

Evan Thompson

# *Empathy and Consciousness*<sup>1</sup>

*This article makes five main points. (1) Individual human consciousness is formed in the dynamic interrelation of self and other, and therefore is inherently intersubjective. (2) The concrete encounter of self and other fundamentally involves empathy, understood as a unique and irreducible kind of intentionality. (3) Empathy is the precondition (the condition of possibility) of the science of consciousness. (4) Human empathy is inherently developmental: open to it are pathways to non-egocentric or self-transcendent modes of intersubjectivity. (5) Real progress in the understanding of intersubjectivity requires integrating the methods and findings of cognitive science, phenomenology, and contemplative and meditative psychologies of human transformation.*

## **I: Preamble**

My aim in this article is to set forth a context for the following essays and a framework for future research on the topic of intersubjectivity in the science of consciousness. To this end, I will present, in broad strokes, an overview of this topic, drawing from three main sources — cognitive science and the philosophy of mind, continental European phenomenology, and the psychology of contemplative or meditative experience. Since my aim is integrative and constructive, I will not offer detailed conceptual and empirical arguments for each step, though I will try to give a taste of some of these arguments along the way.

---

[1] This article draws from two earlier papers, the first a consulting report prepared for the Fetzer Institute (Kalamazoo, MI) that also served as the discussion paper for the meeting I convened at Fetzer on ‘The Intersubjectivity of Human Consciousness: Integrating Phenomenology and Cognitive Science’ (September 24–27, 1999), and the second my opening address to this meeting. The discussions at this meeting have greatly influenced this article; therefore a special debt of gratitude is here acknowledged to all the participants: Yoko Arisaka, Jonathan Cole, Natalie Depraz, Joel Elkes, Shaun Gallagher, Al Kaszniak, Jim Laukes, Eduard Marbach, Lis Nielsen, Alva Noë, Sue Savage-Rumbaugh, Marilyn Schlitz, Brian Smith, S. Kay Toombs, Francisco J. Varela, Dan Zahavi, and Arthur Zajonc. I also wish to thank Margaret Donaldson for helpful critical comments on an earlier draft. Finally, I gratefully acknowledge the support of the McDonnell Foundation for a grant received through the McDonnell Project in Philosophy and the Neurosciences (Thompson, 1999), as well as helpful discussions with two of the McDonnell Project Advisors, Vittorio Gallese and Ralph Adolphs.

## II: Introduction

The theme of this article is that the individual human mind is not confined within the head, but extends throughout the living body and includes the world beyond the biological membrane of the organism, especially the interpersonal, social world of self and other. This theme, long central to the tradition of continental European phenomenology, derived from Edmund Husserl (1859–1938), has lately begun to be heard in cognitive science. Indeed, there is a remarkable convergence between these two traditions, not simply on the topic of intersubjectivity, but on virtually every area of research within cognitive science, as a growing number of scientists and philosophers have discussed (Varela, 1996; Gallagher, 1997; Petitot *et al.*, 1999). In the case of intersubjectivity, much of the convergence centres on the realization that *one's consciousness of oneself as an embodied individual in the world is founded on empathy* — on one's empathic cognition of others, and others' empathic cognition of oneself. Yet despite this convergence, to be explored in this article, many questions remain about how to understand the relationship between the cognitive scientific and the phenomenological treatments of consciousness. In the end, these questions all come back to the question of what kind of science the science of consciousness is or can be. Put another way, *if we are to have a cognitively and ethically satisfying understanding of consciousness, what form should this understanding take?*

To frame my discussion here, let me propose two key points that go to the heart of the matter. I call these points the Core Dyad:

### THE CORE DYAD

- **Empathy is the precondition (the condition of possibility) for the science of consciousness.**
- **Empathy is an evolved, biological capacity of the human species, and probably of other mammalian species, such as the apes.**

The first side of the Core Dyad comes from phenomenology. I will explain its meaning more fully later, but the basic idea is that the mind as a scientific object is an abstraction from, and hence presupposes, our empathic cognition of each other. The second side of the Dyad comes from cognitive science and is comparatively straightforward. My aim in putting the two together, side-by-side, is to create a kind of hub or axis for all of the many different issues that can be raised about intersubjectivity and consciousness as seen from the viewpoints of phenomenology and cognitive science. Underlying all these issues is the fundamental question of how to conceptualize or understand the relationship between these two poles. I will come back to this question later in the article.

## III: Enactive Cognitive Science and the Embodied Mind

The development of cognitive science over the past two decades or so has seen a movement from the classical, cognitivist view that an inner mind represents an outer world using symbols in a computational language of thought, to the view that mental processes are embodied in the sensorimotor activity of the organism and embedded in the environment (see Clark, 1999, and Beer, 2000, for recent discussions). This viewpoint has come to be known as *enactive* or *embodied* cognitive science (Varela *et al.*,

1991; Clark, 1997). Enactive cognitive science, as Francisco J. Varela and I currently conceive of it (Thompson & Varela, forthcoming), involves the following three theses:

- **Embodiment.** The mind is not located in the head, but is embodied in the whole organism embedded in its environment.
- **Emergence.** Embodied cognition is constituted by emergent and self-organized processes that span and interconnect the brain, the body, and the environment.
- **Self–Other Co-Determination.** In social creatures, embodied cognition emerges from the dynamic co-determination of self and other.

### *Embodiment*

Visual perception serves as a good illustration of the embodiment thesis. The perception of visual space, for instance, does not arise from a unified model of space in the brain, but from numerous spatial maps, many of which are located in cortical areas involved in the control of bodily movements (of the eyes, head, arms, and so on) (Rizzolatti *et al.*, 1994). Perceptual space is not a uniform external container, but rather a medium molded by our sensing and moving bodies: our movements ‘progressively carve out a working space from undifferentiated visual information’ and this ‘movement-based space . . . becomes then our experiential peripersonal visual space’ (Rizzolatti *et al.*, 1997, p. 191).

In general, from the enactive perspective, visual perception — or more simply, seeing — is a way of acting: it is visually guided exploration of the world (Thompson *et al.*, 1992; Thompson, 1995). As Kevin O’Regan and Alva Noë put it in a recent article: ‘Activity in internal representations does not generate the experience of seeing. . . . The experience of seeing occurs when the organism masters what we call the governing laws of sensorimotor contingency’ (O’Regan & Noë, in press).

### *Emergence*

Emergence pertains to systems in which local elements and rules give rise to global patterns of activity. What enactive cognitive science stresses is that emergence *via* self-organization is a two-way street involving circular causality (Kelso, 1995; Freeman, 1999; Thompson & Varela, in press). In addition to the ‘upwards’ causation of local interactions giving rise to global patterns, there is the reciprocal ‘downwards’ causation of global patterns controlling and modulating local interactions (e.g., by setting their context and boundary conditions). Thus, in addition to the ‘upwards’ causation of personal consciousness by neural and somatic activity, there is the reciprocal ‘downwards’ causation of neural and somatic activity by the active animal or person as a conscious agent. For example, Francisco J. Varela’s group in Paris has shown in epileptic patients that purposeful cognitive activity on the part of the patient changes the neurodynamic patterns of epileptic activity (Le Van Quyen *et al.*, 1997; Thompson & Varela, in press). Similarly, J.A. Scott Kelso has shown that the conscious intention of an agent to move a finger in a certain way is able (within limits) to stabilize one dynamic pattern of neural-somatic activity and destabilize another (Kelso, 1995, pp. 145–53). Thus ‘downwards’ (global-to-local) causation is no metaphysical will-o’-the-wisp, but a typical feature of complex (nonlinear) dynamical systems, and may occur at multiple levels in the coupled dynamics of brain, body and environment, including that of conscious cognitive acts in relation to local neural

activity. In Kelso's words: 'Mind itself is a spatiotemporal pattern that molds the metastable dynamic patterns of the brain' (1995, p. 288).

These two theses — the embodiment thesis and the emergence thesis — although by no means uncontroversial, have been reasonably well explored compared with the third thesis, explored in this article.

### *Self–other co-determination*

According to this thesis, embodied cognition emerges from the dynamic co-determination of self and other.<sup>2</sup> What recent cognitive science has begun to drive home is that the embodied mind is intersubjectively constituted at the most fundamental levels. Our own human self-consciousness, for example, emerges from a primordial and preverbal sense of self, present in newborn infants, that is inseparably coupled to the perceptual recognition of other human beings (Gallagher & Meltzoff, 1996; Meltzoff & Moore, 1999). To what extent this experiential coupling of self and other, operative from birth, is distinctively human or present in other primate or mammalian species remains an open issue for current research.

The thesis of self–other co-determination is linked to the (re)discovery of the importance of affect and emotion in cognition. Classical cognitive science was cognocentric: it conceived of cognition as the manipulation of affectless representations. New developments, especially in affective neuroscience, have shown that affect and emotion lie at the basis of the mind (Damasio, 1994; 1999; Panksepp, 1998a,b), particularly in the domain of social cognition, in which impaired social behaviour can be linked to fundamental deficits in affective cognition (see Adolphs, 1999, for a recent review).

The prominence of affect reinforces the two enactive theses of embodiment and emergence. Douglas F. Watt (1998) describes affect as 'a prototype "whole brain event"', but we could go further and say that affect is a prototypical whole-organism event. Affect has numerous dimensions that bind together virtually every aspect of the organism — the psychosomatic network of the nervous system, immune system, and endocrine system; physiological changes in the autonomic nervous system, the limbic system, and the superior cortex; facial-motor changes and global differential motor readiness for approach or withdrawal; subjective experience along a pleasure-displeasure valence axis; social signalling and coupling; and conscious evaluation and assessment (Watt, 1998). Thus the affective mind isn't in the head, but in the whole body; and affective states are emergent in the reciprocal, co-determination sense: they arise from neural and somatic activity that itself is conditioned by the ongoing embodied awareness and action of the whole animal or person.

Now, having just described affect as a prototypical whole-organism event, I wish to go one step further and say that much of affect is a prototypical *two-organism event*, by which I mean a prototypical *self–other event*. I now turn to review some of the diverse evidence for this point, evidence that gives more substance to the thesis of self–other co-determination and to the second pole of the Core Dyad (that empathy is an evolved, biological capacity).<sup>3</sup>

[2] See **Arisaka** (this volume) for the elaboration of a parallel point in the context of twentieth-century Japanese philosophy.

[3] This review is by no means exhaustive. Rather, it is meant as a representative sample of some of the most important relevant lines of research.

#### IV: Intersubjectivity — Views From Cognitive Science

Affective neuroscience: basic emotional operating systems of the mammalian brain  
 Jaak Panksepp (1998a,b) has proposed that the panoply of emotional states we experience can be analysed into certain core affective compartments, probably common to all mammals, that depend on distinct, basic emotional operating systems in the brain, and are tied to an animal's social and biological relationships to conspecifics and members of other species. These core affective compartments are seeking/expectancy, rage/anger, fear, nurturance/sexuality, social bonding/separation distress, and play/joy. (**Cheyne**,<sup>4</sup> this volume, provides a fascinating exploration of the link between fear and the endogenous activation of a hypervigilant state — as in the face of a carnivorous predator — and the sensed presence of the Other during sleep paralysis and hypnagogic hallucinations.) Panksepp hypothesizes that each of these affective compartments is subserved by its own core neural network in the midbrain-diencephalon. Each network has certain key chemical neuromodulators, and all project to an area called the periaqueductal grey (PAG). Panksepp proposes that the PAG serves as the substrate for a primordial sense of self, again probably common to all mammals, because in this area 'there is a massive convergence of a diversity of basic emotional systems (fundamental value schema), various simple sensory abilities (perceptual schema), and primitive but coherent response systems (action schema)' (1998b, p. 568).

##### *Cognitive ethology: empathy in primate life*

In primates, especially apes and humans, affective compartment clearly plays a huge role in the interpretation and understanding of mental states, both of others and of oneself (**Savage-Rumbaugh, Fields & Tagliatela**, this volume; **Smuts**, this volume). Higher primates excel at interpreting others as psychological subjects on the basis of their bodily presence — their facial expressions, postures, vocalizations, and so on (Povinelli & Preus, 1995). It is here that we see affective compartment blossoming into empathy, in the sense of a meta-affective cognitive capacity for grasping another's point of view.

The presence and extent of empathy among animals is currently a subject of much debate (Gallup, 1998; Povinelli, 1998). One central thread in this debate concerns the 'mirror test' for self-recognition, which Gordon Gallup introduced in the 1970s. An individual unknowingly received a dot of red dye placed above the eyebrow so that it would be invisible without a mirror. Chimpanzees and orangutans, as well as human children more than 18 months old, guided by their reflection in a mirror, rubbed the spot with their hands and inspected their fingers after touching it, thus apparently recognizing that the dot on the reflected image was on their own face. Other animals, including a variety of primates, failed to connect the reflected image to their own bodies. Thus it seems that chimpanzees and orangutans have a self-concept and a capacity for self-recognition, whereas other animals do not. Gallup went on to argue that self-recognition implies self-awareness and 'that such self-awareness enables these animals to infer the mental states of others. In other words, species that pass the mirror test are also able to sympathize, empathize, and attribute intent and emotions

---

[4] Names in bold type refer to the contributors' articles in this volume.

in others — abilities that some might consider the exclusive domain of humans' (Gallup, 1998, p. 66).

The mirror test is controversial: most agree that the test is evidence for possession of a self-concept and that such a concept appears to be restricted to humans and the great apes, but some argue that such a capacity for self-recognition in a mirror does not imply awareness of one's own psychological states and the understanding that others possess such states (Povinelli, 1998). For this reason, it seems better not to take the mirror test in isolation, and instead to ask the more general question of which elements of human intersubjectivity are recognizable in other animals (de Waal, 1996, p. 79).

The primatologist Frans de Waal, in his landmark book, *Good Natured: The Origins of Right and Wrong in Humans and Other Animals* (de Waal, 1996), makes the important point that empathy is not an all-or-nothing phenomenon:

Many forms of empathy exist intermediate between the extremes of mere agitation at the distress of another and full understanding of their predicament. At one end of the spectrum, rhesus infants get upset and seek contact with another as soon as one of them screams. At the other end, a chimpanzee recalls a wound he has inflicted, and returns to the victim to inspect it (de Waal, 1996, p. 69).

De Waal shows that the animal kingdom exhibits a wide range of other-involved behaviour of various degrees of complexity and sophistication — parental care-giving, succorant behaviour to endangered individuals other than progeny; emotional contagion (vicarious arousal by the emotions of others), and cognitive empathy and sympathy (see also Sober & Wilson, 1998). Underlying all caring behaviour is mutual attachment and bonding, and therefore the ultimate evolutionary source for this kind of behaviour is parental care. Succorant behaviour — helping or providing care to distressed individuals other than progeny — emerges from parental care-giving, but radiates outward to include the social group. To this can be added emotional contagion, which, at its simplest, takes the form of 'total identification without discrimination between one's feelings and those of the other' (de Waal, 1996, p. 80). Cognitive empathy emerges as a further step, in which there is recognition of the other's experience as belonging to the other, without losing the distinction between self and other in emotional contagion. Cognitive empathy in turn makes possible the moral emotions of sympathy and compassion, in which we feel genuine concern for the other. These emotions require certain cognitive abilities and a well developed sense of self, both of which may not be widespread in the animal world: they seem limited to humans and, to varying extents, our evolutionary cousins, the great apes.

#### *Evolutionary neurobiology: sociability and large brains*

We have seen that parental care is the evolutionary source of attachment and bonding, and that the affective compartments involved in these forms of sociality — such as nurturance/sexuality and social bonding/separation distress — depend on basic emotional operating systems of the limbic system in the mammalian brain. According to John Morgan Allman (1999), it was the formation of the extended family as a social support structure for the nurturing of slowly developing offspring that drove the evolution of large brains in apes and humans: 'the development of the brain to the level of



complexity we enjoy — and that makes our lives so rich — depended on the establishment of the human family as a social and reproductive unit' (Allman, 1999, p. 2).

*Developmental neuroscience:*

*care-giving and early environmental regulation of brain development*

Evidence is now accumulating that experience-dependent brain activity in particular environmental contexts plays a huge role in the development of the individual brain. Rather than being a collection of pre-specified modules, the brain appears to be an organ that constructs itself in development through spontaneously generated and experience-dependent activity (Quartz & Sejnowski, 1997; Quartz, 1999; Karmiloff-Smith, 1998), a developmental process made possible by robust and flexible developmental mechanisms conserved in animal evolution (Gerhart & Kirschner, 1997). In mammals — especially human beings, who undergo a protracted period of development even in comparison with other primates — much of early experience takes place in the social context of parental care. It has been shown that the care infants receive regulates the synaptic and chemical properties of the brain, including even altering the rates at which certain genes are expressed, with implications for the long-term viability and health of the organism (Meany *et al.*, 1996). Such findings show that the affective interaction of self and other in infancy can modify the very constitution of the living body. (See also **Savage-Rumbaugh *et al.***, this volume, for related points on brain development in bonobo chimpanzee infants raised in a mixed chimpanzee/human or *Pan/Homo* culture).

*Developmental psychology: infant imitation*

There is now a large amount of evidence showing that human infants possess, at birth, interpersonal body schemas for emotional contagion and facial imitation, and that these schemas underlie the development of more sophisticated empathic abilities (Meltzoff & Moore, 1999). Studies have shown that newborns (less than an hour old in some cases) can imitate the facial gestures of another person (Meltzoff & Moore, 1994). This kind of imitation is known as 'invisible imitation' because the infant uses parts of his body invisible to himself to imitate the other's movements. For this kind of imitation to be possible, the infant must be able to match a visual display (the facial movements of the other) to his own motor behaviour; therefore, he must have a developed body schema, which organizes his experience of his own body's position and movement, and to which he can relate the visible gestures of the other person. If imitation requires such a body schema, and newborns can imitate, then such a schema must be operative from birth, rather than having to wait upon the infants' acquisition of a visual image of themselves (contrary to what earlier theorists, such as Piaget, proposed).

How would such a schema work? The basic idea is that the infant, faced with novel gestures, uses her proprioceptive awareness of her own unseen facial movements to copy what she sees in the face of the other person. This performance depends upon a 'supramodal' body schema that enables the infant to recognize equivalencies between herself and the other person (Meltzoff & Moore, 1994; 1999; Gallagher & Meltzoff, 1996). The schema links the perceptual modalities of vision and proprioception, both to each other and to the motor processes of action. As a result, the gestures of the other are recognizable to the infant in the terms of her own

proprioceptive awareness (*via* the intermodal link between vision and proprioception), and she is able to move so that her proprioceptive awareness of her own body coincides with what she sees (*via* the intermodal link between visuo-proprioception and motor action). There is no need to learn to translate back-and-forth between vision, proprioception, and action, because from the start the senses are linked to each other and to possibilities of action in a supramodal and interpersonal body schema.

These findings about infant imitation call for the revision of earlier conclusions made by Merleau-Ponty (1962) about the phenomenology of infant experience. Merleau-Ponty seems to have believed, on the basis of the psychology available to him at the time, that development unfolds ‘from the inside out’: the sensorimotor equivalence between vision and proprioception must first be established in one’s own case, and only then is transferable to one’s perception of the world and others. The research just mentioned, however, implies that the trajectory of the interpersonal dynamic is precisely *not* ‘from the inside out’. Although an intracorporeal schema makes possible the interpersonal dynamic, the schema operates *intercorporeally* from the start:

No . . . transfer [from self to other] is necessary because it is already accomplished, and already intersubjective. A supramodal code already reaches across the child’s relations with others. . . . From early infancy . . . the visual experiences of the other person communicate in a code that is related to the self. This communication is organized on the basis of an innate system that does not necessarily give priority to body experience over and against the experience of the other. . . . The body schema, working systematically with proprioceptive awareness, operates as a proprioceptive self that is always already ‘coupled’ with the other (Gallagher & Meltzoff, 1996, pp. 225–6).

Merleau-Ponty also believed, again on the basis of the psychology of his time, that the infant’s early experience did not involve a differentiation of self and other, because such a differentiation must wait for the acquisition of a body schema. Yet invisible imitation in newborns clearly shows that such a body schema is operative from birth, and therefore does not have to be acquired in the way Merleau-Ponty envisioned:

The phenomenon of newborn imitation suggests that much earlier [than self-recognition in a mirror] there is a ‘primordial’ or ‘embryonic’ notion of self, what we might call a proprioceptive self — a sense of self that involves a sense of one’s motor possibilities, body postures, and body powers, rather than one’s visual features. The newborn infant’s ability to imitate others, and its ability to correct its movement, which implies a recognition of the difference between its own gesture and the gesture of the other, indicates a rudimentary differentiation between self and non-self. This may be a bare framework of self based on an innate body system, but it serves to introduce a disruptive moment into the supposed indifferentiation of the earliest hours. Furthermore, it suggests that this earliest period is not a ‘pre-communication’ phase, but is already an experience of pre-verbal communication in the language of gesture and action. And this, we note, would actually support some of Merleau-Ponty’s other views about the relation between the infant and language (Gallagher & Meltzoff, 1996, p. 227).

This intersubjective framework of self and other becomes increasingly refined as children grow and develop. Between the ages of two and five, children begin to be able to interpret themselves and others in the human psychological framework of thoughts, feelings, beliefs, desires, and perceptions (Astington, 1993). This



interpretive ability reflects a particular type of social intelligence, which, from an evolutionary standpoint, is likely to be a recent innovation, one that probably emerged before the divergence of humans and apes, for it seems to be present among great apes such as chimpanzees (de Waal, 1996; Gallup, 1998; **Savage-Rumbaugh et al.**, this volume; **Smuts**, this volume). Indeed, there are striking similarities in the development of children's and chimpanzees' psychological abilities, although there are crucial differences too, ones which exemplify the cognitive refinements in the evolution of the human interpersonal dynamic of self and other (see Povinelli & Preuss, 1995; Povinelli, 1998).

*Cognitive neuroscience: mirror neurons*

Giacomo Rizzolatti, Vittorio Gallese, and their colleagues have uncovered in area F5 of the premotor cortex in macaque monkeys a class of neurons they call 'mirror neurons' (see **Gallese**, this volume). These neurons display the same pattern of activity, both when the animal accomplishes certain goal-directed hand movements, and when the animal observes the experimenter performing the same actions. Of particular note is that the activity of the neurons is correlated with specific motor acts (defined by the presence of a goal) and not with the execution of particular movements, such as contractions of particular muscle groups. The neurons can be classified according to the type of action, such as 'grasp with the hand', 'grasp with the hand and mouth', 'reach', and so on. All the neurons of the same type encode actions that meet the same objective. On the basis of these properties, mirror neurons appear to form a cortical system that matches the observation and the performance of motor actions. There is also evidence in humans for such a mirror neuron system for gesture recognition (Gallese & Goldman, 1998).

These findings are notable for several reasons. First, the neural system for recognizing the intentional meaning of the actions of another agent appears to be primarily of a practical nature, rather than inferential or judgmental, for it involves the direct pairing or matching of the bodies of self and other. There seems to be an immediate pairing between the animal's understanding of its own actions and its understanding of those of another, an understanding whose structure is not that of an initial perception of a non-interpreted bodily movement followed by a judgement that attributes meaning to the movement and thereby interprets it as an action. Rather, the movement of the other is already understood as a goal-directed action because of its match to a self-performed action. It seems that primates recognize actions made by others because the neural pattern of activity in their premotor areas when they observe an action is similar to that internally generated to produce the same type of action. As we will see, this kind of non-inferential bodily pairing of self and other is one of the hallmarks of the phenomenological analysis of empathy. Indeed, the mirror neuron findings support Husserl's position that our empathic experience of another depends on one's 'coupling' or 'pairing' with the other (**Depraz**, this volume), rather than some kind of affective fusion, as some of Husserl's contemporaries held (see Petit, 1999).

Second, Rizzolatti and Arbib (1998) have proposed that the gesture recognition mirror neuron system may be part of the basis for the development of language (see also Corballis, 1998). The evidence for a mirror neuron system for gesture recognition in humans includes Broca's area, an area that is known to be involved in speech and is probably the homologue of area F5 of the monkey premotor cortex. Rizzolatti

and Arbib propose that ‘the development of the human lateral speech circuit is a consequence of the fact that the precursor of Broca’s area was endowed, before speech appearance, with a mechanism for recognizing actions made by others. This mechanism was the neural prerequisite for the development of interindividual communication and finally of speech’ (Rizzolatti & Arbib, 1998, p. 190).

This development, as Rizzolatti and Arbib envision it, comprises two main steps, corresponding to two gaps that have to be bridged on the path from action recognition to speech — first, the gap between recognizing actions made by others and sending and receiving messages with communicative intent; and second, the gap between gestural communication and speech. In general, premotor areas are activated when an individual is about to perform an action or observes another individual performing an action. There are usually mechanisms that inhibit the observer from emitting a motor behaviour that mimics the observed action, and that inhibit the actor from initiating the action prematurely. But the premotor system sometimes will allow a brief prefix of the movement to be exhibited, and this prefix will be recognizable by the other individual:

This fact will affect both the actor and the observer. The actor will recognize an intention in the observer, and the observer will notice that its involuntary response affects the behavior of the actor. The development of the capacity of the observer to control his or her mirror system is crucial in order to emit (voluntarily) a signal. When this occurs, a primitive dialogue between the observer and actor is established. This dialogue forms the core of language (pp. 190–1).

Thus the first gap between action recognition and communication is bridged in the form of a mimetic dialogue.

Merlin Donald (1991) has proposed that the evolution of the mimetic capacity — a capacity that still figures centrally in human culture — was a necessary precursor to the evolution of language. Rizzolatti and Arbib build on this proposal. They argue that in the case of individual-to-individual communication, the gestures that were most likely to be used first were oro-facial ones, because these are used extensively by monkeys, apes, and humans for communication. It seems unlikely, however, that speech arose from oro-facial gestures alone, because they limit communication to two actors at a time. But if manual gestures (e.g., pointing) are associated with oro-facial ones, then the communicative possibilities increase considerably:

These considerations suggest that, at a certain stage, a brachio-manual communication evolved complementing the oro-facial one. This development greatly modified the importance of vocalization and its control. Whereas during the closed oro-facial stages, sounds could add very little to the gestural message . . . their association with gestures allowed them to assume the more open, referential character that brachiomanual gestures had already achieved. An object or event described gesturely . . . could now be accompanied by vocalization. If identical sounds were constantly used to indicate identical elements . . . a primitive vocabulary of meaningful sounds could start to develop (p. 193).

This development bridges the second gap between gestural communication and speech.

### *Philosophy of mind and psychology: theories of ‘mind-reading’*

In the philosophy of mind and psychology, intersubjectivity and empathy have been discussed in the context of the debate between the ‘theory-theory’ and the

‘simulation-theory’ of ‘mind-reading’ — our capacity to attribute mental states to ourselves and others, and to understand our behaviour and actions in light of those attributions (Davies & Stone, 1995a,b; Carruthers & Smith, 1996). According to the theory-theory (TT), normal human adults are able to mind-read because they possess a commonsense or folk-psychological ‘theory of mind’ that they employ to explain and predict human behaviour. Many advocates of the theory-theory (though not all) consider this folk-psychological body of knowledge to be essentially equivalent to a scientific theory: mental states, according to this view, are unobservable entities (like electrons), and our attribution of mental states to each other involves causal-explanatory generalizations (comparable in form to those of physics) that relate mental states to each other and to observable behaviour. According to the simulation-theory (ST), on the other hand, mind-reading depends not on the possession of a tacit psychological theory, but on the ability to mentally ‘simulate’ another person, that is, on being able to use the resources of one’s own mind to create a model of another person and thereby identify with him or her, projecting oneself imaginatively into his or her situation.

Although there are different versions of these two theories, as well as various ways in which they might be reconciled or combined, there remain important differences of emphasis between them. As Gallese and Goldman (1998, p. 497) observe in their discussion of mirror neurons and the simulation theory: ‘The core difference between TT and ST, in our view, is that TT depicts mind-reading as a thoroughly “detached” theoretical activity, whereas ST depicts mind-reading as incorporating an attempt to replicate, mimic, or impersonate the mental life of the target agent’ (see also **Gallese**, this volume). Similarly, Gordon (1996, p. 11) describes the theory-theory as a ‘*cold* methodology . . . that chiefly engages our intellectual processes, moving by inference from one set of beliefs to another, and makes no essential use of our own capacities for emotion, motivation, and practical reasoning,’ and the simulation-theory as a ‘*hot* methodology, which exploits one’s own motivational and emotional resources and one’s own capacity for practical reasoning.’

Given this difference, it is not surprising that empathy figures prominently in the simulation-theory account of mind-reading (Goldman, 1993; 1995a,b; Gordon, 1996). According to this account, the simulation heuristic by which we understand the mental states of other subjects has three main elements:

The initial step . . . is to imagine being ‘in the shoes’ of the agent. . . . This means pretending to have the same initial desires, beliefs, or other mental states that the attributer’s background information suggests the agent has. The next step is to feed these pretend states into some inferential mechanism, or other cognitive mechanism, and allow that mechanism to generate further mental states as outputs by its normal operating procedure. . . . More precisely, the output should be viewed as a pretend or surrogate state, since presumably the simulator doesn’t feel the *very same* affect or emotion as a real agent would. Finally, upon noting this output, one ascribes to the agent an occurrence of this output state. Predictions of behavior would proceed similarly. . . . In short, you let your own psychological mechanism serve as a ‘model’ of his (Goldman, 1995a, p. 189).

Empathy, on this view, is a special case of mental simulation, in which the output states are affective or emotional states: ‘empathy consists of a sort of “mimicking” of one person’s affective state by that of another’ (Goldman, 1995a, p. 198).

Overall, the simulation-theory seems more attractive than the theory-theory, because of the role it gives to affect, emotion, and empathy in the co-determination of self and other. The theory-theory is cognocentric in the manner of classical cognitive science: it emphasizes belief-like representations and has very little to say, if anything, about emotion. The simulation-theory seems more congruent with the message coming from affective neuroscience that affect and emotion are basic to social cognition (Adolphs, 1999).

Nevertheless, I think that the simulation-theory does not provide a satisfactory account of empathy. First, it is not clear whether the hypothetical simulation heuristic is supposed to be a subpersonal mechanism of the ‘cognitive unconscious’ or a structure of personal consciousness (but see **Gallese**, this volume, for further discussion). Second, in either case, the strategy of the simulation-theory is to begin from the individual self and then try to work outward to other selves through the mechanisms of mimicry and imaginative projection. But these mechanisms on their own cannot account for the openness of the self to the other; on the contrary, the self must already be ‘intersubjectively open’ in its very structure for these mechanisms to function effectively at all. (This notion of ‘intersubjective openness’ will be explained more precisely in the next section.) Mimicry and the imaginative transposition of oneself to the place of the other are no doubt elements of empathy, but they are founded on more fundamental pre-reflective couplings of self and other at the level of the lived body: it is the passive (not voluntarily initiated), pre-reflective experience of the other as an embodied being like oneself that sets the stage, as it were, for mimicry and the more elaborate mental act of imaginative self-transposal (see Section VI). As Max Scheler pointed out in his *The Nature of Sympathy*: ‘imitation, even as a mere “tendency,” already presupposes some kind of acquaintance with the other’s experience, and therefore cannot explain what it is supposed to do . . . the impulse to imitate only arises when we have already apprehended the gesture as an expression of fear or joy’ (1954, p. 10, as quoted in Hamrick, 2000; this article presents a number of other important criticisms of the simulation-theory from a phenomenological perspective).

Finally, the very terms of the simulation-theory versus theory-theory debate about mind-reading seem problematic. The presupposition both theories share is that mind-reading is primarily a ‘spectatorial’ process of explanation and prediction (McGeer, 1999). On this way of thinking, self and other stand in relation to each other as observer and observed, and mind-reading is a proto-scientific activity, conceived either as causal-explanatory generalization or as model-building and simulation. This view of intersubjective relations seems distorted, for as phenomenologists have long argued, these relations are primarily embodied, practical, and mutually defining in nature (**Gallagher**, this volume). For this reason, **Victoria McGeer’s** ‘know-how’ or ‘regulative’ account of folk psychology (this volume; see also McGeer, 1996; 1999) is an important advance beyond the theory-theory versus simulation-theory debate, one that can establish a meeting ground for Anglo-American philosophy of mind and continental European phenomenology:

There is something right about understanding our folk-psychological abilities in terms of the predictive/explanatory power they buy for us. But this has to be understood in a larger context. I claim that our mindreading skills are part of a more fundamental capacity, or set of capacities, for regulating our own minds in concert with others. That is, we regulate our minds in accord with the intersubjective norms of folk-psychology, becoming in the

process agents that are well predicted and explained from the intentional stance. In becoming well-regulated folk-psychologists, we also develop the capacity for reading the minds of others, especially (normal) human others whose patterns of thought and action are similarly stabilized by the regulative function of folk-psychological norms. A plausible evolutionary account of why this capacity developed in our species concerns its value for us as social beings, to be sure, but, unlike the standard account, it focuses on the prior need to create and maintain stable patterns of behaviour amongst ourselves in order that we can predict and explain one another in a reliably robust way. On this view, a developmental account of the human mind must explain how we develop an active capacity for intentional self-regulation, rather than simply a spectatorial capacity for mindreading (McGeer, 1999).

From the perspective of phenomenology, our active capacity for intentional self-regulation is grounded on the perceptuo-motor capacities of our lived body. The theory-theory and the simulation-theory both take mind-reading to be a matter of how we infer from outward behaviour that others possess unobservable inner mental states (what they disagree about is the nature of the internal representations we use to make these inferences), and thereby they foster a conception of the mental as an inner realm separated from outward behaviour by an epistemic gulf that can be crossed only by inference.<sup>5</sup> In contrast, phenomenologists have long emphasized the importance of affective engagement and perceptuo-motor schemas as the basis of empathy and intersubjective understanding; thus from the start the phenomenological analyses have always been grounded on the idea that self and other recognize each other first and foremost as *persons* (Kern & Marbach, this volume), and hence as living bodily subjects or embodied agents, not as inner mental spectators of the outer world. Let us now take a closer look at this perspective.

### V: Phenomenology and Intersubjectivity

The topic of intersubjectivity in phenomenology is vast and cannot possibly be reviewed comprehensively here (see Zahavi, this volume, for an overview). For this reason, I will concentrate on certain general features of the approach to intersubjectivity taken in Husserlian phenomenology. Many Anglo-American analytic philosophers believe that Husserl's final position on the issue of intersubjectivity is to be found in his *Cartesian Meditations* (Husserl 1960), a position that is arguably solipsistic and hence incapable of accounting for intersubjectivity (though, it must be pointed out, the putative solipsism is transcendental, certainly not empirical). Yet during the last period of his thought Husserl revisited over and over again the issue of intersubjectivity and developed a much richer account, both of intersubjective experience and of phenomenology as an intersubjective endeavour (Husserl, 1973; see also Marbach and Kern, this volume; Depraz, this volume; Zahavi, this volume; as well

---

[5] The same conception plagues most discussions of the 'hard problem' of consciousness. Thus David Chalmers writes: 'It also seems that this is as good a solution to the problem of other minds as we are going to get. We note regularities between experience and physical or functional states in our own case, postulate simple and homogeneous underlying laws to explain them, and use those laws to infer the existence of consciousness in others. This may or may not be the reasoning that we implicitly use in believing that others are conscious, but in any case it seems to provide a reasonable justification for our beliefs' (1996, p. 246). The inadequacies of this type of account had already been detailed by Max Scheler in 1912 (see Zahavi, this volume).

as Depraz, 1995; Steinbock, 1995; Zahavi, 1999). There are two key features of this richer account that I wish to emphasize here: the first is that consciousness is intrinsically ‘intersubjectively open’, that is, it is structurally open to the Other in advance of any actual, concrete encounter of self and other (Zahavi, 1996; 1997); the second is that one’s awareness of oneself as an embodied individual embedded in the world depends on empathy, in particular on one’s empathic grasp of the Other’s empathic grasp of oneself (this second point is taken up in Section VI).

### *Open intersubjectivity*

Drawing from Husserl, Dan Zahavi (1997) has presented a phenomenological analysis of perceptual experience that reveals how the first-person perceptual experience of an objective thing is founded on the open intersubjectivity of consciousness. The point of this analysis is to show that the open intersubjectivity of consciousness is fundamental not simply for the encounter between oneself and other persons, but for any experience of the world at all.

When I perceive a thing before me, the thing is given to me (experienced by me) as having sides or profiles that I do not currently see. Indeed, it belongs to the very sense or meaning ‘objective thing’ — a sense or meaning implicit in my very perception — that, at any given moment, the thing comprises a plurality of co-existing profiles. If this aspect were not implicit in my perception, then my perceptual experience would present to me a two-dimensional image, not a three-dimensional thing. But despite long prejudice to the contrary in the empiricist tradition, our immediate experience is not of two-dimensional sensory patches, on the basis of which we then infer to the presence of things in the world. Rather, we automatically ‘appresent’ to ourselves the absent sides in and through our perception of the present sides.<sup>6</sup> How exactly are we to understand the structural relationship in perceptual experience between the present profile and the absent profiles? More pointedly, is it the case that this structural relationship can be accounted for entirely in terms that do not imply any reference to another perceiving consciousness besides myself? Husserl’s answer is no: the co-intended or appresented profiles must be understood as the correlates of the possible perceptions of another subject.

To retrace the main steps leading to this conclusion, let us suppose, first, that the absent profiles are appresented as profiles given in *my* past perceptions or possible future perceptions. The problem with this interpretation is that it makes the unity of the perceptual object — its ‘thingness’ for experience — the result of a series of temporally separated profiles, but this does not match our experience: the profile facing me is not experienced as present with respect to a past or future absent profile, but as present with respect to other co-present, absent profiles. Perhaps, then, we should say that the absent profiles are appresented as the correlates of the perceptions that I would have right now if I were over there looking at the thing, rather than here. In other words, perhaps the absent profiles are appresented as the correlates of my fictitious co-present perceptions (fictitious because it is impossible for them to be had simultaneously by me: I cannot see the thing from more than one perspective at once).

---

[6] Husserl uses the term ‘appresentation’ to refer to one’s intending in a mental act the presence of one thing on the basis of the actual presence of another: thus, in perception, we appresent or co-intend the backside of an object, which is not immediately present to us, on the basis of what is present, the front side.



But this interpretation too cannot account for the perceptual unity of the object, because it makes the unity a composite of fictitious slices (the correlates of the fictitious perceptions). What the object looks like from another vantage point is a contingent matter, but that there are other vantage points from which it can be seen is a necessity, and therefore cannot be explicated as a fictitious (non-actual) possibility.

The upshot of these considerations is that the co-intended absent profiles cannot be correlated with *my* possible perceptions, but must be understood as the correlates of the possible perceptions of an *Other*.<sup>7</sup> Yet clearly there need be no actual Other present for me to be able to co-intend the absent profiles, nor need there even be any others present in the world at all (as Husserl says, I might be the last survivor of a universal plague). What must be the case, however, is that the very meaning or sense of my perceptual experience refer to the perceptions of possible others. Thus first-person perceptual experience is essentially *intersubjectively open*, ‘and precisely for that reason . . . is incompatible with any solipsism which, in principle, would deny the possibility of a plurality of subjects’ (Zahavi, 1997, p. 312).

#### *Open intersubjectivity and concrete intersubjectivity*

One of the main contributions of Husserlian phenomenology has been to provide an ‘experiential logic’ (Depraz, this volume) of some of the different kinds of intersubjectivity. In addition to the open intersubjectivity of consciousness just discussed, two other kinds of intersubjectivity can be distinguished in Husserl’s analyses — the concrete bodily experience of the Other, and the generative/generational intersubjectivity of communally handed-down norms, conventions, and historical traditions (Zahavi 1996; 1997; Steinbock 1995). For the moment, I wish to consider only the relation between the open intersubjectivity of consciousness and intersubjectivity as the concrete bodily experience of the Other (here I follow Zahavi, 1997, pp. 313–19).

One might think that the intersubjective openness of consciousness depends on one’s concrete perceptual experience of the Other — in other words, that one’s concrete perceptual experience of the Other is the basis for the intersubjective openness of consciousness. One could argue, for instance, that when I experience another person as experiencing me, I realize (tacitly or pre-reflectively) that I am an other for the Other, that I am given to the Other as the Other is given to me, and thus that I am only one among many in a context of Others. In this way, my consciousness becomes intersubjectively opened, as it were, from the outside.

The problem with this account, however, is that it gets things backwards: for me to *perceive* the Other, the open intersubjectivity essential to perceptual experience must already be in play. Thus ‘the actual experience of another embodied subject is founded upon an *a priori* reference to the Other’ (Zahavi, 1997, p. 315). For the same reason, the intersubjective openness of consciousness cannot be reduced to any

[7] One might argue that although the co-intended absent profiles cannot be correlated with the correlates of my fictitious *co-present* perceptions, they can nevertheless be correlated with the correlates of my perceptions were I to walk around the thing and look at it from over there. The problem with this interpretation, however, is that it makes the absent profiles the correlates of my possible future perceptions, which has already been shown to be inadequate. Moreover, this interpretation itself involves the open intersubjectivity of consciousness, in the form of the *alterity* or *otherness* built into consciousness, for it requires that one imagine or otherwise mentally grasp oneself as *altered* or *othered* with respect to one’s present self.

contingent and factual relation of self and other; it must belong *a priori* to the very structure of subjectivity (**Zahavi**, this volume). At the same time, there is clearly much more to our embodied experience of other subjects than its mere dependence on the open intersubjectivity of consciousness: ‘Whereas the contribution of the *open intersubjectivity* is primarily at play in the formal structure of our intentionality, *concrete intersubjectivity* (which for the first time allows dissension and thus a sharpened experience of the *alterity* of the Other) is the condition of possibility for the central change in our categories of validity [from subjective to intersubjectivity validity]’ (Zahavi, 1997, p. 317).

What I propose to do now is to consider one aspect of concrete intersubjectivity in more depth, namely, the nature of empathy.

### VI: Empathy

According to Husserl, and those who have followed his analysis, such as Edith Stein in her 1916 doctoral dissertation prepared under Husserl’s direction, *On the Problem of Empathy* (Stein, 1964), empathy is a unique and irreducible kind of intentional experience: although it is based on sense perception and may involve inference (in difficult or problematic situations), it is not reducible to some additive combination of two, after the fashion of the theory that we understand others by perceiving their bodily behaviour and then inferring or hypothesizing that their behaviour is caused by particular experiences or inner mental states. Rather, we experience another person as a unified whole through empathy.

(Let me note parenthetically that **Steinbock**, this volume, provides an important critical discussion of the assumption that the only manner in which beings who are persons can be experienced is in the mode or manner of perceptual ‘presentation’, that is, in the mode of givenness belonging to perceptual objects and, according to this assumption, to perceptually empathized subjects. His discussion of ‘revelation’ as the mode in which the *person as such* is given, and of ‘loving’ as the act by which the person is ‘revealed’, resonates well with **Pitkin’s** contribution, and should be compared also to **Zahavi’s** discussion of intersubjectivity ‘beyond empathy’.)

Stein describes empathy as the experience of feeling led by an experience that is not one’s own, and distinguishes in it ‘three levels or modalities of accomplishment, even if in a concrete case people do not always go through all levels but are often satisfied with one of the lower ones’ (1964, p. 11). First, the experience of another emerges before me: ‘it arises before me all at once, it faces me as an object (such as the sadness I “read in another’s face”)’ (p. 10). Second, I can inquire into the content of the experience and ‘its implied tendencies’, in which case I become directed toward the object of the experience, that is, I imaginatively transpose myself to the place of the other subject to comprehend the object of the subject’s experience from his or her point of view. Third, once this clarification of the other’s experience is complete, the experience faces me again, but now in a clarified or explicated way. Stein refers to these three levels as: ‘(1) the emergence of the experience, (2) the fulfilling explication, and (3) the comprehensive objectification of the explained experience’ (p. 11).

It is also possible for these levels of empathy to be ‘reiterated’ back onto me, such that I can empathetically grasp the other’s empathic experience of me. In other words,

‘among the acts of another that I grasp empathetically there can be empathetic acts in which the other grasps another’s acts. This “other” can be a third person or me myself. In the second case we have “reflexive sympathy” where my original experience returns to me as an empathized one’ (p. 18).

Within the full performance of empathy, then, we can distinguish at least four possible kinds of empathy (**Depraz**, this volume):

- (1) The passive association of my lived body with the lived body of the Other
- (2) The imaginative transposal of myself to the place of the Other
- (3) The interpretation or understanding of myself as an Other for you
- (4) Ethical responsibility in the face of the Other

The first sort of empathy is passive (not voluntarily initiated on the part of the ego), pre-reflective, and bodily; it serves as the support for the others. When we see another person, we do not perceive his or her body as a mere physical thing, but rather as a lived body like our own. Thus empathy is not simply the grasping of another person’s particular experiences (sadness, joy, and so on), but on a more fundamental level the experience of another as an embodied subject of experience like oneself.

This sort of empathy occurs through the immediate ‘pairing’ or ‘coupling’ of the bodies of self and other in action. We find here a clear connection between phenomenology and recent cognitive neuroscience, in particular to the mirror neuron findings discussed earlier (see Petit, 1999; **Gallese** this volume). ‘Every time we are looking at someone performing an action, the same motor circuits that are recruited when we ourselves perform that action are concurrently activated’ (Gallese & Goldman, 1998, p. 495). It is in part because of this neural-somatic match to a self-performed action that the Other’s movement is understood as a goal-directed action. Thus the mirror neuron findings indicate some of the biological depth of empathy at the level of the passive association of the living bodies of self and other in embodied action.

Following Stein, this sort of empathic experience can be explicated further. In experiencing another as an embodied subject, we perceive the Other (1) as animated by his or her own fields of sensation; (2) as animated by general feelings of life or being in one’s living body (growth, development, aging, health and sickness, vigour and sluggishness, and so on); (3) as expressive of his or her own subjective experience; (4) as another centre of orientation in space; and (5) as capable of voluntary action (see **Toombs**, this volume, for discussion of these levels in the context of illness experience and clinical practice).

The empathic grasping of another as animated by his or her own fields of sensation Stein calls ‘sensual empathy’ or ‘sensing-in’. To take Stein’s example: ‘The hand resting on the table does not lie there like the book beside it. It “presses” against the table more or less strongly; it lies there limpid or stretched; and I “see” the sensations of pressure and tension’ (p. 54). So far this example includes only the first level of accomplishment in empathy — the emergence of the experience of another. The second level involves delving into the content of the Other’s experience. If this happens, then there is a movement from empathy as the passive association of our two lived bodies to empathy as the imaginative transposal of myself to the place of the Other: ‘my hand is moved (not in reality but “as if”) to the place of the foreign one. It is moved into it and occupies its position and attitude, now feeling its sensations, though not primordially [i.e., not in the original] and not as being its own . . . the foreign hand

is continually perceived as belonging to the foreign physical body so that the empathized sensations are continually brought into relief as foreign in contrast with our own sensations' (p. 54).

Clearly, for this kind of sensual empathy to be possible, one's own body and the Other's body must be of a similar type. What the limits of this type might be is an open and important question. Stein notes that 'empathy is quite successful with men's and children's hands which are very different from mine' (p. 54), and then raises the crucial point: 'The type "human physical body" does not define the limits of the range of my empathic objects, more exactly, of what can be given to me as a living body.' For example: 'Should I perhaps consider a dog's paw in comparison with my hand, I do not have a mere physical body, either, but a sensitive limb of a living body... I may sense-in pain when the animal is injured.' Nevertheless, 'the further I deviate from the type "man," the smaller does the number of possibilities of fulfillment become' (p. 55) (see also the contributions by **Gallese**, **Savage-Rumbaugh et al.**, and **Smuts**).

Interwoven with sensual empathy is the experience of the Other as animated by general feelings of life (health, vitality, sickness, and so on), and as expressive of subjective experience: 'we "see" shame "in" blushing, irritation in the furrowed brow, anger in the clenched fist' (p. 70). As blushing shows, the facial expression of feeling and emotion is a paradigm of these aspects of empathy (see **Cole**, this volume).<sup>8</sup> As Jonathan Cole has recently written:

The face involves an injunction not only to express, and to observe expressions, but to immerse oneself in what is expressed and to feel something of it oneself. Though complementary to body language, in this it may go beyond what is usually considered to be expressed through posture. Expressions actually help in constituting what is within. A face, therefore, is not only an expression of a self available for others to read, but to some extent the self is constituted in the face and developed, and experienced, in the interaction between faces (Cole, 1997, p. 482).

Another ingredient of empathy is the experience of the Other as being another centre of orientation in the space of the world. In general, one's experience of space and one's sense of self-identity are tied together: we perceive things to be arrayed around us, while we are 'here', at the centre or 'zero-point' of our orientation in space. This differentiation between 'here' and 'there' does not belong to space considered as a medium independent of one's body; it belongs to bodily space, to what philosophers call 'egocentric' space. When we perceive another, we perceive her as 'there' in relation to us 'here', and we grasp her as having her own egocentric space, defined by her own bodily movements. Furthermore, we perceptually grasp that her body is capable of voluntary movement (the third aspect in the list from Stein above). We do not experience another's movements as merely mechanical, but as alive and spontaneous. Neither sentience (having fields of sensation) nor spatial orientation (having an egocentric space) can be separated from voluntary movement in our empathic grasp of another. In empathetically experiencing another person as a sentient being capable of voluntary movement, we experience her as occupying her own 'here', in relation to which we stand 'there'.

---

[8] The 'Face' is one of the central motifs of the phenomenology of Emmanuel Levinas (1969), but to discuss the role this notion plays in his thought would take us too far afield (see **Zahavi**, this volume, and **Pitkin**, this volume).

Once again, this experience can remain at the first level of accomplishment — the emergence of experience, where it remains tacit and prereflective, a matter of passive association — or it can proceed to the second level — the fulfilling explication, where it unfolds according to the imaginative self-transposal to the place of the Other. This imaginative self-transposal presupposes the open intersubjectivity of consciousness discussed earlier. It enables us to gain a new spatial perspective on the world, that of the Other. At the same time, we continue to have (and must always have) our own centre of spatial orientation. Thus the open intersubjectivity of consciousness and its concrete articulation in empathy make it possible for us to comprehend an intersubjective field in which there is no one single zero-point or bodily centre of orientation. To put the point another way: the intersubjective openness of consciousness and empathy are the preconditions for our experience of inhabiting a common, intersubjective, spatial world. Empathy, as we have just seen, provides a viewpoint in which one's centre of orientation becomes one among others. Clearly, the space correlated to such a viewpoint cannot be one's own egocentric space, for that space is defined by one's own zero-point, whereas the new spatial perspective contains one's zero-point as simply one spatial point among many others.

This experiential grasp of intersubjective space is a condition of possibility for one's ability to experience one's own living body as a physical body like other physical things of the world. If one were confined to one's own first-person point of view, such that one had absolutely no empathic openness to others (an impossibility because of the open intersubjectivity of consciousness), and hence to how one would be experienced by another (empathy as the experience of myself as being an other for you), one would be incapable of grasping that one's own body is a physical object equivalent to the other physical things one perceives. A physical object is something that can stand before one in perception, but the living body, from an exclusively first-person point of view, cannot stand before one in this way. No matter how one turns, one's body is always 'here', at the zero-point, never 'there'; one cannot walk around it to behold it from all sides (see Merleau-Ponty, 1962, pp. 90–7). In general, one's body, as that by which one experiences a world, cannot show up as a fully present object in the world; it is always 'absently available' (Gallagher, 1986; Leder, 1990). Therefore, as long as we consider the living body simply from the first-person singular perspective, it seems like no other physical object, indeed like 'the strangest object' (Stein, 1964, p. 38), something radically 'incomplete' (p. 58). In Husserl's words: 'The same Body [*Leib*] which serves me as means for all my perception obstructs me in the perception of it itself and is a remarkably imperfectly constituted thing' (1989, p. 167). It is through empathy as the experience of oneself as an other for the alter-ego that one gains a viewpoint of one's own embodied being beyond the first-person singular perspective.

Stein elaborates this important point in terms of 'reiterated empathy'. In reiterated empathy, I see myself from your perspective. Stated more precisely, I empathetically grasp your empathic experience of me. As a result, I acquire a view of myself not simply as a physical thing, but as a physical-thing-empathetically-grasped-by-you-as-a-living-being. In other words, I do not merely experience myself as a sentient being 'from within', nor grasp myself as also a physical thing in the world; I experience myself as recognizably sentient 'from without', that is, from your perspective, the perspective of another. In this way, one's sense of self-identity, even at the most

fundamental levels of embodied agency, is inseparable from recognition by another, and from the ability to grasp that recognition empathetically.

Let us now explicitly link empathy back to the open intersubjectivity of consciousness discussed earlier. Empathy — like imagination, recollection, and reflection — can be described as a ‘self-displacing’ or ‘self-othering’ act (see Zahavi 1999, p. 150). Empathy involves a displacement or fission between my empathizing self and the empathized other; recollection between my present recollecting self and my past recollected self (whom I ‘see’ from the vantage point of the Other who is me now); imagination between myself imagining and myself imagined (whom I ‘see’ from the vantage point of the Other who is me imagining); and reflection between my reflecting self and the experiences I reflect upon. What such self-displacing experiences indicate is that, as **Natalie Depraz** puts it (this volume), the ego is structured or inhabited by many ‘inner splittings’ or ‘inner openings,’ openings that intrinsically involve otherness or alterity, and thus manifest the open intersubjectivity of consciousness. In Dan Zahavi’s words: ‘even if consciousness could turn its attention so completely toward itself that everything else were excluded, it would not escape the confrontation with Otherness’ (Zahavi 1999, p. 125). (For a moving meditation on the complex and intertwined dimensions of self-displacing experience in the context of illness experience, see **Varela**, this volume).

### VII: The Core Dyad Revisited

We have now reached the point where we can return to the Core Dyad, beginning with the phenomenological side.

#### *Empathy as the precondition of the science of consciousness*

If the phenomenological analysis of empathy and the open intersubjectivity of consciousness is on the right track, then it follows that the naturalistic perspective of cognitive science presupposes empathy as its condition of possibility, in particular the reciprocal empathy by which self and other are concretely co-determined. By this assertion I do not simply mean that cognitive science is an intersubjective enterprise that depends on the shared, pre-theoretic, lived experience of the scientists themselves. I mean something more radical, namely, that the very object of cognitive science — the embodied mind as a natural entity — is constituted as a scientific object through reciprocal or reiterated empathy in the human life-world.

Husserl himself makes this type of point about the human organism, indeed the whole of the realm of nature as conceived by empirical science. The basic idea is that the living body (or all of nature), understood as an objective entity governed by natural laws, is the intentional correlate of a certain mental attitude, the naturalistic attitude, and that this attitude presupposes (conceptually, epistemologically, and developmentally), the personalistic attitude, in which we relate to each other empathically as living bodily subjects. In Zahavi’s words: ‘It is through the Other that I learn to carry out an objectifying, ideative, and abstractive apprehension of my own body, which conceives it as a part of nature, as a mere complex of physiological organs embedded within and determined by causal relations in the world’ (1999, p. 161).

This point has important implications for the science and philosophy of consciousness. Consider what is often taken to be the central challenge faced by consciousness



studies, that of the so-called ‘hard problem’ of consciousness (Chalmers, 1996; 1997), also known as the problem of the ‘explanatory gap’ between consciousness and nature. Many philosophers have argued that there seems to be a gap between the objective, naturalistic facts of the world and the subjective facts of conscious experience. The hard problem is the conceptual and metaphysical problem of how to bridge this apparent gap. There are many critical things that can be said about the hard problem (see Thompson & Varela, forthcoming), but what I wish to point out here is that it depends for its very formulation on the premise that the embodied mind as a natural entity exists ‘out there’ independently of how we configure or constitute it as an object of knowledge through our reciprocal empathic understanding of one other as experiencing subjects. One way of formulating the hard problem is to ask: if we had a complete, canonical, objective, physicalist account of the natural world, including all the physical facts of the brain and the organism, would it conceptually or logically entail the subjective facts of consciousness? If this account would not entail these facts, then consciousness must be an additional, non-natural property of the world. One problem with this whole way of setting up the issue, however, is that it presupposes we can make sense of the very notion of a single, canonical, physicalist description of the world, which is highly doubtful, and that in arriving (or at any rate approaching) such a description, we are attaining a viewpoint that does not in any way presuppose our own cognition and lived experience. In other words, the hard problem seems to depend for its very formulation on the philosophical position known as transcendental or metaphysical realism. From the phenomenological perspective explored here, however — but also from the perspective of pragmatism à la Charles Saunders Peirce, William James, and John Dewey, as well as its contemporary inheritors such as Hilary Putnam (1999) — this transcendental or metaphysical realist position is the paradigm of a nonsensical or incoherent metaphysical viewpoint, for (among other problems) it fails to acknowledge its own reflexive dependence on the intersubjectivity and reciprocal empathy of the human life-world.

Another way to make this point, one which is phenomenological, but also resonates with William James’s thought (see Taylor, 1996), is to assert the primacy of the personalistic perspective over the naturalistic perspective. By this I mean that our relating to the world, including when we do science, always takes place within a matrix whose fundamental structure is I-You-It (this is reflected in linguistic communication: I am speaking to You about It) (Patocka, 1998, pp. 9–10). The hard problem gives epistemological and ontological precedence to the impersonal, seeing it as the foundation, but this puts an excessive emphasis on the third-person in the primordial structure of I–You–It in human understanding. What this extreme emphasis fails to take into account is that the mind as a scientific object has to be constituted as such from the personalistic perspective in the empathic co-determination of self and other.

The upshot of this line of thought with respect to the hard problem is that this problem should not be made the foundational problem for consciousness studies. The problem cannot be ‘How do we go from mind-independent nature to subjectivity and consciousness?’ because, to use the language of yet another philosophical tradition, that of Madhyamika Buddhism (**Wallace**, this volume), natural objects and properties are not intrinsically identifiable (*svalaksana*); they are identifiable only in relation to the ‘conceptual imputations’ of intersubjective experience.

*Empathy as an evolved, biological capacity of the human and other mammalian species*

It is important to realize that even though the embodied mind is constituted as a scientific object on the basis of the lived body and reciprocal empathy, it does not follow that the embodied mind and the lived body are two different entities, or belong to two different orders of reality. Rather, they are, we might say, two *aspects* of one single spatiotemporal individual.<sup>9</sup> Husserl himself wrote an intriguing statement to this effect in 1934 near the end of his life: ‘The lived body is at one with the physical body, membered thus and so... [it is] precisely organ and system of organs’ (1973, Vol. III, p. 643, as quoted by Welton, 1999, p. 51). It is precisely because the lived body is ‘at one’ with the organism that phenomenology needs to go outside itself to embrace other ways of knowing, natural science in particular. On the one hand, phenomenology is needed to understand the lived body as the primordial subject of lived experience. The lived body is the presupposition of our ability to know anything, in particular of our ability to know anything of the embodied mind as an object of biological and cognitive scientific study. As Merleau-Ponty says, it is not possible to comprehend the lived body without ‘abandoning the body as an object . . . and . . . going back to the body which I experience at this moment’ (1962, p. 75). On the other hand, without biology and cognitive science we can know nothing of the lived body as an organism in the vast web of life that is always anterior to ‘the body which I experience at this moment’ and that claims this body as its own. Accordingly, these two forms of understanding, phenomenology and cognitive science, should not be opposed, but must be joined together in a relationship of mutual illumination.

To move forward on this task two steps need to be taken. First, from the phenomenological side, phenomenology needs to give more attention to the development of methods for the careful and sustained examination of lived experience, methods that are intersubjective and open to the objective, empirically based descriptions of biology and cognitive science.<sup>10</sup> The gesture of mindful awareness on which such methods need to be based is a basic human capacity, but one whose exercise needs cultivation and practice (Depraz *et al.*, 2000). To recognize the necessity of cultivating this gesture is, I believe, the first step toward a ‘mature’ science of consciousness (Varela *et al.*, 1991; Donaldson, 1991; Varela & Shear, 1999).

Second, from the cognitive science side, we need to pursue an understanding of consciousness that is commensurate with phenomenology, and that looks to the phenomenology of the lived body and intersubjectivity as its complement. Among other things, this means incorporating phenomenological methods and data into the research protocols of cognitive science.

[9] Robert Hanna and I are currently engaged in a research project supported by the Center for Consciousness Studies at the University of Arizona, Tucson, titled ‘The Spontaneity of Consciousness: Neurophenomenology and Dual Aspect Metaphysics.’ One aim of this project is to spell out a coherent and defensible dual aspect metaphysics of consciousness using the metaphysical terminology and argumentation of analytic philosophy, but grounded on neurophenomenology (in the sense of Varela, 1996). See Hanna and Thompson (forthcoming).

[10] A common misconception is that phenomenological *reflection* is a form of solipsistic, first-person singular *introspection*. For one thing, phenomenological reflection is fundamentally different in its methods and aims from introspection. For another, as we have seen, phenomenology emphasizes that otherness or alterity, and intersubjective openness, are built into the very structure of consciousness: thus the first-person singular always already inhabits the first-person plural.

A growing number of researchers believe that cognitive science and phenomenology must complement each other in this fashion if there is to be a cognitively and ethically satisfying science of consciousness. Although a full exploration of this ethical dimension is beyond the scope of this article, in the final section I wish to explore in a tentative fashion a few aspects of this important dimension, having to do with the developmental possibilities of empathy as ethical responsiveness to the Other.

### VIII: From Intersubjectivity To Interbeing

One of the unique possibilities that human empathy affords is the development of non-egocentric or self-transcendent modes of consciousness. To understand these modes we need to explore the path that leads from intersubjectivity to what we can call, borrowing a term from the Vietnamese Buddhist teacher Thich Naht Hanh (1987), ‘interbeing’.

Pathways of self-transcendence, although long familiar to the world’s spiritual or wisdom traditions, have barely begun to be acknowledged by cognitive science. What needs to be realized, however, is that cognitive science, as a result of its own internal development, is now beginning to find itself in the position of being able to understand, in concepts commensurate with its own scientific approach to the mind, that there are such pathways and that they exemplify one of the most significant aspects of the human mind — its intersubjective and empathic openness.

#### *Emotion and value feeling*

Within Western moral philosophy there is a long tradition, going back to Immanuel Kant, that privileges reason over feeling: to act out of duties legislated by reason is thought to have greater moral worth than acting on the basis of feeling or sentiment. Yet as Frans de Waal observes, echoing David Hume: ‘Aid to others in need would never be internalized as a duty without the fellow-feeling [sympathy] that drives people to take an interest in one another. Moral sentiments came first; moral principles second’ (de Waal, 1996, p. 87). As discussed earlier, the precondition, both logically and evolutionarily, for moral sentiments such as sympathy and compassion, is cognitive empathy, and cognitive empathy is not a disembodied and affectless comprehension of the Other, but rather the feeling of being led by another’s experience, to use Stein’s formulation. Feeling, in this context, does not mean simply bodily sensation, but also value feeling or emotion. In Margaret Donaldson’s words: ‘how do “emotions” differ, if at all, from “feelings”? The crux is that emotions are our value feelings. They mark importance. We experience emotion only in regard to *that which matters*’ (Donaldson, 1991, p. 12).

According to this conception, emotions are a subclass of feelings, the value feelings. Emotions enact or constitute the world of values: ‘a new object realm is constituted in feeling. This is the world of values. In joy the subject has something joyous facing him, in fright something frightening, in fear something threatening’ (Stein 1964, p. 83). We have many feelings that are not emotions, such as hunger, fatigue, or pain. Such feelings are, of course, typically accompanied by emotions, such as anger or fear, which, as value feelings, reflect our evaluation of the situation in which we find ourselves. Emotions mark importance, and therefore involve conceptually

structured meaning and the evaluative stance of an intersubjective, personal self (Lazarus, 1991; Lazarus & Lazarus, 1994).

If it is feeling in the sense of bodily affect — especially proprioception and kinaesthesia — that makes one experience one's body as one's own (Sheets-Johnstone, 1999, pp. 41–87), then it is emotion or value feeling that makes one experientially aware of one's personal self. As Stein puts it: 'as it [the subject] feels it not only experiences objects, but it itself. It experiences emotions as coming from the "depths of its 'I'" . . . the "I" experienced in emotion has levels of various depths. These are revealed as emotions arise out of them . . . in feelings we experience ourselves not only as present, but also as constituted in such and such a way. They announce personal attributes to us' (1964, pp. 89–90).

Thus emotions, as value feelings, make possible the evaluative experience of oneself and the world, and therefore are the very precondition of moral perception, of being able to 'see' a situation morally before deliberating rationally about it. As Arne Johan Vetlesen (1994, p. 4) has recently argued:

We experience the objects of moral judgments through emotion. . . . Judgment presupposes perception in the sense that perception 'gives' judgment its object; we pass moral judgment on things that are already given, or disclosed, to us through acts of perception. . . . It is on this level, which logically precedes that of judgment . . . that we locate the emotions. Emotions anchor us to the *particular* moral circumstance, to the aspect of a situation that addresses us immediately, to the *here and now*. To 'see' the circumstance and to see oneself as addressed by it, and thus to be susceptible to the way a situation affects the weal and woe of others, in short, to identify a situation as carrying *moral significance* in the first place — all of this is required in order to enter the domain of the moral, and none of it would come about without the basic emotional faculty of empathy.

The issue that we need to consider now is the range of developmental possibilities open to us in empathy and value feeling, in particular the possibility of cultivating self-transcendent or non-egocentric value feelings.

### *Modes of mind in human development*

The developmental psychologist Margaret Donaldson, in her remarkable book, *Human Minds: An Exploration* has shown that the human potential for emotional development is as great as that for intellectual development, despite the present imbalance between the two in our culture (Donaldson, 1991). Her argument rests on a developmental map of the human mind, in which there are four main 'modes of mind', each made up of various mental components — perception, action, thought, and emotion — and each defined by its own 'locus of concern':

- The Point Mode  
Locus of concern: here and now  
Components: perception, action, thought, and emotion
- The Line Mode  
Locus of concern: there and then  
Components: thought and emotion
- The Construct Mode  
Locus of concern: somewhere/sometime (no specific place or time)  
Components: thought and emotion

- The Transcendent Mode
  - Locus of concern: nowhere (i.e., not in space and time)
  - Components: thought and emotion

These modes emerge progressively in human development, without one replacing the other; each is retained along the way and influences the others. Their progressive unfoldment involves a process of ‘opening out’ or ‘disembedding’ of the mental context: ‘to begin with, the mind functions in the context of its own totality and in the external context of people, things, and happenings. . . . Then step by step the unity breaks up and new ways of doing and experiencing become possible’ (1991, p. 17).

The point mode emerges first and is the only one available to the young infant (under eight months). In this mode the locus of concern is always the present moment, the directly apprehensible ‘here and now’. (This is not an extensionless point, but rather the kind of short span of duration that William James called the specious present.) Adult experience in the point mode is highly absorbed perception in the moment, in which past and future drop away, as can occur in losing oneself in a piece of music.

The line mode expands the locus of concern to include the personal past and the personal future, a development that apparently begins to occur at eight to ten months: ‘it is very likely that the baby, in the first half-year of life at least, while having a sort of “rolling” sense of movement from immediate past to immediate present to immediate future, has no sense of an extended past in which specific events can be located — and likewise no sense of a future filled with events yet to come’ (p. 54). The line mode, however, enables one to look forward as well as back, so that one can locate oneself in relation to a remembered past and a possible future.

The step from the line mode to the next mode, the construct mode, consists in a movement away from personal happenings toward the impersonal nature of things: ‘Instead of here/now or there/then the mind will next begin to concern itself with a locus conceived as somewhere/sometime or anywhere/anytime. Thus in the third mode we are no longer restricted to a consideration of episodes in our own experience — or even those we have heard about from others. We start to be actively and consciously concerned about the general nature of things’ (p. 80). This mode is called the construct mode because its context is not provided by perception, memory, or anticipation, but depends instead on a deliberate constructive act of imagination. It apparently begins to emerge around the age of three.

When thought and emotion occur on an equal footing in the construct mode, then Donaldson speaks of the *core construct mode*. In some construct mode activity, however, the intention is to think unemotionally or dispassionately, and hence thought predominates over emotion. This effort gives rise to what Donaldson calls the *intellectual construct mode*, in which one’s concern is with the nature of impersonal phenomena in space and time.

The next development is the movement from the intellectual construct mode to the *intellectual transcendent mode*, which Donaldson believes emerges around the age of nine. Whereas in the former case, the locus of concern is still bound to space and time, in the latter case, a ‘disembedding’ from this context has been achieved: ‘the fourth mode is “spaceless”. It needs no local habitation, present, remembered, foreseen or imagined. To speak paradoxically, we may say that the locus of concern is nowhere’

(p. 126). The paradigmatic mental activities of the intellectual transcendent mode are logic and mathematics. One's concern is no longer with spatiotemporal things themselves, but rather with the patterns of relationships into which things can enter.

We now come to the key question Donaldson raises in her book: do there exist, or could there exist, developments of the emotions that parallel those of the intellectual modes? Donaldson shows that such parallels do indeed exist, and she calls them the *value-sensing modes*, 'with the proviso that the values in question must transcend personal concerns' (p. 143).

### *The value-sensing modes*

There are two value-sensing modes — the *value-sensing construct mode* and the *value-sensing transcendent mode*: 'A value-sensing construct mode would be one where the main component of experience was an apprehension of transpersonal importance, powerfully felt, but where the functioning of the mode depended upon the support of the imagination . . . [In] the value-sensing transcendent mode . . . the need for a constructed context is gone, so that self-transcending values can now be experienced and responded to without the props provided by the working of the imagination (pp. 150–1).

As Donaldson goes on to say, 'once [the] defining features [of these two modes] have been recognized it is not hard to find evidence that they have indeed formed part of the repertoire of at least some human minds' (p. 152). The main body of evidence, some of which Donaldson discusses, comes from the world's contemplative wisdom traditions — Christian mysticism; Jewish Cabala; Sufism; Buddhism (Theravada, Mahayana, and Vajrayana); Advaita Vedanta; and Taoism and Neo-Confucianism.

One of her examples is St John of the Cross who, in explaining 'how to reach divine union quickly,' distinguishes between meditation, which involves focusing on an image, such as beautiful light, and contemplation, which involves emptying oneself of all images so as to attain a heightened, open, and receptive awareness of divine love. Donaldson comments: 'It is clear that meditation in John's sense belongs in the value-sensing construct mode. But in his view this mode, though useful, is strictly limited in its value for us since the divine reality far surpasses anything that we can know through the imagination' (p. 153). Proficiency in contemplation, on the other hand, belongs to the value-sensing transcendent mode: 'The locus of concern is certainly not in space–time. Concern centres rather on something conceived as infinite and eternal — something conceived also as having supreme value. Accordingly, the response evoked is one of deep emotion' (p. 154).

Another example, notable for its precision, can be found in the Tibetan Buddhist practice of *samatha* meditation, which aims at achieving clear and attentive stability of mind (Wallace, this volume; see also 1998; 1999). One of the main methods involves focusing the attention upon a mental image, while cultivating mindfulness (attention to the image without forgetfulness) and introspection (repeated checking of the quality of one's attentiveness). There are said to be nine distinct stages in the development of *samatha*, leading to firm and sustained attention upon the meditative object. These stages, given their image-based character, fall within the value-sensing construct mode. On the other hand,

[w]ith the full achievement of *Samatha*, one disengages the attention from the previous meditative object, and the entire continuum of one's attention is focused single-



pointedly, non-conceptually, and internally in the very nature of consciousness. . . . Only the aspects of sheer awareness, clarity, and joy of the mind appear, without the intrusion of any sensory objects (Wallace, 1999, p. 182).

Upon attaining *Samatha*, by focusing attention on the *sheer* clarity and the *sheer* cognizance of experience, one attends to the defining characteristics of consciousness alone, as opposed to the qualities of other *objects* of consciousness (p. 183).

This description clearly suggests a transition from the value-sensing construct mode to the value-sensing transcendent mode. (The point mode is also present, because *samatha* is sustained attention to the here and now.) Indeed, the final aim of *samatha* is ‘to realize the ultimate nature of awareness, free of all conceptual mediation and structuring, transcending even the concepts of existence and nonexistence. Such primordial awareness, known in this tradition as “the Buddha nature”, is said to be our essential nature, and it is the fathomless well-spring of intuitive wisdom, compassion, and power’ (p. 186).

### *Compassion*

The progression of the value-sensing modes amounts to an ‘opening out’ or ‘disembedding’ of the egocentric sense of self. Many wisdom traditions testify that the natural and spontaneous expression of this disembedding is compassion (see **Pitkin** and **Wallace**, this volume) or love (see **Steinbock**, this volume). Compassion is the heart of interbeing, and is the superlative expression of the human capacity for empathy.

Compassion is not merely an expression of nonegocentric value-feeling, one that can emerge only as a result of inward meditative disembedding, for it plays a guiding role in moving from one mode to another, in the expansion of the value-sensing repertoire. This is the reason that practices of compassion, benevolence, or love are emphasized so strongly right from the start in the practices of many wisdom traditions. The example most familiar to Westerners is Jesus’ injunction to ‘love thy neighbour’ and the Golden Rule ‘Do to others as you would have them do to you’. Another example is the central Confucian virtue of benevolence or human-heartedness (*ren*). When asked to explain benevolence, Confucius replied: ‘Do not impose on others what you yourself do not desire’ (*Analects* XII: 2); and he explained the method of benevolence as ‘the ability to take as analogy what is near at hand’ (VI: 30). ‘What is near at hand’ means oneself; hence the method consists ‘in using oneself as a measure to gauge others’ (IV: 15), that is, in placing oneself in the position of the Other and asking what one would like or dislike. In the Bodhisattva path of Mahayana Buddhism, as described by the eighth century Indian philosopher Shantideva (1997), the cultivation of *bodhicitta* or awakened mind, whose main attributes are wisdom and compassion, requires the specific contemplative practices of ‘meditation on the equality of self and other’ and ‘meditation on the exchange of self and other’. In the first, one aims to transcend the egocentric opposition of self and other through considering the sufferings of others as one’s own; in the second, one puts oneself in the place of others to understand how they feel, and how one appears in their eyes (see **Wallace**, this volume).<sup>11</sup>

Earlier, when discussing the phenomenological conception of empathy, I raised the question of the limits of empathy: how far can empathy radiate beyond the human case? Here the wisdom traditions give different answers. The extension of empathy

and compassion to the nonhuman world seems rather foreign to the Judaeo-Christian tradition (at least until recently), but is central to the Buddhist ideal of compassion for all sentient beings, and to the Neo-Confucian ideal of ‘forming one body with the universe’ (Tu, 1985; see also Leder, 1990, pp. 156–73).

The Neo-Confucian perspective, which arose as a synthesis of Confucianism, Buddhism and Taoism in the eleventh century CE in China, presents us with an example of the great openness that some have found in the human capacity for empathy and compassion. Neo-Confucianism builds on the philosophy of Mencius (371–289 BCE), who held that benevolence was innate to the human ‘heart-mind’ (*xin*):

When I say that all men have the mind which cannot bear to see the suffering of others, my meaning may be illustrated thus: Now, when men suddenly see a child about to fall into a well, they all have a feeling of alarm and distress, not to gain friendship with the child’s parents, nor to seek the praise of their neighbors and friends, nor because they dislike the reputation [of lack of humanity if they did not rescue the child]. From such a case, we see that a man without the feeling of commiseration is not a man. . . . The feeling of commiseration is the beginning of humanity (Chan, 1963, p. 65).

The Neo-Confucian philosopher, Wang Yang Ming (1472–1529), took this line of thought much further, saying that even when the mind of the ‘small man’

sees a child about to fall into a well, he cannot help a feeling of alarm and commiseration. This shows that his humanity forms one body with the child. It may be objected that the child belongs to the same species. Again, when he observes the pitiful cries and frightened appearance of birds and animals about to be slaughtered, he cannot help feeling an ‘inability to bear’ their suffering. This shows that his humanity forms one body with birds and animals. It may be objected that birds and animals are sentient beings as he is. But when he sees plants broken and destroyed, he cannot help a feeling of pity. This shows that his humanity forms one body with plants. It may be said that plants are living things as he is. Yet, even when he sees tiles and stones shattered and crushed, he cannot help a feeling of regret. This shows that his humanity forms one body with tiles and stones. This means that even the mind of the small man necessarily has the humanity that forms one body with all (Wang, 1963, p. 272).

Our culture today finds it hard to comprehend such a perspective. Yet the basic premise — that heart–mind and empathy are what distinguish human beings from other animals — is familiar from both cognitive ethology and phenomenology. Whereas cognitive ethology shows us the evolutionary roots of the human heart–mind, and phenomenology shows us the experiential structure of empathy, traditions such as Neo-Confucianism and Mahayana Buddhism challenge us to consider how we as members of a Western scientific culture might conceive the range of empathy in relation to what David Abram (1996) has called the more-than-human-world (see also Tucker & Berthrong, 1998). (In fact, a path to this idea can be found in the recent discussions by biologists and environmentalists of ‘biophilia’, the idea that humans have a biologically rooted affinity for the natural world. See Wilson, 1984; Kellert, 1997). Whatever we might decide, we cannot know the range of empathy and compassion short of investigating the human heart–mind with the full range of tools available — cognitive science, phenomenology, and the contemplative and meditative psychologies of the wisdom traditions.

### IX: The Next Step

It has become commonplace to say that we — contemporary Westerners — are children of the European Enlightenment and its discovery of reason and science. Some believe that we now need to abandon the Enlightenment, but the path followed in this article confirms Margaret Donaldson's suggestion that what we need is a second 'value-sensing enlightenment' to complement the first:

The very possibility of emotional development that is genuinely on a par with — as high as, level with — the development of reason is only seldom entertained. So long as this possibility is neglected, then if reason by itself is sensed as inadequate where else can one go but back? Thus there arises a regressive tendency, a desire to reject reason and all that was best in the Enlightenment, a yearning for some return to the mythic, the magical, the marvelous in old senses of these terms. This is very dangerous; but it has the advantage that it is altogether easier than trying to move forward into something genuinely new.

Now we have clearly seen that the cultivation of the advanced value-sensing modes is not of itself new. It has ancient roots. What *would* be new would be a culture where both kinds of enlightenment were respected and cultivated together. Is there any prospect that a new age of this kind might be dawning? (1991, p. 264).

The path that I have followed in this article leads not so much to a fixed conclusion, but to a clear sense of the next step that needs to be taken for such a second enlightenment: *we need to pursue a 'science of interbeing' (Varela, in press) that integrates the methods and findings of cognitive science, phenomenology, and the contemplative and meditative psychologies of the world's wisdom traditions.*

What steps can be taken now in this endeavour? Clearly the first step is to bring together cognitive scientists, phenomenological philosophers and psychologists, and contemplative practitioners to explore more thoroughly the common ground presented in this article. Especially important would be to establish a research programme for the following tasks:

- Clarification of the phenomenological method of investigating experience and exploration of its scope and limits in relation to cognitive science and contemplative practices.
- Development of more refined taxonomies of empathy and value-sensing.
- Development of experimental techniques to explore the neurocognitive dimensions of value sensing.
- Development of experimental techniques to assess the neurocognitive effects of value-sensing training (e.g., the Mahayana Buddhist meditations on equality of self and other, and exchange of self and other).
- Consideration of the implications of this research for biomedicine (e.g., psychoneuroimmunology), illness experience, and for empathy in the physician-patient relationship.
- Consideration of the implications of this research for our treatment of animals, especially in the context of scientific research on consciousness.

It is my hope that the articles collected in this volume will provide some of the resources for creating the research programs needed to pursue these tasks.

## References

- Abram, D. (1996), *The Spell of the Sensuous: Perception and Language in a More-Than-Human-World* (New York: Vintage Books).
- Adolphs, R. (1999), 'Social cognition and the human brain', *Trends in Cognitive Sciences*, 3, pp. 469–79.
- Allman, J.M. (1999), *Evolving Brains* (New York: Scientific American Library).
- Astington, J.W. (1993), *The Child's Discovery of the Mind* (Cambridge, MA: Harvard University Press).
- Beer, R.D. (2000), 'Dynamical approaches to cognitive science', *Trends in Cognitive Sciences*, 4, pp. 91–9.
- Carruthers, P. & Smith, P.K. (ed. 1996), *Theories of Theories of Mind* (Cambridge: Cambridge University Press).
- Chalmers, D.J. (1996), *The Conscious Mind: In Search of a Fundamental Theory* (New York: Oxford University Press).
- Chalmers, D.J. (1997), 'Moving forward on the problem of consciousness', *Journal of Consciousness Studies*, 4 (1), pp. 3–46.
- Chan, W-T. (1963), *A Source Book in Chinese Philosophy* (Princeton, NJ: Princeton University Press).
- Clark, A. (1997), *Being There: Putting Brain, Body, and World Together Again* (Cambridge, MA: The MIT Press/A Bradford Book).
- Clark, A. (1999), 'An embodied cognitive science?', *Trends in Cognitive Sciences*, 3, pp. 345–51.
- Confucius (1979), *The Analects*. trans. D.C. Lau. (Harmondsworth, Middlesex: Penguin Books).
- Corballis, M.C. (1998), 'Cerebral asymmetry: motoring on', *Trends in Cognitive Sciences* 2, pp. 152-157.
- Cole, J. (1997), 'On being "faceless": Selfhood and facial embodiment', *Journal of Consciousness Studies*, 4 (5–6), pp. 467–84.
- Damasio, A.R. (1994), *Descartes' Error. Emotion, Reason, and the Human Brain* (New York: Grosset/Putnam).
- Damasio, A.R. (1999), *The Feeling of What Happens: Body and Emotion in the Making of Consciousness* (New York: Harcourt Brace).
- Davies, M. & Stone, T. (ed. 1995a), *Folk Psychology. The Theory of Mind Debate* (Oxford: Blackwell).
- Davies, M. & Stone, T. (ed. 1995b), *Mental Simulations: Evaluations and Applications* (Oxford: Blackwell).
- Depraz, N. (1995), *Transcendence et incarnation: le statut de l'intersubjectivité comme altérité à soi chez Husserl* (Paris: Librairie Philosophique J. Vrin).
- Depraz, N., Varela, F.J. & Vermersch, P. (2000), 'The gesture of awareness: an account of its structural dynamics', in *Investigating Phenomenal Consciousness*, ed. Max Velmans (Amsterdam/Philadelphia: John Benjamins Publishing).
- Donald, M. (1991), *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition* (Cambridge, MA: Harvard University Press).
- Donaldson, M. (1991), *Human Minds: An Exploration* (London: Penguin Books).
- Freeman, W. (1999) *How Brains Make Their Minds* (London: Wiedenfeld & Nicholson).
- Gallagher, S. (1986), 'Lived body and environment', *Research in Phenomenology*, 16, pp. 139–70.
- Gallagher, S. (1997), 'Mutual enlightenment: recent phenomenology and cognitive science', *Journal of Consciousness Studies*, 4 (3), pp. 195–214.
- Gallagher, S. & Meltzoff, A. (1996), 'The earliest sense of self and others: Merleau-Ponty and recent developmental studies', *Philosophical Psychology*, 9, pp. 211–33.
- Gallese, V. & Goldman, A. (1998), 'Mirror neurons and the simulation theory of mind-reading', *Trends in Cognitive Sciences*, 2, pp. 493–501.
- Gallup, Jr., G. (1998), 'Can animals empathize? Yes', *Scientific American*, 9, pp. 66–71.
- Gerhart, J. & Kirschner, M. (1997), *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variability and Evolutionary Adaptability* (Malden, MA: Blackwell Science Publications).
- Goldman, A. (1993), *Philosophical Applications of Cognitive Science* (Boulder, CO: Westview Press).
- Goldman, A. (1995a), 'Empathy, mind, and morals', in Davies. & Stone (1995a).
- Goldman, A. (1995b), 'In defense of the simulation theory', in Davies & Stone (1995b).
- Gordon, R.M. (1996), "'Radical" simulationism', in Carruthers & Smith (1996).
- Hamrick, W.S. (2000), 'Empathy, cognitive science, and literary imagination,' *Journal of the British Society for Phenomenology*, 31, pp. 116–30.
- Hanna, R. & Thompson, E. (forthcoming), 'The spontaneity of consciousness'.
- Husserl, E. (1960), *Cartesian Meditations*, trans. Dorian Cairns (The Hague: Martinus Nijhoff).
- Husserl, E. (1973), *Zur Phänomenologie der Intersubjektivität I, II, & III. (Husserliana Vols. 13, 14, & 15.)* (The Hague: Martinus Nijhoff).
- Husserl, E. (1989), *Ideas Pertaining to A Pure Phenomenology and to a Phenomenological Philosophy, Second Book*, trans. R. Rojcewicz & A. Schuwer (Dordrecht: Kluwer Academic Publishers).
- Karmiloff-Smith, A. (1998), 'Development itself is the key to understanding developmental disorders', *Trends in Cognitive Sciences*, 2, pp. 389–98.
- Kellert, S.R. (1997), *Kinship to Mastery: Biophilia in Human Evolution and Development* (Washington, DC: Island Press/Shearwater Books).

- Kelso, J.A.S. (1995), *Dynamic Patterns: The Self-Organization of Brain and Behavior* (Cambridge, MA: The MIT Press).
- Lazarus, R. (1991), *Emotion and Adaptation* (New York: Oxford University Press).
- Lazarus, R. & Lazarus, B.N. (1994), *Passion and Reason: Making Sense of Our Emotions* (New York: Oxford University Press).
- Le Van Quyen, M., Adam, C., Lachaux, J-P., Martinerie, J., Baulac, M., Renault, B. & Varela, F.J. (1997), 'Temporal patterns in human epileptic activity are modulated by perceptual discriminations', *NeuroReport*, **8**, pp. 1703–10.
- Leder, D. (1990), *The Absent Body* (Chicago: University of Chicago Press).
- Levinas, E. (1969) *Totality and Infinity: An Essay on Exteriority*, trans. Alphonso Lingis (Pittsburgh, PA: Dusquesne University Press).
- McGeer, V. (1996), 'Is "self-knowledge" an empirical problem? Renegotiating the space of philosophical explanation', *Journal of Philosophy*, **93**, pp. 483–515.
- McGeer, V. (1999), 'Rethinking "mindreading": philosophical, psychological, and neurological implications of a regulative account of our "folk-psychological" capacities', Project Description, McDonnell Project in Philosophy and the Neurosciences. <http://www.sfu.ca/neurophilosophy>.
- Meany, M.J., Diorio, J., Francis, D., Widdowson, J., LaPlante, P., Caldji, C., Sharma, S., Seckl, J.R. & Plotsky, P.M. (1996), 'Early environmental regulation of forebrain glucocorticoid receptor gene expression: implications for adrenocortical responses to stress', *Developmental Neuroscience*, **18**, pp. 49–72.
- Meltzoff, A. & Moore, M.K. (1994), 'Imitation, memory, and the representation of persons', *Infant Behavior and Development*, **15**, pp. 479–505.
- Meltzoff, A. & Moore, M.K. (1999), 'Infant intersubjectivity: broadening the dialogue to include imitation, identity and intention', in *Intersubjective Communication and Emotion in Early Ontogeny*, ed. S. Braten (Cambridge: Cambridge University Press).
- Merleau-Ponty, M. (1962), *Phenomenology of Perception*, trans. Colin Smith (London: Routledge).
- O'Regan, J.K. & Noë, A. (in press), 'A sensorimotor account of vision and visual consciousness', *Behavioral and Brain Sciences*.
- Panksepp, J. (1998a), *Affective Neuroscience. The Foundations of Human and Animal Emotions* (New York: Oxford University Press).
- Panksepp, J. (1998b), 'The periconscious substrates of consciousness: Affective states and the evolutionary origins of self', *Journal of Consciousness Studies*, **5** (5–6), pp. 566–82.
- Patocka, J. (1998), *Body, Community, Language, World*, trans E. Kohak (Chicago & La Salle, IL: Open Court Publishing).
- Petit, J-L. (1999), 'Constitution by movement: Husserl in light of recent neurobiological findings', in Petitot *et al.* (1999).
- Petitot, J., Varela, F.J., Pachoud, B. & Roy, J-M. (ed. 1999), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* (Stanford, CA: Stanford University Press).
- Povinelli, D.J. (1998), 'Can animals empathize? Maybe not', *Scientific American*, pp. 67–75.
- Povinelli, D.J. & Preuss, T.M. (1995), 'Theory of mind: evolutionary history of a cognitive specialization', *Trends in Neurosciences*, **18**, pp. 418–24.
- Putnam, H. (1999), *The Threefold Chord: Mind, Body, and World*. New York: Columbia University Press.
- Quartz, S. (1999), 'The constructivist brain', *Trends in Cognitive Sciences*, **3**, pp. 48–57.
- Quartz, S. & Sejnowski, T.J. (1997), 'The neural basis of cognitive development: a constructivist manifesto', *Behavioral and Brain Sciences*, **20**, pp. 537–96.
- Rizzolatti, G. & Arbib, M. (1998), 'Language within our grasp', *Trends in Neurosciences*, **21**, pp. 188–94.
- Rizzolatti, G., Riggio, L. & Sheliga, B.M. (1994), 'Space and selective attention', in *Attention and Performance IV. Conscious and Nonconscious Information Processing*, ed. C. Umiltà & M. Moscovitch (Cambridge, MA: The MIT Press).
- Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (1997), 'The space around us', *Science*, **277**, pp. 190–1.
- Shantideva (1997), *The Way of the Bodhisattva*, trans. the Padmakara Translation Group. (Boston, MA: Shambala Press).
- Sheets-Johnstone, M. (1991), *The Primacy of Movement* (Amsterdam/Philadelphia: John Benjamins Publishing Company).
- Sober, E. & Wilson, S.W. (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior* (Cambridge, MA: Harvard University Press).
- Stein, E. (1964), *On the Problem of Empathy*, trans. Waltraut Stein (The Hague: Martinus Nijhoff).
- Steinbock, A. (1995), *Home and Beyond: Generative Phenomenology After Husserl* (Evanston, IL: Northwestern University Press).
- Taylor, E. (1996), *William James: On Consciousness Beyond the Margin* (Princeton, NJ: Princeton University Press).
- Thompson, E. (1995), *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception*. (London: Routledge Press).



- Thompson, E. (1999), 'Empathy and the affective mind-brain: the importance of affective neuroscience for understanding everyday mind-reading', Project Description, McDonnell Project in the Philosophy and the Neurosciences. .
- Thompson, E. & Varela, F.J. (in press) 'Radical embodiment: neural dynamics and conscious experience', *Trends in Cognitive Sciences*.
- Thompson, E. & Varela, F.J. (forthcoming), *Why the Mind Isn't in the Head* (Cambridge, MA: Harvard University Press).
- Thompson, E., Palacios, A. & Varela, F.J. (1992) 'Ways of coloring: comparative color vision as a case study for cognitive science', *Behavioral and Brain Sciences*, **15**, pp. 1–26.
- Tu, W-M.. (1985), *Confucian Thought: Selfhood as Creative Transformation* (Albany, NY: State University of New York Press).
- Tucker, M.E. & Berthrong, J. (ed. 1998), *Confucianism and Ecology: The Interrelation of Heaven, Earth, and Humans* (Cambridge, MA: Harvard University Press).
- Varela, F.J. (1996), 'Neurophenomenology: a methodological remedy for the hard problem', *Journal of Consciousness Studies*, **3** (4), pp. 330–50.
- Varela, F.J. (in press ), 'Steps to a science of interbeing: unfolding the dharma implicit in modern cognitive science', in *The Psychology of Awakening*, ed. S. Batchelor, G. Claxton & G. Watson.
- Varela, F.J. & Shear, J. (ed. 1999), *The View from Within: First-Person Approaches to the Study of Consciousness* (Thorverton, UK: Imprint Academic). Also published as *Journal of Consciousness Studies* **6** (2–3), 1999.
- Varela, F.J., Thompson, E., & Rosch, E. (1991), *The Embodied Mind: Cognitive Science and Human Experience* (Cambridge, MA: The MIT Press).
- Vetlesen, A.J. (1994), *Perception, Empathy, and Judgment: An Inquiry into the Preconditions of Moral Performance* (University Park, PA: Pennsylvania State University Press).
- de Waal, F. (1996), *Good Natured: The Origins of Right and Wrong in Humans and Other Animals* (Cambridge, MA: Harvard University Press).
- Wallace, B.A. (1998), *The Bridge of Quiescence: Experiencing Tibetan Buddhist Meditation* (Chicago and LaSalle, IL: Open Court).
- Wallace, B.A. (1999), 'The Buddhist tradition of *Samatha*: Methods for refining and examining consciousness', in Varela & Shear (1999). Also published in *Journal of Consciousness Studies*, **6** (2–3), pp. 175–87.
- Wang, Y-M. (1963), *Instructions for Practical Living and Other Neo-Confucian Writings*, trans. by Wing-Tsit Chan (Princeton: Princeton University Press).
- Watt, D.F. (1998), 'Emotion and consciousness: implications of affective neuroscience for extended reticular thalamic activating system theories of consciousness', <http://server.philvt.edu/assc/watt/default.htm>.
- Welton, D. (1999), 'Soft, smooth hands: Husserl's phenomenology of the lived body,' in *The Body*, ed. Donn Welton (Oxford: Basil Blackwell).
- Wilson, E.O. (1984), *Biophilia* (Cambridge, MA: Harvard University Press).
- Zahavi, D. (1996), 'Husserl's intersubjective transformation of transcendental philosophy', *Journal of the British Society for Phenomenology*, **27**, pp. 228–45.
- Zahavi, D. (1997), 'Horizontal intentionality and transcendental intersubjectivity', *Tijdschrift Voor Filosofie*, **59**, pp. 304–21.
- Zahavi, D. (1999), *Self-Awareness and Alterity. A Phenomenological Investigation* (Evanston, IL: Northwestern University Press).