

Forthcoming in *Consciousness and Cognition*. Please quote only from published version.

## **Simulation, projection and empathy**

Dan Zahavi  
Center for Subjectivity Research  
Department of Media, Cognition and Communication  
University of Copenhagen  
Njalsgade 140-142  
DK-2300 Copenhagen S  
Denmark  
[dza@hum.ku.dk](mailto:dza@hum.ku.dk)  
Phone: +45 35328690  
Fax: +45 35328681

## Abstract

Simulationists have recently started to employ the term "empathy" when characterizing our most basic understanding of other minds. I agree that empathy is crucial, but I think it is being misconstrued by the simulationists. Using some ideas to be found in Scheler's classical discussion of empathy, I will argue for a different understanding of the notion. More specifically, I will argue that there are basic levels of interpersonal understanding - in particular the understanding of emotional expressions - that are not explicable in terms of simulation-plus-projection routines.

Keywords: Simulation, theory of mind, empathy, phenomenology, social cognition, intersubjectivity, Scheler

### Simulation, projection and empathy

#### 1. *Simulating others*

Different versions of the simulation theory are currently defended by figures like Goldman, Gordon, and Gallese. Although simulationists of various denominations hold that we don't need a general psychological theory in order to understand others, the opinions differ somewhat when it comes to a more positive account of what simulation amounts to. However, a frequent claim has been that we can gain information about other minds by pretending to be in their "mental shoes." We can use our own mind as a model, use it to "mirror" or "mimic" the minds of others. That is, our ability to predict and explain the actions of others is frequently taken to depend crucially on our ability to *project* ourselves imaginatively into their situation. It is no coincidence that Goldman, whose version of simulationism I – for reasons that will become evident in a moment – primarily will be discussing, considers a more apt name for the whole process to be simulation-plus-projection (Goldman 2006, 40).

In making their case against the theory-theory of mind, simulationists have frequently appealed to certain persuasive examples. Consider the following scenario. You are engaged in a play of chess. If you wish to predict your opponent's next move, how do you go about doing that? Rather than relying on some general theory about how chess players act in certain specific situations, one plausible suggestion is that you simply imagine yourself being in your opponent's position and after having decided what you would choose to do, you then predict that he would do the same.

Of course, nobody is claiming that simulation is a foolproof method, and there might be various confounding factors – if you happen to be a chess master, and your opponent is a novice, it wouldn't make much sense to ascribe your preferred move to him. Thus, if the simulation routine is to succeed, the attributor must quarantine his own idiosyncratic desires and beliefs (Goldman 2006, 29). If he fails to do so, he will not reach an understanding of the other, but merely come to understand how he himself would have acted under different circumstances. But even granted this complication, the example still seems to favor the simulation approach rather than the theory-theory approach.

On closer consideration, however, there is something strikingly dissatisfying about the example. It should be obvious that our understanding of others comes in many shapes and forms. Even if no single model can do justice to the whole variety, any account that aspires to be the core theory, the default account of how we gain access to other minds, should at the very least be able to capture prototypical everyday situations. But how representative is the kind of interpersonal understanding we encounter in chess playing? Consider the following twist to the story: While facing your opponent and being absorbed in trying to figure out what his next move is going to be, he startles you by suddenly removing all the pieces from the board with a sweep of his hand while his face is contorted in anger. How plausible is it to claim that you need imaginatively to put yourself in his shoes in order to determine that you are faced with a minded opponent, and that he is frustrated and angry, rather than, say, blissfully happy?

As I see it, any convincing theory of social cognition should be able to account for our face to face encounters with others, should be able to deal with our faced-based "mindreading" abilities. However, I

have difficulties detecting any crucial role for pretense and imaginative projection on this very basic level. As Wittgenstein once remarked, “Do you look into *yourself* in order to recognize the fury in *his* face?” (Wittgenstein 1980, § 927). More generally speaking, I think there has been a tendency to focus too narrowly on our ability to *explain* and *predict* the actions of others, as if our social life was exclusively a question of ascribing causally efficacious inner mental states (mainly beliefs and desires) to others in order to facilitate our prediction, explanation, and controlling of their behavior. This is not to deny that we occasionally do engage in either, say, if we are asked by a third party to provide reasons for somebody’s behavior or if we happen to observe somebody behave unexpectedly or puzzling. But the “engineering” look on sociality, to use a recent expression of RoCHAT (2008), has become too dominant.

In his recent book *Simulating Minds* Goldman concedes that an account of mindreading should be able to cover the whole range of mental states, including sensations, feelings and emotions. It shouldn’t just address the issue of belief-ascription (Goldman 2006, 20). As a consequence, Goldman distinguishes what he calls low-level mindreading from high-level mindreading (Goldman 2006, 43). Goldman is not alone among simulationists in making such a distinction. One can find a very comparable move in Stueber’s recent book *Rediscovering Empathy* where Stueber distinguishes *basic empathy*, which he defines as a mechanism of inner imitation that underlies our theoretically unmediated quasi-perceptual ability to recognize other creatures directly as minded creatures, from *reenactive empathy*, which he defines as involving the use of our cognitive and deliberative capacities to reenact or imitate the thought processes of others (Stueber 2006, 20-21).

But let us stay with Goldman for the moment. Goldman argues that we need to recognize the existence of a low-level ability to attribute basic emotions such as fear, anger and disgust to others on the basis of their facial expressions (Goldman & Sripada 2005). This ability is simple, primitive and automatic. In order to explain this kind of basic mindreading, Goldman considers and appraises four different models – the *generate and test model*, the *reverse simulation model*, the *reverse simulation with as-if loop model*, and the *unmediated resonance model*. Goldman shows a certain preference for the latter model (Goldman 2006, 132), which appeals to interpersonal mirroring mechanisms and he suggests that the perception of the target’s emotional expression might directly trigger activation of the neural substrate of the same type of emotion in oneself, thereby making the process a kind of unmediated matching (Goldman 2006, 128). One obvious question to ask though, is why such a low-level direct mirroring process should count as a form of simulation? For one, it doesn’t involve any pretend states. But on Goldman’s view, pretend states are *not* essential to simulation; rather they are only to be found in high-level forms of mindreading. In his view, a process P can be called a simulation of another process P\*, as long as P duplicates, replicates or resembles P\* in some significant respects (Goldman 2006, 36). Given that basically everything resembles everything else in some respect, it is obviously important to specify what is meant by significant. On a minimal reading, it is the existence of matching emotions in target and attributor that makes the process one of simulation. As Goldman puts it in joint article with Sripada: “In the case of the successful simulation, the experienced state matches that of the target.” (Goldman & Sripada 2005, 208). Indeed, if I understand Goldman right, he occasionally suggests that the observation of another’s emotional expression automatically triggers the experience of that emotion in myself, and that this first-personal experience then serves as the basis for my third-person ascription of the emotion to the other. As he writes – in the context of discussing disgust expressions – “the evidence points towards the use of one’s disgust experience as the causal basis for third-person disgust attributions” (Goldman 2006, 137). Goldman also appeals to neuroscientific findings according to which the same neural substrate is activated regardless of whether we experience an emotion ourselves or recognize the emotion in others, and to data indicating that impairments in experiencing a given emotion are paired with impairments in recognizing that emotion in others (Goldman 2006, 110, 115).

At one point, Goldman writes that he considers mindreading an extended form of *empathy* (Goldman 2006, 4). As already indicated, he is not the only simulationist who has started to employ that term when characterizing our basic mindreading abilities. Indeed, recently, it has even been claimed that simulationists are today’s equivalents of empathy theorists (Stueber 2006, ix). I, for my part, also think that empathy conceived as a primitive form of interpersonal understanding plays a crucial role, but I have some worries about the way it is being construed by simulationists. More specifically, what I intend to do

in the following is to question whether simulationism really offers us a compelling account of *empathy*. My way of approaching this issue will be somewhat indirect, however, since I simultaneously wish to fry another fish. In *Simulating Minds* Goldman starts out by giving us a quick overview of the modern debate. On Goldman's construal the debate goes back roughly 50 years to Ryle and Wittgenstein, though Goldman does acknowledge that simulationist themes can be found scattered in earlier theorists, such as Lipps and Dilthey (Goldman 2006, 18). What I find surprising about this overview is that it contains no reference to the discussion of social cognition found in the phenomenological tradition. I am not merely thinking of the significant and substantial contribution to an understanding of intersubjectivity found in the works of Husserl, Heidegger, Sartre and Merleau-Ponty, but also of more specific works such as Edith Stein's *Zum Problem der Einfühlung*, Aaron Gurwitsch's *Die mitmenschlichen Begegnungen in der Milieuwelt* and Max Scheler's classic *Wesen und Formen der Sympathie*. There is obviously no way I can correct this omission here, but I intend to take a closer look at Scheler's discussion of empathy, since it contains ideas that are of direct contemporary relevance.<sup>1</sup> In particular, it articulates what I would consider a *non-simulationist conception of empathy*.

Let me make it clear that my dissatisfaction with simulation theory – when appraised as an account of our most basic understanding of other minds – is not meant as an indirect defense of some version of the theory-theory. Rather, in my view the forced choice between theory-theory and simulation-theory is a false choice. By criticizing the simulation theory I hope to make room for what I consider to be an interesting and significant – though overlooked – alternative.

## 2. Scheler on empathy

Scheler's work *Wesen und Formen der Sympathie* (1923) is frequently listed as an example of a phenomenological investigation of emotional life. But in addition to presenting us with careful analyses of various emotional phenomena, the work must also be considered a significant contribution to the phenomenology of intersubjectivity and social cognition. Indeed, on Scheler's account, our relation to others is not primarily a theoretical matter, but rather something rooted in our emotional life.

I won't be able to cover all aspects of Scheler's discussion in the following, but let me start with some of his phenomenological distinctions. Scheler discusses two groups of cases. In the first he asks us to compare the following three situations:

Consider first the situation where you see the face of a crying child, but rather than seeing it as expressing discomfort or distress, you merely see a certain distortion of the facial muscles, i.e., you basically don't see it as emotionally expressive. Compare this (pathological case) with the situation where you see the same face as emotionally expressive, but without feeling any compassion, i.e., while remaining indifferent. And finally consider the situation where you also feel compassion or concern for the child. For Scheler, the last situation counts as a case of sympathy, which he considers an ethically relevant act. But in order to feel sympathy, in order to feel compassion with, say, somebody's suffering, you need to believe that the other is indeed suffering. More basic than sympathy is a way of experiencing others, which Scheler designates with various terms,<sup>2</sup> and which I in the following will render as

---

<sup>1</sup> For further discussions of phenomenological theories of intersubjectivity, see Zahavi 1996, 2001a, 2001b, 2002.

<sup>2</sup> Unfortunately, Scheler doesn't stick to a single term when referring to this basic form of understanding. Rather he uses terms such as *Nachfühlen* (reproduction of feeling), *Nachleben* (reproduction of experience), *Nacherleben* (visualizing of experience), *Verstehen* (understanding), or *Fremdwahrnehmung* (perception of other minds) (Scheler 1954, 9, 238). In some of the cases, the English translation might not be ideal, but Scheler himself must also be blamed for the inevitable confusion. How can *Nachfühlen* and *Fremdwahrnehmung* refer to one and the same phenomenon? As we will see in a moment, Scheler rejects the view that our understanding of the emotional experience of others is based on an imitation or reproduction of the emotion in question, but why does he then himself use a term like *Nachfühlen*? The fact remains, however, that Scheler is quite unequivocal in his rejection of the view that our understanding

*empathy*.<sup>3</sup> In short, whereas empathy has to do with a basic understanding of expressive others, sympathy adds care or concern for the other.

Now, apart from stressing the difference between empathy and sympathy, the point of Scheler's example is also to remind us that it is possible to empathize with somebody without feeling any sympathy (Scheler 1954, 8-9). Just think of the skilled interrogator or the sadist. Sadistic cruelty doesn't merely consist in failing to notice the other's pain, but in empathically enjoying it (Scheler 1954, 14).

Consider now a second group of cases. You might enter a bar and be swept over by the jolly atmosphere. A distinctive feature of what is known as *emotional contagion* is that you literally catch the emotion in question (Scheler 1954, 15). It is transferred to you. It becomes your own emotion. Indeed you can be infected by the jolly or angry mood of others without even being aware of them as distinct individuals. But this is precisely what makes emotional contagion different from both empathy and sympathy. In empathy and sympathy, the experience you empathically understand or sympathetically care for remains that of the other. In both of the latter cases, the focus is on the other, the distance between self and other is preserved and upheld. Another distinctive feature of emotional contagion is that it concerns the emotional quality rather than the object of the emotion. You can be infected by cheerfulness or hilarity, without knowing what it is about. This is what makes emotional contagion different from what Scheler calls *emotional sharing*. Think of the situation where a father and mother stand next to the corpse of a beloved child. For Scheler, this situation exemplifies the possibility of sharing both an emotion (sorrow or despair) and the object of the emotion. But emotional sharing must on its part still be distinguished from both empathy and sympathy. Consider the situation where a common friend approaches the despairing parents. He can empathize or more likely sympathize with their sorrow, without experiencing the despair in question himself, which is why his state of mind differs qualitatively from either of theirs. Indeed their sorrow and his empathy or sympathy are clearly two distinct states. Their sorrow is the intentional object of his empathy or sympathy (Scheler 1954, 12-13).

So far, we find Scheler insisting on the difference between emotional contagion, emotional sharing, empathy and sympathy. Using more contemporary terms, he would also subscribe to the difference between a basic and automatic form of empathy (which is what he is mainly interested in analyzing) and a more advanced cognitive type which might involve perspective-shifting, imaginative projection or inferential attribution.

On Scheler's view, empathy isn't simply a question of intellectually judging that somebody else is undergoing a certain experience. It is not the mere thought that this is the case; rather, Scheler defends the view that we are empathically able to *experience* other minds (Scheler 1954, 9). It is no coincidence that Scheler repeatedly speaks of the perception of others (*Fremdwahrnehmung*), and even entitles his own theory *a perceptual theory of other minds* (Scheler 1954, 220). Empathy is a basic, irreducible, form of intentionality that is directed towards the experiences of others. It is a question of understanding other experiencing subjects. But this doesn't entail that the other's experience is literally transmitted to us. Rather, it amounts to experiencing, say, the other person's emotion without being in the corresponding emotional state yourself. You experience the emotion in a way that differs from the way you would experience the emotion if it were your own. Thus empathy doesn't entail that we ourselves undergo the emotion we observe in the other. It might – as a consequence – but it is not a requisite. We might encounter a furious neighbor and become furious ourselves, but our empathic understanding of our neighbor's emotion might also elicit the reverse response, namely the feeling of fear. In either case,

---

of the emotional experiences of others requires us to have the same emotion ourselves (Scheler 1954, 9-10), and one must consequently simply note that his choice of terms left something to be desired.

<sup>3</sup> "Empathy" is usually considered the standard translation of *Einfühlung*, and it so happens that Scheler himself only used the latter term rather sparingly and when he did frequently rather dismissively. However, Scheler's reservation was mainly due to his dissatisfaction with Lipps' projective theory of empathy, and for want of a better term, I have decided to use "empathy" as the best way of capturing what Scheler was referring to when he spoke of a basic experience of others. It is telling that other contemporary phenomenologists reasoned in a similar way. Both Stein and Husserl also referred to Scheler's theory of empathy (*Einfühlung*) (Stein 1989, 27, Husserl 1999, 147).

however, our emotional reaction is exactly that – a reaction. It is a consequence of our understanding of the other's emotion, and not a precondition or prerequisite for this understanding, contrary to what Goldman seems to be suggesting. Indeed how plausible is it after all to claim that I need to become furious myself, if I am to recognize the fury in the face of my assailant. The same goes for sympathy. As Goldie recently put it, sympathy involves *caring* about rather than *sharing* the other's sufferings (Goldie 1999, 420). It doesn't feel (physically) painful to sympathize with, i.e. care for, somebody who is in physical pain.

On Scheler's account empathy neither involves some kind of analogical inference, nor some kind of projection, simulation or imitation (Scheler 1954, 12). He rejects the proposal that emotional contagion or mimicry should be the paradigm of automatic empathy (cf. Hodges and Wegner 1997, Darwall 1988, 264-266), and ridicules the suggestion that empathy should be based on a direct association between cues coming from others and one's own similar past experiences, as if one's ability to empathize with, say, parents who have lost their only child, would not only be conditional on one having oneself gone through such an ordeal, but would even entail an actual reproduction of that past experience. As he points out, if that were to happen, we would be faced with something very different from empathy, namely the phenomenon of what has more recently been called *egoistic drift*. Ultimately, Scheler insists that it is necessary to reject the projective theory of empathy in all its forms (Scheler 1954, xlvi), and he is also very dismissive of the attempt to account for the experience of others in terms of some imaginative transformation. In basic empathy the focus is on the other, on his thoughts and feelings, and not on myself, not on how it would be like for me to be in the shoes of the other (Scheler 1954, 39).

### 3. Analogy and expression

Scheler's investigation of empathy and interpersonal understanding is restricted to the personal level. He is not concerned with the various sub-personal mechanisms that might be involved in interpersonal understanding. His main objection against competing theories seems to be that they are phenomenologically inadequate and that they fail to do justice to our actual experience. On this basis, it might be natural to conclude that his project is exclusively descriptive. It seeks to describe how we experience other minds, but it doesn't address the normative question concerning the justification or validity of that experience. But this is a mistake. Not only does Scheler provide more systematic arguments against the argument from analogy – the grandmother of the kind of simulation theory espoused by Goldman (cf. Gordon and Cruz 2003) – but ultimately he also seeks to provide an account of the nature of experience that makes it comprehensible how we can experience other minds, and why such an experience can be justified.

According to Scheler, the argument from analogy presupposes that which it is meant to explain. In order for the argument to work, there has to be a similarity between the way in which my own body is given to me, and the way in which the body of the other is given to me. But if I am to see a similarity between, say, my laughing or crying and the laughing or crying of somebody else, I need to understand the bodily gestures and behavior as expressive phenomena, as manifestations of joy or pain, and not simply as physical movements. If such an understanding is required for the argument of analogy to proceed, however, the argument presupposes that which it is supposed to establish. To put it differently, in some cases we do employ analogical lines of reasoning, but we only do so when we are already convinced that we are facing minded creatures but are simply unsure about precisely how we are to interpret the expressive phenomena in question (Scheler 1954, 240, cf. Gurwitsch 1979, 14, 18).

In addition, Scheler also questions two of the basic presuppositions behind the whole approach. First, it assumes that my point of departure is my own consciousness. This is what is at first given to me in a quite direct and unmediated fashion, and it is this purely mental self-experience that is then taken to precede and make possible the recognition of others. Second, the argument also assumes that we never have direct access to another person's mind. We can never *experience* her thoughts or feelings. We can only infer that they must exist based on that which is actually given to us, namely her bodily behavior. Although both of these two assumptions might seem perfectly obvious, Scheler rejects both. As he points

out, as philosophers it is our duty to question the obvious. We should pay attention to what is actually given, rather than letting some theory dictate what can be given (Scheler 1954, 244). On his view, the argument from analogy underestimates the difficulties involved in self-experience and overestimates the difficulties involved in the experience of others (Scheler 1954, 244-246). We should not ignore what can be directly perceived about others and we should not fail to acknowledge the embodied and embedded nature of self-experience. Scheler consequently denies that our initial self-acquaintance is of a purely mental nature, as if it anteceded our experience of our own expressive movements and actions, and as if it took place in isolation from others. He considers such an initial purely internal self-observation a mere fiction.

Scheler also denies that our basic acquaintance with others is inferential in nature. As he argues, there is something highly problematic about claiming that intersubjective understanding is a two-stage process of which the first stage is the perception of meaningless behavior and the second an intellectually based attribution of psychological meaning. Scheler argues that we in the face-to-face encounter are neither confronted with a mere body, nor with a pure soul, but with the unity of an embodied mind. He speaks of an “expressive unity” (*Ausdruckseinheit*), and claims that the notion of behavior is a psychophysically undifferentiated concept. It is only subsequently, through a process of abstraction, that this unity is divided and our interest then proceed “inwards” or “outwards” (Scheler 1954, 218, 261).

Foreshadowing something that both Sartre and Lévinas would later discuss in more detail, Scheler writes that I experience, say, the hostility or love in the expression of another’s gaze long before I can specify the color of his eyes (Scheler 1954, 244). Indeed on Scheler’s account, our primary knowledge of nature is knowledge of expressive phenomena. He finds this claim corroborated by newborns’ preferential interest for expressive faces and human voices. This knowledge of a living world is taken to precede our knowledge of a dead and mechanical world. So for Scheler, it is not the case that we first see inanimate objects and then animate them through a subsequent addition of mental components. Rather, at first we see everything as expressive, and then we go through a process of de-animation. Learning is a question of *Entseelung* rather than *Beseelung* (Scheler 1954, 239).

To sum up, Scheler opposes the view according to which our encounter with others is first and foremost an encounter with bodily and behavioral exteriorities devoid of any psychological properties. According to such a view, which has been defended by behaviorists and Cartesians alike, behavior, considered in itself, is neither expressive nor significant. All that is *given* is physical qualities and their changes. Seeing a radiant face means seeing certain characteristic distortions of the facial muscles. But as Scheler points out, this account presents us with a distorted picture, not only of behavior but also of the mind. It is no coincidence that we use psychological terms to describe behavior and that we would be hard pressed to describe the latter in terms of bare movements. In the majority of cases, it is quite hard (and artificial) to divide a phenomenon neatly into its psychological and behavioral aspect: think merely of a groan of pain, a handshake, an embrace, a leisurely stroll. In his view, affective and emotional states are not simply qualities of subjective experience, rather they are given *in* expressive phenomena, i.e., they are expressed in bodily gestures and actions, and they thereby become visible to others. In fact, only thereby do they gain full visibility for ourselves, since a repression of the emotional expression will necessarily lead to a reduction of the felt quality of the emotion (Scheler 1954, 251).

Instead of attempting to secure an access to the minded life of others through technical detours, Scheler argues that we need a new understanding of the given. If the realm of expressive phenomena is accepted as the primary datum or primitive stratum of perception, the access to the mind of others will no longer present the same kind of problem. What we see is the body of the other as a field expressive of his or her experiences (Scheler 1954, 10). Indeed, on Scheler’s view, expressive phenomena – in particular facial expressions and gestures, but also verbal expressions – can present us with a direct and non-inferential access to the experiential life of others. As he writes in what must count as a *locus classicus*:

For we certainly believe ourselves to be directly acquainted with another person’s joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands, with his love in his look of affection, with his rage in the gnashing of his teeth, with his threats in the clenching of his fist, and with the tenor of his thoughts in the

sound of his words. If anyone tells me that this is not 'perception', for it cannot be so, in view of the fact that a perception is simply a 'complex of physical sensations', and that there is certainly no sensation of another person's mind nor any stimulus from such a source, I would beg him to turn aside from such questionable theories and address himself to the phenomenological facts (Scheler 1954, 260).

It should by now be clear that Scheler takes a solution to the problem of other minds to require a correct understanding of the relation between mind and body. And of course, the mind-body relation in question is not the mind-brain relation. Scheler is not concerned with the search for the neural correlates of consciousness. Rather he is interested in the relation between experience and expressive behavior, and he obviously takes the connection to be quite close.<sup>4</sup>

#### 4. *Objections and clarifications*

Scheler might not provide us with a fully developed theory of social cognition,<sup>5</sup> but enough has been said to make it clear that his understanding of empathy differs markedly from the views espoused by recent simulationists, who again and again have claimed that empathy is a projective process, that it involves the imaginative adaptation of another person's point of view or some form of inner or mental imitation (cf. Goldman 2006, 40, Stueber 2006, 15, 28).

Does our most basic understanding of other minds – including our face-based recognition of emotions in others – involve simulation? One proposal was that all that was required for A to count as a low-level simulation of B was for A to resemble B in some significant respect (Goldman 2006, 36). If we apply this principle on the personal level, it must presumably mean that whereas there is no simulation involved in my perception of a green triangle, since my perception is neither greenish nor triangular, there is simulation involved when I observe somebody who is in pain or furious, since this somehow requires me to be in pain or furious. Indeed, if the simulation is to serve any explanatory purpose, my own felt pain or fury must precede rather than follow my conscious recognition of the pain or fury in the other.

As already indicated, I have difficulties with this view. Not only do I find the proposed definition of simulation far too watered-down, I also find the personal-level description phenomenologically implausible, and finally I have serious doubts about the basic assumption behind the whole setup. Simply put, the general idea is that if I want to reach an understanding of other minded creatures, be it their emotions, their desires, their beliefs etc., I must project what I know about my own mind into their minds. Why is this circuit through self deemed necessary? I need to project what I know about my own mind into the mind of others, because the only mind I have any direct and non-inferential knowledge of is my own, whereas the minds of others are closed territory. They remain concealed and hidden. All that I have experiential access to is their behavior. Without a commitment to a view like this, the constant appeal to internal imitations and projections would make little sense. This commitment is rarely spelled out explicitly, but for a rather unequivocal endorsement, consider the following quote by Baron-Cohen, an erstwhile defender of the theory-theory approach. Baron-Cohen writes, "Empathy involves a leap of imagination into someone else's headspace. While you can try to figure out other persons' thoughts and feelings by reading their faces, their voices, and their postures, ultimately their internal worlds are not transparent, and to climb inside their heads requires imagining what it must like to be them" (Baron-Cohen 2005, 170).

This quote seems to indicate that Baron-Cohen must now be counted as belonging to the simulationist camp, but I don't think it is very hard for a defender of the theory-theory approach to endorse part of what Baron-Cohen is saying here. Indeed, despite their many differences, the theory-theory of

---

<sup>4</sup> See Zahavi 2007 and Overgaard 2007 for more on this topic.

<sup>5</sup> Many further ideas can be gleaned from other phenomenologists such as Husserl, Heidegger, Gurwitsch, Merleau-Ponty and Sartre. For a concise discussion of how other phenomenologists have dealt with the issue of intersubjectivity, see Zahavi 2001a.



mind and the simulation theory of mind both deny that it is possible to *experience* other minded creatures. It is precisely because of the absence of an experiential access to others that we need to rely on and employ either theoretical inferences or internal simulations. Both accounts share the view that the minds of others are hidden, and they consider one of the main challenges facing a theory of social cognition to be the question of how and why we start ascribing such hidden mental entities or processes to certain publicly observable bodies. If one accepts this outlook, it is difficult to avoid the verdict that simulationism remains stuck in an egocentric predicament. Its focus remains intrapersonal and it is ultimately unable to account for real interpersonal understanding.<sup>6</sup> To put it differently, the simulation-plus-projection procedure imprisons me within my own mind – consider that Goldman explicitly defines projection as “the act of assigning a state of one’s own to someone else” (Goldman 2006, 40) – and prevents me from ever achieving a true understanding of *others*.<sup>7</sup>

Let me end by considering three possible simulationist rejoinders.

A.

One obvious objection to consider is the following. Doesn’t my criticism of simulationism entail a confusion of levels? Couldn’t simulationists accept the phenomenological description *tout court* but simply insist that this description remains a personal-level description, one that doesn’t consider the underlying sub-personal mechanisms, which are the ones that simulationists are interested in. To be more specific, couldn’t it be argued that most if not all low-level simulation is non-conscious and sub-personal, for which reason a phenomenological motivated criticism loses its force? Moreover, isn’t the existence of such sub-personal simulation routines now being confirmed by various neuroscientific findings? Although Goldman in many of his writings has referred explicitly to the experiential level in his account of simulation, I suspect this is how he would respond if pressed on the issue.

In reply, let me make it clear that I am obviously not denying that my conscious recognition of, say, emotional expressions in others is subserved by various sub-personal mechanisms. The crucial question though is to what extent such mechanisms involve routines that merit the name simulations.<sup>8</sup> There is a crucial difference between claiming that my recognition of a certain emotion in you requires me to experience the very same kind of emotion immediately prior to ascribing it to you and claiming that the same neural substrate subserves both the experience of an emotion and the recognition of the same kind of emotion in others. The latter claim is considerably weaker. Even though damage to the brain area in question might impede attributions of emotions to both self and other, the fact that these impairments are paired, the fact that the same brain area is involved in both processes, doesn’t show that there is simulation involved.

Furthermore, does it really make sense to apply personal-level terms such as anger, happiness, sadness, fear, disgust etc. to non-conscious, sub-personal mechanisms? Does it make sense to claim that the presence of a specific neural activity – even in the absence of the characteristic experiential and behavioral manifestation – is sufficient to constitute a token of the emotion in question? If one foregoes this questionable attempt – the presence of a specific non-conscious process might be a necessary and enabling condition, but it cannot be sufficient, since emotions are defined in terms of their experiential

---

<sup>6</sup> Of course, simulationists readily concede that we can compensate for our personal bias, and quarantine those of our own states that differs from the target, and that this is indeed required if we are to understand the other, but this also seems to get things backward. How can I quarantine those states of my mind that differ from the target’s, unless I already know his or her mind?

<sup>7</sup> I recognize that simulationism is something of an umbrella term, and that quite a few simulationists would distance themselves from what they consider to be Goldman’s overly Cartesian construal of simulation. However, my criticism is not merely directed at Goldman’s position but at any version of simulation theory insofar as it 1) aims to account for our most basic understanding of other minds, and 2) takes simulation to involve imitation, projection or imagination.

<sup>8</sup> For a more extensive criticism, see Gallagher and Zahavi 2008, chap. 9.

qualities and behavioral display – the basis for any reference to simulation is further weakened, since it is no longer legitimate to speak of matching emotions in target and attributor.

Whereas it is relatively easy to understand what is meant by simulation as long as the term denotes personal-level processes, the use of the term to denote sub-personal processes might increase the plausibility of the claim that simulation routines are pervasive but at the cost of making the meaning of the term quite unclear. Why speak of the involved sub-personal processes in terms of simulation, and not rather in terms of, say, resonance mechanism that detect and represent the target's emotional state? The latter phrasing would avoid the misleading connotations of the former.

B.

Another possible retort to the claim that we are able to *experience* other minds might be that such a claim is preposterous, since it overlooks the fact that we do not have the same kind of access to the minds of others that we have to our own. To phrase it differently, any convincing account of our understanding of others must respect the asymmetry between self-ascription and other-ascription of mental states, must respect that whereas I enjoy a first-person perspective on my own mental life, I do not have first-personal access to the minds of others.

I fully agree. When I claim that we are able to *experience* others, and as a consequence do not exclusively have to rely on and employ theoretical inferences, internal simulations or imaginative projections, I am not claiming that I can experience the other in precisely the same way as she herself does, nor that the other's consciousness is accessible to me in precisely the same way as my own is. Second- (and third-) person access to psychological states *do* differ from first-person access. But we shouldn't make the mistake of restricting and equating experiential access with first-person access. It is possible to experience minds in more than one way. When I experience the facial expressions or meaningful actions of an other, I am *experiencing* foreign subjectivity, and not merely imagining it, simulating it or theorizing about it. I am experiencing the other him- or herself, and not merely some theoretical or imagined construct, some simulation or simulacrum. For comparison, consider the case of perception. 1) I can theorize about the taste of Papaya juice, 2) I can imagine what its taste must be like, and 3) I can drink and experience the taste. To see the sorrow and pain in the other's face has more in common with 3, than with 1 or 2.

The fact that I can be mistaken and deceived is no argument against the experiential character of the access. As Moran has pointed out, the second- (or third-) person access only "falls short" of the first-person access if it is assumed that the latter is privileged and that it is the internal aspiration of the former to approximate the latter as closely as possible (Moran 2001, 157). We should recognize that each type of access has its own strengths and weaknesses.

Moreover, the fact that my experiential access to the minds of other differs from my experiential access to my own mind is not an imperfection or shortcoming. On the contrary, it is a difference that is constitutional. It is precisely because of this difference, precisely because of this asymmetry, that we can claim that the minds we experience are *other* minds. Indeed, a more precise way of capturing what is at stake is by saying that we experience bodily and behavioral expressions as expressive of an experiential life that transcends the expression. There is so to speak more to the mind of the other than what we are grasping, but this doesn't make our understanding non-experiential.

To sum up, I think it is crucial to avoid construing the mind as something visible to only one person and invisible to everyone else. The mind is not something exclusively inner, something cut off from the body and the surrounding world. Some mental states are directly apprehended in the bodily expressions of people whose mental states they are. Indeed, our understanding of affective and emotional states, such as sorrow, shame, love, gratitude, hate, pity, disgust, fear, pride etc., is informed and influenced by their bodily manifestations. The latter is partly constitutive of the meaning of the former. Expressive movements and behavior is soaked with the meaning of the mind; it reveals the mind to us. This is not behaviorism. The idea is not to identify mental states with or reduce them to behavior, nor does it rule out that some experiential states are covert but not all mental states can lack a natural expression if intersubjectivity is to get off the ground.

Tomasello has recently proposed that our social cognition takes three forms. We can understand others 1) as animate beings, 2) as intentional agents, and 3) as mental agents. In his view, the ontogenetic relevance of this tripartition is straightforward. Whereas infants are able to distinguish animate beings from non-animate beings already from birth onwards, they are only able to detect goal-directed behaviour from around 9-12 months of age (as evidenced in phenomena such as joint attention, gaze following, joint engagement, imitative learning, etc.), and they only become aware of others as mental agents with beliefs that might differ from their own at around 4-5 years of age. Why does the last step take so much longer? The answer provided by Tomasello is twofold. On the one hand, he calls attention to the different role of expressive behaviour. Whereas the animacy (and certain emotions) of others is directly expressed in their behaviour, intentions are also expressed in actions, but are at the same time somewhat divorced from them, since on occasion they may remain unexpressed or be expressed in different ways. Finally when it comes to thoughts and beliefs these might lack natural behavioural expressions altogether (Tomasello 1999, 179), which is what makes them so much more difficult to grasp. On the other hand, Tomasello argues that the more advanced form of social cognition emerges as late as it does because it depends on prolonged real-life social interaction (Tomasello 1999, 198). In order to understand that other persons have beliefs about the world that differ from their own, children need to engage them in discourses where these different perspectives become clearly apparent, be it in disagreements, misunderstandings, requests for clarification or reflective dialogues (Tomasello 1999, 176, 182). Tomasello is certainly right in pointing to the fact that our understanding of others gradually becomes more sophisticated, and that there are dimensions of the mind that are not as readily accessible as others. Moreover, he is also right in pointing to the cultural and social dimension of this developmental process. Rather than being the result of an automatic maturation of certain innate cognitive modules, it seems plausible to view these more sophisticated forms of social cognition as abilities that develop in tandem with increasingly complex forms of social interaction.

C.

As I pointed out in the beginning, it is important for a theory of social cognition to consider and account for very basic forms of interpersonal understanding, thus the discussion of face-based emotion recognition. However, a too narrow focus on expressions would also be misleading. Indeed this might be one of the weaknesses in Scheler's proposal. First of all, we shouldn't forget that emotions are intentional. They are about something, and in order to understand the emotions fully, it is not enough simply to pay attention to their expressions; we also need to look at the context, in order to determine what they are about. Secondly, to engage with others as thematic objects of cognition is the exception rather than the rule. When we encounter others, we encounter them in a shared world, and our understanding of each other takes place in a specific context and situation (Gurwitsch 1979, 35-36, 95, 106). It is precisely within such situations that expressive phenomena occur. When working or conversing with my partner, he might shake his head or wrinkle his brow. But these facial expressions and bodily gestures are not unambiguous. They do not reveal psychological states simply or uniformly. The "same" shaking of the head can take on different meanings in different situations. But this is rarely a problem, since the expressions occur in a given context, and our understanding of the context, of what comes before and after, facilitates and aids our understanding of the expression. To put it differently, it is crucial to avoid what might be called a snapshot view of emotional understanding. Outside psychology labs we very rarely encounter expressions in isolation.

But doesn't this concession raise doubts about the extent to which our understanding of others' emotional expressions might be direct and theoretically unmediated? It obviously depends on how one understands the notions of direct and theoretical. For comparison consider the case of vision. Vision usually counts as the paradigm of direct experience. I can theorize about Aurora Borealis, I can imagine what it must be like to see it, and I can see and experience it in all its splendor, but – and this is an old insight – when we perceive an object, we perceive it in a perceptual field. We are conscious of it in a particular setting, and the way it is given to us is influenced by what is co-given with it. I see no conflict between this insight and the claim that the perceptual object is directly given. Similarly, consider the case

of utensils, say, an iPod. For something to be intended as an iPod, for something to appear as an iPod, a whole network of equipmental contexture, to use a Heideggerian phrasing, must be in place. But again, this fact doesn't make the perception of an iPod indirect and theory-laden in the same way as our positing of black holes or sub-atomic particles. It doesn't make our access to the iPod non-experiential; it doesn't turn the iPod into an unobservable and theoretically postulated entity.

To put it differently, there is no contradiction in defending the direct and contextual character of perception at the same time. Similarly, when it comes to interpersonal understanding. I would consequently maintain that there are ways of understanding others that are as direct as our perception of medium sized objects – which is not to deny that there are also important differences between our perception of objects and our understanding of other subjects. Granted that one accepts this line of thought, the fact that the behavioural manifestation of emotions may be socially influenced and, as a result, vary considerably from culture to culture doesn't threaten my general point. It is true that the expression of, say, sorrow might differ among Ifaluks, Balinese and Americans (cf. Nussbaum 2001, 157), just as there might be no direct English equivalent to the Czech notion of *Litost*, which designates an emotional state of torment created by the sudden insight into one's own misery. But one can concede that emotions and emotional expressions are conditioned in various degrees by cultural factors without drawing the conclusion that emotions are merely theoretical constructs. One can concede that our typical understanding of others is contextual without endorsing the view that our engagement with others as minded creatures is primarily a question of attributing hidden mental states to them.

In closing, let me stress that I am obviously not suggesting that our experience and understanding of others are infallible. Other people can fake or conceal their experiences. But there is a decisive difference between our everyday uncertainty about what precisely others might be thinking about, and the nightmare vision of the solipsist. Although we might be uncertain about the specific beliefs and intentions of others, this uncertainty does *not* make us question their very mindedness. To put it differently, our recognition of others as minded creatures is not only more fundamental and certain than our ascription of specific beliefs and emotions to others; but the latter practice is firmly rooted in the former conviction. One can define the mind as something purely interior and private, as something that is not visible in meaningful actions and expressive behavior. But the obvious question is why one should opt for such a narrow mentalistic understanding of the mind in the first place. From a developmental perspective it is certainly not primary. Long before the child starts to wonder about your specific belief-content, it has interacted with and treated you as a social partner (cf. Stern 1985, Rochat 2001).

As I mentioned in the beginning of my paper, one of my aims has been to question whether Goldman offers us a compelling account of our most basic experience of other minds. According to Goldman the “distinctive characteristics of the simulationist approach is to hypothesize that (normal) attributors execute face-based emotion attribution by means that somehow involve the production of that very emotion” (Goldman & Sripada 2005, 199). Indeed, on his account “an attributor arrives at a mental attribution by simulating, replicating or reproducing in his own mind the same state as the target's, or by attempting to do so” (Goldman & Sripada 2005, 194). Not only have I argued that such an account fails to do justice to our experiential access to others – its focus is on the experience of one's own mental states and on the projection and attribution of those states to others, rather than on one's ability to experience the mental states of others – but I have also shown that we in Scheler find an alternative account. One that insists on the existence of a basic and automatic form of empathy that neither involves perspective-shifting, imaginative projection or inferential attribution. The difference between Scheler's and Goldman's account should be obvious. Whereas Goldman argues that a necessary condition for mindreading “is that the state ascribed to the target is ascribed as a result of the attributor's instantiating, undergoing, or experiencing, that very state” (Goldman & Sripada 2005, 208), Scheler explicitly denies this, and claims that such an account conflates empathy with other kinds of interpersonal understanding and fails to capture the fact that we can and do *experience* others.

On my account, empathy provides us with an experiential access to other minds. But to avoid misunderstandings, it is first of all important to realize that empathy rather than being some mysterious form of telepathy simply amounts to an experience of the embodied mind of the other, i.e., simply refers

to our ability to access the life of the mind of others in their bodily and behavioral expressions; an ability that moreover can improve with familiarity, learning, and salience. Secondly, although it is important to recognize the importance of empathy, it is also important to recognize its limitations. There is a limit to how far empathy (plus sensitivity to the immediate context) can get us. Our everyday understanding of others draws on other resources as well. If we wish to unearth *why* somebody is feeling the way he does or *why* he is acting the way he does, we might need to consider the larger social, cultural, and historical context, and this understanding cannot be provided by empathy (cf. Stueber 2006).<sup>9</sup>

## REFERENCES

- Baron-Cohen, S. 2005. Autism – ‘autos’: literally, a total focus on the self? In T. E. Feinberg & J.P. Keenan (Eds.), *The Lost Self: Pathologies of the Brain and Identity*, pp. 166-180. Oxford: Oxford University Press.
- Darwall, S. 1998. Empathy, Sympathy, Care. *Philosophical Studies*, 89, 261-282.
- Gallagher, S. & Zahavi, D. 2008. *The phenomenological mind. An introduction to philosophy of mind and cognitive science*. New York and London: Routledge.
- Goldie, P. 1999. How we think of others' emotions. *Mind and Language*, 14/4, 394-423.
- Goldman, A. & Sripada, C.S. 2005. Simulationist models of face-based emotion recognition. *Cognition*, 94, 193-213.
- Goldman, A. 2006. *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford: Oxford University Press.
- Gordon, R. and Cruz, J. 2003. Simulation theory. *Encyclopedia of Cognitive Science*. London: Nature Publishing.
- Gurwitsch, A. 1979. *Human Encounters in the Social World*. Trans. F. Kersten. Pittsburgh: Duquesne University Press.
- Hodges, S., & Wegner, D. M. 1997. Automatic and controlled empathy. In W. J. Ickes (Ed.), *Empathic Accuracy*, pp. 311-339. New York: Guilford.
- Husserl, E. 1999. *Cartesian Meditations*. Trans. D. Cairns. Dordrecht: Kluwer Academic Publishers.
- Moran, R. 2001. *Authority and Estrangement: An essay on self-knowledge*. Princeton: Princeton University Press
- Nussbaum, M. 2001. *Upheavals of thought: The intelligence of emotions*. Cambridge: Cambridge University Press.
- Overgaard, S. 2007. *Wittgenstein and Other Minds: Rethinking Subjectivity and Intersubjectivity with Wittgenstein, Levinas, and Husserl*. New York and London: Routledge.

---

<sup>9</sup> Thanks to Alvin Goldman, Thor Grünbaum, Joel Krueger and Karsten Stueber for various comments to earlier versions of the paper. The study was funded by the Danish National Research Foundation.

- Rochat, P. 2001. *The infant's world*. Cambridge, MA: Harvard University Press
- Rochat, P. 2008. Mutual recognition as foundation of sociality and social comfort. In T. Striano & V. Reid (Eds.), *Social Cognition: Development, Neuroscience and Autism*. Oxford: Blackwell Publishing. In press.
- Scheler, M. 1954. *The Nature of Sympathy*. Trans. Peter Heath. London: Routledge & Kegan Paul.
- Stein, E. 1989. *On the problem of empathy*. Trans. W. Stein. Washington, D.C.: ICS Publications.
- Stern, D. N. 1985. *The interpersonal world of the infant*. New York: Basic Books.
- Stueber, K. 2006. *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, MA.: MIT Press.
- Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Wittgenstein, L. 1980. *Remarks on the Philosophy of Psychology*, Volume I. Ed. G. E. M. Anscombe and G. H. von Wright, trans. G. E. M. Anscombe. Oxford: Blackwell.
- Zahavi, D. 1996. Husserl's intersubjective transformation of transcendental philosophy. *Journal of the British Society for Phenomenology* 27/3, 228-245.
- Zahavi, D. 2001a. Beyond Empathy. Phenomenological approaches to intersubjectivity. *Journal of Consciousness Studies* 8/5-7, 151-167.
- Zahavi, D. 2001b. *Husserl and transcendental Intersubjectivity*. Athens: Ohio University Press.
- Zahavi, D. 2002. Intersubjectivity in Sartre's *Being and Nothingness*. *Alter* 10, 265-281.
- Zahavi, D. 2007. Expression and empathy. In D.D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-Assessed*, pp. 25-40. Dordrecht: Springer.