a human in half does more than produce a corpse. It changes the type of thing it is from animate to inanimate. By undergoing spatial fragmentation, animalia lose their “psychic” component (p. 36). Animalia are given in our experience as beings with a psychic life. That is, we perceive them to have a first-person perspective and we explain their behavior in intentional terms (e.g., the cat wants to play). Animalia are given both as living bodies and as “interiorities” (that is, as having a point of view or consciousness; p. 171).

Husserl’s empathic perception is canonically applied to humans, but he understands it to be the way all living bodies, human and nonhuman, are perceived. His theory of social perception is also a theory of the human perception of nonhuman animals. This empathic perception is direct and noninferential (Painter, 2007); nonhuman consciousness is not “analogically understandable, thus it does not receive [its anima] from egoity” (Husserl, 2013, p. 6). Although the animal’s consciousness is not originarily perceived in the same way that its body is (i.e., it is not present, and I have no access to it), it is nonetheless perceived (or appresented) as co-present or coterminous with their body (1989, p. 208).
“[T]he cat is present there in the flesh—specifically, as a physical thing with sensing surfaces, sense organs, etc. The stratum of sensation is not there as something beside the physical thing; what is there is a Body, a Body which has physical and aesthesiological qualities as one. Likewise, the Body is also experienced as a Body of a soul...” (Husserl, 1989, p. 185).

According to Husserl (1980), we directly perceive nonhuman animals to be conscious or minded. Indeed, few argue that nonhuman animals are completely nonconscious, even though we have no access to what nonhuman consciousness might be like. Although the pioneering ethologist Konrad Lorenz refused to speculate on the existence of nonhuman consciousness, he also noted that no one can practically deny they are conscious in their everyday dealings (Merleau-Ponty, 1968/1995, p. 259). For Husserl, our social perception of nonhuman animals is not different in kind from our social perception of other humans as minded. Not only do we perceive them as minded, but we furthermore understand their actions in intentional terms: the cat is mad at me for annoying her, the dog wants to protect me from the stranger at the door, the eagle is protecting her young. Although Husserl thought only organic bodies could be perceived in this way, research on human-robot interaction suggests that this can extend to inorganic agents (see Abubshait & Wiese, 2017; Kim, Kwon, & Kim, 2018). For Husserl, our relation to nonhuman animals is empathic—i.e., we directly perceive them as being minded, intentional, being in pain, being hungry, or being excited to see us. This empathic relation to nonhuman animals is largely prereflective. Although we may verbally reflect on the state and motivations of nonhuman animal at times (e.g., maybe the dog is barking because they are hungry?), the bulk of our mentalizing towards them is directly given in experience. Our notorious anthropomorphizing of animal behavior is a consequence of our empathic social perception; we anthropomorphize because we understand them in terms of our own bodies and experiences (Kwan, Gosling, & John, 2008).

In Ideas III, Husserl (1980) states that this perception of “alien” Leiber or lived bodies is ultimately based on the perceiver’s direct perception and experience of their own body. Husserl retains the apperceptive, analogical structure of empathy from Cartesian Meditations. However, in this text he emphasizes embodiment. Whereas in Ideas II the alien consciousness is more like the absent, dark side of the coin that I never directly perceive but nonetheless intuit as there, in Ideas III Husserl (1980) claims that aspects of its experience are also perceived. We see, as an element of our experience, “its pain when pricked, its sensuous pleasure while eating, and thus everything specifically psychic” (p. 8). For Husserl, we do not simulate their mental processes; pain and pleasure are directly perceived in the world.

We cannot have an empathic perception of plants, although they are living beings, because they lack the kind of “concrete organs” we are familiar with. Although he here uses the language of “interpretation” and “analogy,” he has not relinquished his claim that the perception of nonhuman animal minds is direct and noninferential (Meacham, 2013). Rather, he means that our empathic perception is only possible when there is a similarity in the “concrete organs” or the living body (Leib) that is presented. Even if the Venus flytrap has a “mouth” that devours its prey,

---

4 The veracity of this perception is not my concern here, nor was it Husserl’s.
5 Although Descartes (1996) is often thought of as holding such a position, see Cottingham (1978) for a refutation.
6 “Empathy” for Husserl does not have an ethical signification. For ethical implications, see Painter (2007).
7 It is important to note that, while “alien” in Cartesian Meditations only refers to other people and minds, in Ideas III it sometimes refers to both humans and nonhumans.
8 The translation has “animate organism” for Leib.
its generally stationary nature prevents it from being perceived as ensouled. As Husserl states in Addendum XXIII of the *Crisis of European sciences and transcendental phenomenology*, life is given as “the self-understanding of the biological dimension” (p. 6). It is our bodies’ forms, “instincts,” and “originary drives” that allow us to prereflectively understand other species (p. 8). Consequently, our understanding of the “what-it-is-like-ness” of animal experience is based within our own embodiment. In fact, in this short work of his final period, Husserl claims that even the concept of “organism” and the science of biology are founded on empathy (Meacham, 2013).

Husserl’s empathic theory is closest to ST. In fact, ST coincidentally happens to also go by the name of “empathy theory” (Goldman, 2006). Some ST theorists even use the language of projection. Husserl writes that my perception of the other entails something like perceiving them as though I were acting from their position. However, Husserl does not argue that we in any way imaginatively pretend to be in another’s state so as to simulate their perspective. He does not think we run an explicit simulation of others’ situations. As Goldman notes, simulation requires that the simulation be “executed with the aim of duplicating or matching” the simulated process (2006, p. 39). For Husserl, the perception of the minded other is direct in the world and does not duplicate anything. Rather, his position seems to be that our mode of embodiment founds the possibility for us to perceive others as minded.

Social cognition is direct and noninferential. Since there is no explicit simulation going on, Husserl’s empathic theory of social cognition is closest to a subpersonal or neural ST position (although he of course does not refer to the mirror neuron system). Intersubjective interaction, which is primary for IT, is enabled by the way monadic consciousnesses are, “by reciprocal empathy, …unified into a nexus” (Husserl, 1980, p. 118). Although a community or nexus of intersubjectivity develops through reciprocal empathy, monadic empathy is primary. Social cognition is ultimately based on a projection of one’s own embodiment and subjectivity into the body of the other. Our own bodies and minds are the model that we use to make the other comprehensible (Husserl, 1989). Rather than making recourse to a neural ToMM, however, Husserl refers to “apperception.” Apperception is not something I have explicit control over; it is automatic. Just as the far, occluded side of the cube is not directly present in my visual field but nonetheless is apperceived as there, the other’s mind is given to me despite not being directly present to my senses. This process is, however, still manifest to consciousness, even if its mechanism can only be isolated by the phenomenological method of bracketing.

The way this direct perception is based in my own embodiment becomes particularly clear by looking to cases of the perception of nonhuman animals. According to Husserl, I can only empathically perceive nonhuman animals that are relatively similar to me or have relatively similar organs or appendages. While chimpanzees, cats, and dogs are easy to empathize with, snails and tapeworms are not. We might understand a dog’s growl as an intentional state of anger towards my intruding presence, for example. Likewise, I do not perceive plants as minded or intentional precisely because they are so radically different in form, even though they are nonetheless living.

Abubshait and Wiese (2017) similarly note in research on human-robot interaction that humans’ view of robots as agents (rather than as mere machines) is predicated on the latter’s similarity in

---

9 In the Addendum XXIII of the *Crisis*, Husserl proposed the direct perception we have of nonhuman animal consciousnesses formed the prereflective basis for biology. Like the later Husserl, Jonas states that “life is only understood by life” (Meacham, 2013, p. X). In *The Phenomenon of Life*, Hans Jonas (1966/2001) develops this empathic approach to biology.

10 Bracketing refers to…

11 Again, Husserl is not interested in whether plants *really are* minded. He is only concerned with how we perceive them (see Husserl, 2013).
form and behavior to humans. Overall, Husserl’s empathy is closest to neural ST, even though it is not actually a neural account of social cognition.

**Heidegger’s Critique of Husserlian Empathy in *Being and Time* and Its Relation to Interaction Theory**

Heidegger’s (1926/2010) critique of Husserl’s theory of empathy in *Being and Time* strengthens my claim that Husserl is closest to ST. Heidegger attacks Husserl’s monadic approach—in a different register, we can call it his “methodological individualism.” Heidegger says, in consonance with IT, that the “explicit disclosure of the other in concern only grow[s] out of one’s primary being-with with him” (1926/2010, p. 121). The perception of other minds or “psychical life of others” is predicated on a prior being-with. Being-with is a structure of Dasein that irreducibly implicates others with whom it shares a world. Transposed to different language, the self is not enclosed off from others, but actually is intersubjective in its very nature. Others are part of the world, and the world is a part of the self: being-in-the-world. In enactivist terms, the organism is a brain-body-environment system. For Heidegger, that “environment” is irreducibly linked with others such that the self is essentially intersubjective.

Heidegger makes it clear that he thinks Husserl’s failure is in founding the perception of the other by an empathic projection from the self to that other. Empathy is supposed “to provide the first ontological bridge from one’s own subject, initially given by itself, to the other subject, which is initially quite inaccessible” (Heidegger, 1926/2010, p. 121). Empathic projection amounts to making the other a “duplicate” of the self. This fails, Heidegger thinks, not only because being-with is primary, but furthermore because being towards others is a unique structure irreducible to being towards oneself. The way I perceive and understand myself is different from the way I perceive and understand others.

Heidegger does not think, however, that empathic processes never occur. It is simply not an “original existential phenomenon” (Heidegger, 1926/2010, p. 121). Heidegger states that empathic projection occurs only when being-with breaks down. It is when I can no longer directly perceive the other and their intentions, or when those become ambiguous, that I empathically project myself into their position. It must be pointed out that Heidegger is changing the meaning of “empathy” here. Whereas for Husserl it is something achieved automatically, Heidegger thinks it is a secondary mode of social perception that only occurs when the primary, direct mode of being-with is interrupted. Because this secondary empathy is not automatic, it is similar to an inferential form of simulation. The idea is that when direct intersubjective perception becomes present-at-hand—e.g., your language tells me you are happy, but your body gestures tell me you are sad, and I have to figure out what you really feel—then I project or simulate your perspective. Overall, Heidegger is in accord with IT, but additionally thinks that explicit simulation can occur when the smooth coping of social perception breaks down due to ambiguities.

**Conclusion**

While IT is influenced by phenomenology, this does not mean that all phenomenologists agree with something like IT. Merleau-Ponty’s concept of intercorporeity has important resonances with IT (Gallagher & Hutto, 2008). Husserl’s empathic theory of social cognition is closest to ST, especially a neural ST. Nevertheless, simulation or projection for him does not occur by a ToMM. Rather, it occurs by fast, automatic processes that are able to be discerned in
consciousness by the phenomenological method of bracketing. Even though it is a process that is manifest to consciousness—even if it requires a trained perceiver to discern its operations—it is not explicit in the way that inferential ST is. I do not knowingly or intentionally simulate the other person’s mind. Rather, such simulation or projection is something that I constantly am doing without awareness. It is not a premeditated or explicit process that I control. Heidegger’s critique of Husserl’s theory of empathy further strengthens my claim that Husserl is closest to ST. Heidegger critiques Husserl’s methodological individualism. He disagrees with the idea of a projection of one’s subjectivity into the other and poses his own theory of being-with as an answer to this view; being-with is a mode of Dasein that irreducibly links it to others.

By working out just what Husserl’s empathy could mean in the context of the discourse of social cognition within contemporary cognitive science, I hope to contribute to the greater dialogue between phenomenologists and philosophers of mind and cognitive science. Furthermore, by showing that Husserl is close to an ST position, or that Heidegger is closer to IT, I am showing the contemporary relevance of phenomenology. It is a living tradition that has asked similar questions contemporary philosophers of mind and cognitive science are asking in the 21st century. This is a particularly important point given the contemporary marginalization of so-called Continental philosophy in the US. Finally, given Heidegger’s early critique of Husserlian empathy in 1926, it would be an interesting question for further research to determine whether all subsequent phenomenologists hold a position in tune with IT. While Gallagher (2013) argues (successfully, I think) that Merleau-Ponty does, what about other phenomenologists, such as Sartre, de Beauvoir, and Michel Henry?
References


