Forthcoming in A. Newen, L. de Bruin, S. Gallagher (eds.): *The Oxford Handbook of 4e Cognition*. Oxford University Press. In press.

Beyond Mirroring: 4E Perspectives on Empathy

Dan Zahavi & John Michael

1. Introduction

The notion of empathy does not have a long history. The German term 'Einfühlung' was introduced into the field of social cognition by the psychologist Theodor Lipps at the beginning of the 20th Century and used as a label for our basic understanding of others; an understanding that, according to Lipps, involved a combination of imitation and projection. It was Lipps' notion that Edward Titchener had in mind when he in 1909 translated 'Einfühlung' as 'empathy' (Titchener 1909).

When considering the current debate on empathy, it quickly becomes evident that a diversity of different definitions of and approaches to the topic are available, and that no consensus seems forthcoming. A recent issue of *Boston Review* entitled *(Against) Empathy* can serve as a good illustration of this.

In his target contribution, Paul Bloom concedes that the term 'empathy' is used in many ways, but maintains that it typically refers to a process whereby one comes to experience the world as others do, be it through imaginative perspective taking or by some kind of affective matching. Bloom further argues that empathy serves to dissolve the boundaries between one person and another, and that it can therefore be a force against selfishness and indifference. It is consequently not surprising that some have seen empathy as a moral virtue and have argued that we need to nurture and expand our empathic powers, since a high degree of empathy might be a requirement for being good and doing good (Bloom 2014: 15).

For a variety of reasons, however, Bloom is quite sceptical about this line of reasoning. This is in part because empathy concerns our relation to specific individuals. If we are striving for a better world, one that might involve an increase of humanitarian aid or a rethinking of the criminal justice system, i.e., policies affecting large groups of people, a keen sense of justice or moral obligation might be far more relevant than any empathic skill. This is all the more true since, according to Bloom, empathy is biased: we tend to empathize more with those whose needs are salient, who are similar to ourselves, and who are close by. If we want to promote impartiality and fairness, we should consequently put empathy aside.

In addition, Bloom is also sceptical about the value of empathy even in relationships with specific other people. To empathize with another person in pain or distress is, according to Bloom, to feel what the other person is feeling. But if the empathizer suffers as a result of empathizing with your suffering, it is not obvious that this is to your advantage. Empathic distress can lead to egoistic drift, where the empathizer becomes more concerned with alleviating her own distress (for instance by absenting himself) than with caring about you. You want the other to respond with care and concern and help, rather than to relive your pain and distress. You want the physician to be calm and confident when she is treating you, not to be overwhelmed by negative emotions (Bloom 2014: 16).

Given the definition of empathy that Bloom starts out with, this line of reasoning might seem compelling.[1] But, as pointed out by various commentators in the *Boston Review* issue, it is by

[1] For a more extensive discussion of the relation between empathy and morality cf. Maibom 2014

no means obvious that Bloom's definition is the most appropriate one, i.e. that we should conceptualize empathy as involving an affective matching between empathizer and target. Moreover, how can I expect to receive care or help from another unless she understands my situation, and isn't that understanding exactly what empathy on some accounts is supposed to provide? Should we then distinguish among different kinds of empathy – say, an affective type and a cognitive type? If so, what is the relation between cognitive empathy and ordinary mindreading? Questions abound. People disagree about the role of affective matching, caring, understanding and imagination in empathy, just as they disagree about the relation between empathy and social cognition in general, and about whether empathy is a natural kind or rather a heterogeneous construct (for an overview, cf. Zahavi 2014a).

In the following, we will not attempt to resolve these disputes or to argue in favour of any one particular way of conceptualizing empathy. Instead, our aim is to open up a new perspective by exploring the potential of applying embodied, extended, enactive and embedded approaches to empathy research. As we shall see, these approaches provide useful resources in thinking about empathy, and in particular in going beyond the notion of affective matching. First, however, it will be useful to begin by articulating the notion of affective matching as clearly as possible, and by illustrating how it, too, can be enriched and sharpened by drawing upon ideas from embodied cognition approaches.[2]

2. Empathy and affective matching

2.1. Lipps

It is natural to start by taking a closer look at Lipps' contribution. Although the German term for empathy (*Einfühlung*) was first used in the domain of aesthetics, it was Lipps who started to use it in the context of interpersonal understanding. Whereas a number of contemporary empathy theorists have taken empathy to denote a particular pro-social attitude vis-à-vis others, the original discussion was to a much larger extent epistemologically oriented. In various of his writings, Lipps argues that there are three distinct domains of knowledge: 1) knowledge of external objects, 2) self-knowledge, and 3) knowledge of others – and he took these domains to have three distinct cognitive sources, namely perception, introspection and empathy (Lipps 1909: 222). Lipps consequently insisted that empathy, which he took to be a psychological and sociological core-concept, qualified as a modality of knowledge *sui generis* (Lipps 1907: 697-698, 710).

Sometimes Lipps refers to what he calls the *instinct of empathy*, and argues that it involves two components, a drive directed towards imitation and a drive directed towards expression (Lipps 1907: 713). In the past, I have been sad, and have experienced an instinctual tendency to express that sadness. The expression was not experienced as something next to or on top of the sadness but as an integral part of the feeling. Now, when I see the expression elsewhere, I have an instinctual tendency to imitate or reproduce it, and this tendency will then evoke the same feeling to which it was intimately connected in the past (Lipps 1909: 229-230, 1907: 719). When I experience the feeling anew, it will remain linked to the expression I am currently perceiving and will be projected into or onto it. In short, when I see an angry face, I will reproduce the expression of anger, this will evoke a feeling of anger in me, and this felt anger that is co-given with the currently perceived facial expression will then be attributed to the other, thereby allowing for a form of interpersonal understanding (Lipps 1907: 717-719).

[2] The following discussion partly draws on and expands on points made in Zahavi 2011, 2014a and 2014b, Michael & Fardo 2014, Michael 2014.

One implication of Lipps' model is that there are rather strict limitations to what I can come to understand empathically of the other. The imitated expression can only evoke an affective state in myself that resembles the affective state of the other if I have had the affective state in question in the past (Lipps 1907: 718-719). Consequently, I can only understand those of the other's experiences which I have already enjoyed myself, or to put it differently, Lipps' account of empathy doesn't allow me to recognize anything in the other that is new, anything that I am not already familiar with, anything that I haven't put there myself.

2.2 Goldman, De Vignemont and Colleagues

Lipps' position has remained influential and has a number of modern heirs. Not surprisingly, it is in particular within the simulationist camp that the notion of empathy has resurfaced as a central category. Indeed, it has even been argued that simulationists are today's equivalents of empathy theorists (Stueber 2006: ix). Goldman has acknowledged that simulationist themes can be found in earlier theorists such as Lipps (Goldman 2006: 18), and in *Simulating Minds,* Goldman explicitly equates *empathy theory* with *simulation theory* (Goldman 2006: 11), and states that mindreading is an extended form of empathy (Goldman 2006: 4).

In recent work, Goldman has emphasized that an account of mindreading should be able to cover the whole range of mental states, including sensations, feelings and emotions – i.e. it shouldn't just address the issue of belief-ascription (Goldman 2006: 20). This is precisely why Goldman now distinguishes what he calls *low-level mindreading* from *high-level mindreading* (Goldman 2006: 43), and argues that we need to recognize the existence of a simple, primitive and automatic ability to attribute basic emotions such as fear, anger and disgust to others on the basis of their facial expressions (Goldman & Sripada 2005).

How can we explain this kind of basic 'mindreading,' this ability to recognize someone's face as expressive of a certain emotion? In *Simulating Minds,* Goldman considers different models, and ultimately opts for one he calls the *unmediated resonance model* (Goldman 2006: 132) that avoids some of the limitations of Lipps' original proposal.

According to this model, the same neural substrate is activated both when we experience an emotion ourselves and when we recognize the emotion in others. This occurs because the perception of a target's emotional expression directly triggers activation of the neural substrate of the same type of emotion in oneself, thereby making the process a kind of unmediated matching, one that bypasses the need for and feedback from facial mimicry (Goldman 2006: 128). When compared to Lipps' model, this proposal has the distinct advantage that it does not require an agent to have had any particular past experience in order to empathize with the experience that some other agent is currently having. Rather, the coupling is hardwired. In principle, observing the facial expressions of others might give rise to new emotions in yourself, emotions you haven't felt before.

Goldman suggests that my 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 toward the use of one's disgust experience as the causal basis for third-person disgust attributions" (Goldman 2006: 137). It is consequently no coincidence that Goldman considers a more apt name for the whole process to be simulation-plus-projection (Goldman 2006: 40), thereby affirming the structural similarity between his own account and the one we found in Lipps.

Like Goldman, de Vignemont and colleagues (De Vignemont and Jacob 2012; cf. also De Vignemont and Singer 2006) have articulated a conception of empathy that also builds upon the core idea found in Lipps. In contrast to Goldman, however, de Vignemont and her co-authors have

insisted on the need for a quite narrow definition of empathy. Rather than seeing empathy simply as another label for (basic) mindreading, they have stipulated necessary and sufficient conditions for empathy which, taken together, make it possible to distinguish empathy from related phenomena such as emotional contagion, sympathy and mindreading[3]. In this vein, De Vignemont and Jacob (2012) have recently proposed the following conditions:

(i) the affectivity condition;
(ii) the interpersonal similarity condition;
(iii) the causal path condition;
(iv) the ascription condition;
(v) the caring condition

Let us briefly examine each of these in turn. The affectivity condition – i.e. condition (i) – demands that an empathizer and her target both experience an affective state. This condition rules out the possibility of empathizing with someone who is not experiencing an affective state of any kind. It also makes it possible to distinguish empathy from standard mindreading: merely identifying that somebody is in a particular affective state would be mindreading, whereas empathy involves the empathizer also coming to experience an affective state.

The interpersonal similarity condition – i.e. condition (ii) – demands that the affective states of the empathizer and the target be similar to each other (De Vignemont and Singer 2006 refer to an 'isomorphism' between the affective states of the empathizer and her target). This is intended to make it possible to distinguish empathy from sympathy. A sympathizer also experiences an affective state, but it is different from the affective state of the target. In fact, according to De Vignemont and Jacob, sympathy is a 'sui generis social emotion'.

The causal path condition – i.e. condition (iii) – requires that the empathizer's affective state is caused by the target person's affective state. This is intended to rule out cases in which two people have similar affective experiences because of some common cause. The ascription condition – condition (iv) – requires that the empathizer be aware that the target person is in an affective state and that this is the source of her own affective state. This distinguishes empathy from contagion. The caring condition – condition (v) – is based on the observation that when empathizing (as when sympathizing) one tends to be concerned about the target person's well-being, i.e. insofar as she is suffering one is motivated to alleviate that suffering. It is worth noting that this condition is also assumed in psychological theories in which empathy is associated with prosocial motivation (Batson 1991; Hoffman 1982).

Taken together, these five conditions present a clear and concise conception that systematically relates empathy to mindreading, sympathy and contagion (for a critical discussion, cf. Zahavi 2011, Zahavi & Overgaard 2012). It is also worth emphasizing that De Vignemont and Jacob (2012) have developed this conception in a manner that incorporates impulses from recent embodied cognition approaches. To see how, consider their analysis of one particular type of empathy, namely empathy for pain. Their account is based on the common view that the experience of pain derives from the processing and integration of nociceptive inputs and complex emotional and cognitive processes, implicating the participation of several pain-specific brain structures that may be functionally distinct. The neural network involved in pain processing is often referred to as the 'pain matrix', the primary components of which are sometimes said to be a sensory-discriminative and an

[3] There is broad agreement that a conceptual model of empathy should make it possible to distinguish empathy from these related phenomena. In future research, it may also be useful to investigate how empathy relates to pity, compassion or other associated phenomena.

affective-motivational network (e.g., Singer et al. 2004; Aydede 2006). On this view, primary and secondary somatosensory and posterior insular cortices are thought to serve the processing of sensory-discriminative features of pain stimuli, such as location, duration, and stimulus intensity. In the affective-motivational domain, anterior cingulate and anterior insular cortices are thought to mediate these aspects of pain processing, for example, the unpleasantness of pain.

De Vignemont and Jacob (2012) suggest that these two components can be dissociated and that this provides a basis for distinguishing pain empathy from the phenomenon of contagious pain. Specifically, they suggest that contagious pain is more likely to recruit the sensory-discriminative component, whereas empathy is more likely to recruit the affective component of the pain matrix. In accordance, they argue that empathy is 'other-centered' insofar as it involves a concern for the other person's affective state, whereas contagion is 'self-centered'. In support of this position, they refer to research suggesting that in pain empathy the affective neural components are selectively activated (Singer et al. 2004; Botvinick et al. 2005). Thus, De Vignemont and colleagues' proposal is an embodied account insofar as it conceptualizes affective matching as a bodily response to the others' experiences, and insofar as it distinguishes empathy from related phenomena partly in virtue of the specific nature of this bodily response (for critical discussion, see Michael and Fardo 2014; Michael 2014).

For De Vignemont and colleagues, empathy is a special kind of third-person mindreading (de Vignemont and Jacob 2012: 310), which is more complex and less direct than standard mindreading (de Vignemont 2010: 292; de Vignemont & Singer 2006: 439). After all, whereas empathy must meet five requirements, the simpler and presumably more widespread standard mindreading only has to meet one requirement – that of attributing a mental state to another (de Vignemont & Jacob 2012: 307). Thus, and this is obviously quite significant, on their proposal, empathy is not what establishes an awareness of the other person's mental state in the first place. Rather, empathy requires a prior understanding of the other's mental life in order to get off the ground, and is then supposed to allow for an enhanced understanding of the other's feeling.

As mentioned before, one noteworthy difference between Goldman and de Vignemont is consequently that whereas the former takes (basic) empathy to be involved in simple mindreading, the latter operates with a far more restrictive use of the term. Despite this difference, though, it bears emphasizing that both Goldman and de Vignemont understand affective matching to be integrated with various other social cognitive processes, from drawing inferences about others' situations and mental states, to being motivated to alleviate others' suffering. This is important in connection with the concerns raised by Bloom about empathy (i.e. that affective matching may lead to biases and may be of little use to a person who is suffering). While Bloom may well be right that affective matching, taken alone, may be prone to such dangers and limitations, empathy may not be because it typically involves a consideration of others' situations and mental states as well as a motivation to alleviate their suffering.

While Goldman and de Vignemont illustrate how approaches to empathy may include affective matching but also incorporate other resources that help to address Bloom's concerns, we shall see in the next section that some 4E approaches, also inspired in important ways by themes within the phenomenological tradition, depart more radically from the debate which structures Bloom's critique of empathy.

3. Insights from 4E approaches

3.1 Embodied Simulation: Gallese and Iacoboni

Lipps' influence on the contemporary empathy debate is wide-reaching. Other authors inspired by

him include Iacoboni and Gallese, who both endorse Lipps' idea that empathy involves a form of inner imitation (Gallese 2003: 519, Iacoboni 2007: 314). However, both of the latter have also been inspired and influenced by the post-Lippsian discussion of empathy found in the phenomenological tradition. Gallese, for instance, references Stein's account of empathy, and Husserl's and Merleau-Ponty's discussion of intersubjectivity and intercorporeity (2001), and is quite explicit in arguing that his own notion of embodied simulation is akin to, and a further development of, the phenomenological proposal (Gallese et al. 2004: 397; cf. Iacoboni 2009).

Building upon similar findings to those discussed by Goldman and De Vignemont, embodied simulationists such as Gallese and Iacoboni draw more far-reaching conclusions. For them, empathy is not a relatively rare instance of an enhanced understanding of others' affective states. Rather, for them empathy constitutes a basic and important form of social understanding which renders further inferences or explicit attribution of mental states to others otiose (Gallese 2001, 2009, Gallese et al 2004). More specifically, the discovery of what have been called *mirror neuron systems* or *neuronal resonance mechanisms* has been interpreted as lending support to the existence of a low-level simulation-based form of empathy; one that explains the ease with which we 'mirror' ourselves in the behaviour of others and recognize them as similar to us. Indeed, it is the neural matching mechanism constituted by mirror neurons that allows for a direct, automatic, non-predicative, and non-inferential empathic link between different individuals (Gallese 2001: 42, 44).

Gallese ultimately claims that all kinds of interpersonal relations – including action understanding, the attribution of intentions, and the recognition of emotions and sensations in others – rely on automatic and unconscious embodied simulation routines (Gallese 2003: 517). The very same neural substrate, which is activated when we execute actions or subjectively experience emotions and sensations, is also activated when we observe somebody else act or experience emotions and sensations. So, when we encounter somebody, and observe their actions, or their displayed emotions or sensations, we don't just see them. In addition to the sensory information we receive from the other, internal representations of the body states associated with the other's actions, emotions and sensations are evoked in us, and it is "as if" we are doing a similar action or experiencing a similar emotion or sensation. It is because of this automatic, non-predicative and non-inferential embodied simulation mechanism, it is because the activation of these neural mechanisms allows us to share actions, intentions, feelings and emotions with others, that we are able to understand others (Gallese 2001: 44-45, 2009: 527).

We shall here not rehearse in any detail the debate concerning the precise contribution of the mirror neurons (see Michael 2011a). For our purpose, the more interesting aspect concerns the extent to which empathy on this model is taken to involve a rupture with standard accounts of social cognition, and ultimately point beyond the dichotomy of behaviour-reading and mindreading (cf. Sinigaglia 2008). According to Iacoboni, mirror neuron activity links self and other in a way that questions traditional Cartesian as well as more recent cognitivist assumptions about how social understanding comes about. Indeed, as Iacoboni also writes, the functioning of mirror neurons only makes sense if we are dealing with agents that interact with other people in a shared environment, where the classical dichotomies (such as action-perception, subject-world or inner-outer) have dissolved (Iacoboni 2007, 2009). This view is, according to Iacoboni, reminiscent of themes found in existential phenomenology, which is why he has labelled his own approach, 'existential neuroscience' or 'neurophysiologic phenomenology' (Iacoboni 2007: 319, 2009: 17). Indeed, for Iacoboni the discovery of mirror neurons has not only for the first time in history provided a plausible neurophysiological explanation for complex forms of social cognition and interaction (Iacoboni 2009: 5). Mirror neurons also seem to explain why, as he puts it, "existential phenomenologists were correct all along" (Iacoboni 2009: 262).

On this account, embodied social cognition might involve an attempt to replicate, imitate or simulate the mental life of the other, but the simulation process in question is automatic, unconscious, prelinguistic, and non-metarepresentational. As Gallese puts it, intercorporeity is more fundamental than any explicit attribution of propositional attitudes to others and remains the main source of knowledge we directly gather about others (Gallese 2009: 524). In other words, while De Vignemont and colleagues conceptualize embodied affective sharing as a component part of empathy, Gallese goes a step further in drawing the conclusion that embodied simulation may suffice for empathic understanding and thus render the attribution of mental states superfluous (i.e. he rejects De Vignemont and colleagues' ascription condition).

Gallese's work on embodied simulation is undoubtedly more in line with 4E approaches to cognition than Goldman's or de Vignemont's. However, this is not to say that there are no tensions within his conception of empathy. As we have seen, Gallese has emphasized the affinities between his own position and that of both Lipps and the classical phenomenologists. This is in itself slightly surprising, since the latter were highly critical of Lipps' account (cf. Zahavi 2014a). Gallese's commitment to the idea that empathy is at bottom to be explained in terms of mirroring or matching mechanisms is also not unequivocal. In a 2009 publication, Gallese observes that the mirror metaphor itself might be misleading, since it suggests the presence of an exact match between object and observer, thereby disregarding individual differences (Gallese 2009: 531), and he has also conceded that imitation cannot really account for interpersonal understanding, since the latter calls for a preservation of difference and otherness (Gallese 2007: 11; 2009: 527).

The idea that empathy, rather than involving identity, similarity and affective matching, might crucially preserve interpersonal difference, and indeed highlight other-centeredness, is something that other phenomenologically inclined empathy theorists have explored further. To see how, let us briefly return to the phenomenological reception of Lipps.

3.2 Other-centeredness

In the wake of Lipps' investigation, a number of phenomenologists engaged in intensive discussions regarding the nature and structure of empathy. While they accepted the idea that empathy must be equated with (a basic form of) other-understanding, they were more critical of Lipps' suggestion that empathy involves a form of inner imitation, and rejected various attempts to explain empathy in terms of mirroring or mimicry.

In a number of recent publications, one of us has offered a systematic reconstruction of the phenomenological discussion of empathy (drawing especially on Husserl, Stein and Scheler), and has defined empathy as "a distinctive form of other-directed intentionality, distinct from both self-awareness and ordinary object-intentionality, which allows foreign experiences to disclose themselves as foreign rather than as own" (Zahavi 2014b: 138). On this account, empathy is a perceptually based experience of another person's mental life, one that more complex and indirect forms of social cognition presuppose as well as rely on. To insist that the empathizer must have the same (kind of) state as the target, is on this account to miss what is distinctive about empathy, namely the fact that it confronts you with the presence of an experience that you are not living through yourself. Rather than blurring the distinction between self and other, rather than leading to some sense of merged personal identities (Cialdini et al 1997), the asymmetry between self-experience and other-experience is consequently crucial for empathy. One might say that empathy provides a special kind of knowledge by *acquaintance*. It is not first-person acquaintance, but rather a distinct other-acquaintance. Empathy denotes a special kind of epistemic access and should not be confounded with sympathy or compassion. Thus, for Zahavi, it is perfectly coherent to think that an expert torturer may rely on empathy in order to work out how best to push her victim's buttons (Zahavi 2011, 2014a,

2014b)

One benefit of this lean account is that it can elegantly distinguish between empathy and emotional contagion – empathy, according to Zahavi, does not require that the empathizer have an affective experience that is similar to that of the target. Moreover, this way of conceptualizing empathy also admits of a clear and straightforward distinction between empathy and sympathy: sympathy, but not empathy, involves concern for the well-being of the target person.

However, a challenge for this view is to articulate a distinction between mindreading and empathy. One option would be to regard empathy as a *kind* of mindreading – for example, as perceptual mindreading. It is no coincidence that Zahavi has occasionally presented his investigation of empathy in the framework of the direct social perception debate (cf. Zahavi 2011). For some, it may seem odd to restrict empathy to perception. After all, why should we exclude cases in which one learns about someone's suffering through a third person and empathizes with them without ever perceiving them? But, as we have already noted, people's intuitions about borderline cases of empathy tend to diverge greatly, so one should be highly cautious about basing theoretical claims on them. Our evaluation of such a restriction would therefore have to be based on other factors, such as the overall coherence of the proposal, its empirical fruitfulness or its explanatory power. For the moment, then, it must be regarded as perfectly legitimate to advance a conception of empathy that very much targets the face-to-face encounter. The guiding idea would be the following: Just as we ought to consider the difference between thinking about a tiger, imagining a tiger, and seeing a tiger, we also ought to acknowledge the difference between referring to Berta's compassion or sadness, imagining in detail what it must be like for her to be compassionate or sad, and being empathically acquainted with her compassion or sadness in the direct face-to-face encounter. In the latter case, our acquaintance with Berta's experiential life has a directness and immediacy to it that is not possessed by whatever beliefs we might have about her in her absence.

But this is not the only option for Zahavi. Another option would be to say that empathy is more basic and fundamental than mindreading proper. The coherency of this proposal obviously depends on what one understands by mindreading. For some, the term 'mindreading' suggests that we come to identify mental states on the basis of bodily behaviour in a manner analogous to the way in which we grasp meaning on the basis of written inscriptions (cf. Heyes & Frith 2014; Apperly 2011). According to this usage, mindreading qua mental state attribution is a skill that has to be acquired just as we need to learn how to read texts (since there is no intrinsic or natural connection between the psychologically meaningful mental states and what is perceptually available). Given such a usage, empathy could be seen as an immediate and direct form of social understanding (involving sensitivity to the animacy, agency and emotional expressivity of others) that is manifest from the outset and which any attempt to explain or predict the other's mental states and behaviours rely on and presuppose.

Even if children from birth onwards might have the empathic ability to distinguish animate creatures from inanimate objects, the introduction of a developmental perspective on empathy complicates matters somewhat. It might, for instance, put pressure on an overly epistemic account of empathy by suggesting that the most basic form of social relatedness isn't emotionally neutral. As Hobson has observed, interpersonal understanding normally involves emotional responsivity and this especially holds true for early infancy, where it is the infant's affective engagement with others that provides it with salient interpersonal experiences encompassing an interplay between similarity and difference, connectedness and differentiation (Hobson 2007; cf. Reddy 2008)? Indeed, such emotionally structured interactions in early infancy might be crucial for the further development of social cognition. A more in-depth exploration of the relationship between interpersonal affectivity and social cognition is obviously of crucial importance, but beyond the scope of this paper.

At any rate, Zahavi's conception is clearly leaner than that espoused by De Vignemont and colleagues. In order to see just how much leaner it is, let us examine which of De Vignemont and colleagues' conditions it incorporates and which it does not. Like De Vignemont and colleagues, Zahavi requires that there be a causal link between the mental state of the empathizer and that of the target person. Thus, he endorses the causal path condition – i.e. condition (iii). In fact, since empathy in his sense is elicited by perception, it appears to exclude some other types of causal path that would satisfy De Vignemont and colleagues' condition (iii), e.g. imagination, memory or communication. In this sense, Zahavi's version of condition (iii) is stricter than De Vignemont and colleagues'. As for De Vignemont and colleagues' requirement that the empathizer ascribes a mental state to the target person (i.e. condition (iv)), Zahavi certainly does maintain that empathy is other-centered or other-directed, but he also allows for cases of empathy where it doesn't involve any specific mental state ascription, but simply an experience of the presence of other-mindedness. Thus, Zahavi endorses the ascription condition (condition (iv)) in a loose sense. However, Zahavi dispenses with the affectivity condition (condition (i)), since he rejects the claim that one can only empathize with affective states. We can see the other's elation or doubt, surprise or attentiveness in his or her face, we can hear the other's trepidation, impatience or bewilderment in her voice, feel the other's enthusiasm in his handshake, grasp his mood in his posture, and see her determination and persistence in her actions. Moreover, Zahavi also rejects the interpersonal similarity condition (ii). In fact, Zahavi is at pains to insist that empathy primarily confronts one with the presence of an experience that one is not living through oneself. Finally, as already noted, empathy in Zahavi's sense does not entail any prosocial motivation. Thus, Zahavi also rejects the requirement that the empathizer be concerned for the well-being of the target person (i.e. condition (v), the caring condition).

3.3 Intentional Alignment

Building upon the theme of other-centeredness, some theorists (within and outside of the phenomenological tradition) do make room for matching of a sort to play an important role in empathy – but it is of a very different sort than that envisaged by proponents of the Lippsian model of projective empathy.

To begin with, Merleau-Ponty maintains that when we perceive an angry gesture, we are perceiving the anger itself and not merely psychologically meaningless behaviour. At the same time, though, he denies that the meaning of the gesture is perceived in the same way as the colour of the carpet. The other's gestures point to an intentional object, and I understand the meaning of those gestures, not by looking behind them, but by attending to the part of the world that they highlight (Merleau-Ponty 2012: 191-192). The idea here is that in perceiving another as a minded being whose experience is directed toward the same world as that toward which my experience is directed, the intentional content or structure of their experience may also become a focus of my experience.

Shaun Gallagher has recently put forth a proposal based upon this idea, which explicitly incorporates elements of enactive and embedded cognition approaches. His starting point is to highlight that emotions at least usually involve not only a qualitative feel but also an intentional object or structure. Thus, while it is not at all necessary for an empathizer to enter into an affective state that matches that of her target with respect to its qualitative feel, it is necessary for her to attune to the same intentional object or state of affairs that is the focus of the target's experience. This, according to Gallagher, is what distinguishes empathy from sympathy:

> 'Empathy: A feels sad [and/or outrage] *about the injustice done to B*, knowing that B also feels sad [and perhaps outrage] about the injustice done to her (A's feeling has a similar intentional structure as B's affective state) (2012: 6).

In contrast, A may also feel sad for B without agreeing with B that an injustice really occurred:

> Sympathy: A feels sad *for B*, who is sad [and perhaps outraged] about an injustice done to B (dissimilar intentional structure)' (2012: 6).

With this distinction in hand, Gallagher is able to distinguish empathy from contagion, since contagion clearly does involve matching with respect to affect states but not with respect to the intentional structures of experiences. Thus, Gallagher's account includes a kind of alignment between empathizer and target, but this alignment is specified in relation to intentional content rather than affective or qualitative experience. One possible objection to this view is that it appears to rule out the possibility of empathizing with someone when one does not know what the focus of their experience is. And yet, if one for example enters a room and finds someone sitting there weeping, it seems odd to suggest that one could not empathize with her. In response to this objection, one reasonable response for Gallagher would be to place the emphasis on the empathizer's motivation to identify the focus of the target person's experience rather than on the successful identification per se. In other words, empathy may be understood to involve the project of identifying and understanding a target's experience (see Goldie 1999 for an account which is not inspired by phenomenology but that also conceptualizes empathy as a project of reconstructing the other's experience without matching its internal features).

In this context, it is interesting to take a second look at the body of research into pain empathy which, as mentioned above, has informed the approach of De Vignemont and colleagues. Specifically, some of this research suggests that the areas associated with the affective component of the pain matrix (anterior cingulate cortex and anterior insula) are in fact involved in modulating attention to reflect what is salient in the environment (Iannetti and Mouraux 2010; Legrain et al. 2011; Mouraux et al. 2011). In other words, whenever there is something salient and attention needs to be directed to it, these areas are recruited, and of course pain is one example of a salient stimulus, but so are lots of other non-painful things. Crucially, registering salience would fit well with the idea of aligning the intentional structure of one's experience with that of the person with whom one is empathizing. Hence, this data raises the possibility that the findings which De Vignemont and colleagues take to provide evidence of embodied affective sharing in fact reveal the matching of intentional structures of experiences rather than of qualitative features of those experiences.

3.4 Complementarity and Reciprocity: The importance of the second-person perspective

In developing the core 4E themes of embodied, enactive, extended and embedded cognition, some recent theorists have also homed in on the overlapping themes of complementarity and reciprocity in empathy – both of which themes are neglected by approaches focusing on affective sharing. Here, again, the writings of Merleau-Ponty have proven to be a fruitful starting point.

In characterizing the internal relation that obtains between my own body and that of the other, for example, Merleau-Ponty claims that the other appears as the completion of the system and that "the other's body and my own are a single whole, two sides of a single phenomenon" (Merleau-Ponty 2012: 370). Thus, to speak, as Merleau-Ponty does, of self and other as "collaborators in perfect reciprocity" (2012: 370), suggests an approach to social cognition where the encounter with the other's actions, rather than simply occasioning a mere replication or simulation of those actions, elicits a dynamic response that takes those actions as affordances for further complementary actions (cf. Gallagher and Miyahara 2012). In order to capture what Merleau-Ponty has in mind, it might consequently be better to liken social understanding to dancing than to mirroring. In any case,

Merleau-Ponty's emphasis on complementarity provides one more reason why proponents of the embodied simulation approach should distance themselves more clearly from the Lippsian model of projective empathy if they wish to retain a link to the classical phenomenological account of empathy.

In the current research landscape, the themes of complementarity and reciprocity provide useful keys to appreciating ongoing debates about the importance of the second-person perspective. Over the past 15 years or so, interest in the second-person perspective has been stoked by dissatisfaction with the two hitherto dominant mainstream positions in the theory of mind debate, the theory theory (in its different versions) and the simulation theory (in its different versions). It has occasionally been argued that a limitation of both of the traditional positions is that they privilege either the first-person perspective (this would be the simulation theory) or the third-person perspective (this would be the theory theory), and that an adequate account of social cognition should also explicitly target the second-person perspective.

However, there is still considerable disagreement about what exactly second-person perspective taking involves. One influential account can be found in a target article in *Behavioral and Brain Sciences*, written by Schilbach and colleagues. For them, the second-person perspective concerns the issue of directly interacting with and emotionally engaging with others (rather than simply observing them from a distance). Thus, the second-person perspective is contrasted with what is called the spectatorial stance (Schilbach et al., 2013). Indeed, given that face-to-face interaction engages complementary affective, motoric and higher-cognitive processes that are not engaged in observational settings, an important new challenge for researchers investigating social cognition is to consider more developing interactive paradigms (Michael, 2011b; Overgaard and Michael 2015).

One aspect that may not have been sufficiently highlighted in Schilbach and colleagues' article, however, is the role of reciprocity (de Bruin *et al.*, 2012; Fuchs, 2013). After all, the second-person perspective involves bidirectionality and reciprocation. In short, to adopt the second-person perspective is to engage in a subject–subject (you–me) relation where I not only respond to the other but am aware of the other as an other, and, at the same time, implicitly aware of myself in the accusative, as attended to or addressed by the other (Husserl, 1973a, 211). This process has been described in much detail by developmental psychologists working on dyadic joint attention (Rochat 2001, Reddy 2008), but also by classical phenomenologists such as Walther, Stein and Schutz who argued that reciprocal empathy is a key to experiential sharing and communal experiences (cf. Zahavi & Rochat 2015, León & Zahavi 2016, Zahavi & Salice 2016).

4. Conclusion

In sum, then, we have seen that impulses from 4E approaches cognition have informed recent attempts to articulate the notion of affective matching and to use it as a basis for conceptualizing empathy and distinguishing it from related phenomena, and that they have also inspired approaches to empathy that do not require affective matching at all. We have also seen that all of these approaches offer resources that help to address the concerns raised by Bloom in his target article. While Bloom is surely right that merely matching one's affective state with that of someone who is suffering is not necessarily much use to them or to anyone else, empathy need not be limited to such a matching relation, and in fact it need not involve such a matching relation at all.

For those who do include affective matching in their conception of empathy, it is crucial to bear in mind that empathy is integrated with various other social cognitive processes, from drawing inferences about others' situations and mental states, to being motivated to alleviate others' suffering. Furthermore, as we have already seen, there are many who simply do not consider affective matching to be a constitutive feature of empathy at all. For them, empathy may in fact include a sort of matching insofar as the empathizer comes to focus on the intentional content or structure of the target's

experience – and indeed this is just the sort of contextual feature that embedded and extended approaches would urge us to take into account. But it is difficult to see how this sort of matching would fall prey to the same dangers as the affective matching criticized by Bloom. Moreover, whatever one might think of this sort of matching as intentional alignment, the themes of other-centeredness, complementarity and reciprocity – all of which are inspired by insights from embodied and enactive approaches – provide further insurance against the perils identified by Bloom. In sum, then, the outlook for empathy is not as bleak as Bloom envisions.

Of course, this still leaves us with the question of how to understand empathy. In the contemporary debate, one can encounter distinctions between mirror empathy, motor empathy, affective empathy, perceptually mediated empathy, reenactive empathy and cognitive empathy, to mention just a few of the options available. As should have become clear by now, one reason why it continues to be so difficult to reach a commonly accepted definition of empathy is that people have been using the notion to designate rather different phenomena. For the same reason, it is not obvious that it makes that much sense to try to determine once and for all what empathy really is. Although one might make the case that one ought to stick to the traditional use of the term – as already mentioned, it was introduced by Lipps as a general term for our understanding of others – instead of identifying it with, say, prosocial behaviour or a very special kind of imaginative perspective taking, it is not evident that such a strategy would be particular productive or illuminating. Thus, rather than promoting a specific account of empathy as the right account, a more reasonable verdict might be that the different analyses of empathy contain various insights that contemporary debates on social cognition and interpersonal understanding ought to incorporate.

References

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of 'Theory of* Mind.' Hove: Psychology Press.

Aydede M (ed) (2006). *Pain: new essays on its nature and the methodology of its study*. MIT Press, Cambridge, MA

Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, *2*(2), 107-122.

Bloom, P. (2014). Against Empathy. *Boston Review* 39/5: 14-19.

Botvinick M, Jhab A, Bylsmaa L, Fabian S, Solomon P, Prkachin K (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage* 251:312–319

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., & Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology*, 73(3), 481–494.

De Bruin, L., van Elk, M., and Newen, A. (2012). Reconceptualizing Second-Person Interaction. *Frontiers in Human Neuroscience*, 6: 1–14.

de Vignemont, F., Jacob. P. (2012). What is it like to feel another's pain? *Philosophy of Science* 79(2):295–316

de Vignemont, F., Singer, T. (2006). The empathic brain: how, when and why? *Trends in Cognitive Sciences* 10:435–441

De Vignemont, F. (2010). Knowing Other People's Mental States as if They Were One's Own. In S. Gallagher & D. Schmicking (Eds.), *Handbook of Phenomenology and Cognitive Science* (pp. 283–299). Dordrecht: Springer.

Fuchs, T. (2013). The Phenomenology and Development of Social Perspectives. *Phenomenology and the Cognitive Sciences*, 12/4: 655–83.

Gallagher, S. (2012). Empathy, simulation, and narrative. *Science in context*, 25/3: 355-381.

Gallagher, S., and Miyahara, K. (2012). Neo-Pragmatism and Enactive Intentionality. In J. Schulkin (ed.), *Action, Perception and the Brain: Adaptation and Cephalic Expression* (pp. 117-146). Basingstoke: Palgrave Macmillan.

Gallese, V. (2001). The "Shared Manifold" Hypothesis: From Mirror Neurons to Empathy. *Journal of Consciousness Studies*, 8/5–6: 33–50

Gallese, V. (2003). The Manifold Nature of Interpersonal Relations: The Quest for a Common Mechanism. *Philosophical Transactions: Biological Sciences*, 358/1431: 517–28.

Gallese, V. (2007). Embodied Simulation: From Mirror Neuron Systems to Interpersonal Relations. *Empathy and Fairness: Novartis Foundation Symposium*, 278: 3–19.

Gallese, V. (2009). Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification. *Psychoanalytic Dialogues*, 19/5: 519–36.

Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A Unifying View of the Basis of Social Cognition. *Trends in Cognitive Sciences*, 8/9: 396–403.

Goldie, P. (1999). How we think of others' emotions. *Mind and Language* 14(4):394–423

Goldman, A. I. (2006). *Simulating Minds.* New York: Oxford University Press.

Goldman, A. I., and Sripada, C. S. (2005). Simulationist Models of Face-Based Emotion Recognition. *Cognition*, 94/3: 193–213.

Heyes, C., Frith, C. (2014). The Cultural Evolution of Mind Reading. *Science* 344/6190. DOI: 10.1126/science.1243091

Hobson, R.P. (2007). Communicative depth: Soundings from developmental psychopathology. *Infant Behavior & Development*, 30, 267–277.

Hoffman, M.L. (1982). Development of prosocial motivation: empathy and guilt. In: Eisenberg N (ed): *The development of prosocial behavior* (pp.281-338). Academic Press, New York, NY.

Husserl, E. (1973). *Zur Phänomenologie der Intersubjektivität II. Texte aus dem Nachlass. Zweiter Teil. 1921-28*. Ed. I. Kern. Husserliana 14. Den Haag: Martinus Nijhoff.

Iacoboni, M. (2007). Existential Empathy: The Intimacy of Self and Other. In T. F. D. Farrow and P. W. R. Woodruff (eds.), *Empathy in Mental Illness* (pp. 310-321). Cambridge: Cambridge University Press.

Iacoboni, M. (2009). *Mirroring People: The Science of Empathy and How We Connect with Others.* New York: Picador.

Iannetti G.D., Mouraux, A. (2010). From the neuromatrix to the pain matrix (and back). *Experimental Brain Research* 205:1–12

León, F., & Zahavi, D. (2016). Phenomenology of experiential sharing: The contribution of Schutz and Walther. In A. Salice & H. B. Schmid (Eds.), *Social reality: The phenomenological approach*. Dordrecht: Springer (in press).

Legrain, V., Iannetti G.D., Plaghki, L., Mouraux, A. (2011). The pain matrix reloaded: a salience detection system for the body. *Prog Neurobiol* 93:111–124.

Lipps, T. (1907). Das Wissen von fremden Ichen. In T. Lipps (ed.), *Psychologische Untersuchungen I* (pp. 694-722). Leipzig: Engelmann.

Lipps, T. (1909). *Leitfaden der Psychologie*. Leipzig: Verlag von Wilhelm Engelmann.

Maibom, H.L. (ed.) (2014). *Empathy and Morality*. Oxford: Oxford University Press.

Merleau-Ponty, M. (2012). *Phenomenology of Perception*, trans. D. A. Landes. London: Routledge.

Michael, J. (2011a). Four Models of the Functional Contribution of the Mirror Neuron System. *Philosophical Explorations* 14 (2) 2011: 185-194.

Michael, J. (2011b). Interaction *and* Mindreading. *Review of Philosophy and Psychology* 2(3): 559-578.

Michael, J. & Fardo, F. (2014). What (if anything) is shared in pain empathy? A critical discussion of De Vignemont and Jacob's (2012) theory of the neural substrate of pain empathy. *Philosophy of Science* 81: 154–160.

Michael, J. (2014). Towards a consensus about the role of empathy in interpersonal understanding. *Topoi* 33 (1): 157-172.

Mouraux, A., Diukova, A., Lee, M.C., Wise, R.G., Iannetti, G.D. (2011). A multisensory investigation of the functional significance of the ''pain matrix''. *Neuroimage* 54:2237–2249.

Overgaard, S. & Michael, J. (2015). The interaction theory of social cognition – a critique. *Philosophical Psychology* 28, 160-183.

Reddy, V. (2008). *How Infants Know Minds*. Cambridge, Mass.: Harvard University Press.

Rochat, P. (2001). *The Infant's World.* Cambridge, Mass.: Harvard University Press.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2013). Toward a Second-Person Neuroscience. *Behavioral and Brain Sciences*, 36/4: 393–414.

Singer, T., Lamm, C. (2009). The social neuroscience of empathy. *Ann NY Acad Sci* 1156:81–89

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R., Frith, C. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303:1157–1162

Sinigaglia, C. (2008). Mirror Neurons: This Is the Question. *Journal of Consciousness Studies* 15/10-11: 70–92.

Stueber, K. R. (2006). *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, Mass.: MIT Press.

Titchener, E. B. (1909). *Lectures on the Experimental Psychology of Thought-Processes*. New York: Macmillan.

Zahavi, D. (2011). Empathy and Direct Social Perception: A Phenomenological Proposal. *Review of Philosophy and Psychology* 2/3, 541-558.

Zahavi, D. (2014a). *Self and Other: Exploring Subjectivity, Empathy, and Sharing*. Oxford: Oxford University Press.

Zahavi, D. (2014b). Empathy and Other-Directed Intentionality. *Topoi* 33/1, 129-142.

Zahavi, D., Overgaard, S. (2012). Empathy without isomorphism: A phenomenological account. In J. Decety (eds.): *Empathy: From Bench to Bedside*. Cambridge, MA.: The MIT Press, 3-20.

Zahavi, D., Rochat, Ph. (2015). Empathy ≠ sharing: Perspectives from phenomenology and developmental psychology. *Consciousness and Cognition* 36, 543-553.

Zahavi, D., Salice, A. (2016). Phenomenology of the we: Stein, Walther, Gurwitsch. In J. Kiverstein (ed.): *The Routledge Handbook of Philosophy of the Social Mind*. London: Routledge. In press.