Mind in practice : a pragmatic and interdisciplinary account of intersubjectivity

Bruin, L.C. de

Epilogue

Some Consequences of Pragmatism

I have a pragmatic interest in seeing how philosophy can address issues that are not purely philosophical (at least not as purely philosophical as defined in what is, in my opinion, an overly technical and narrow sense of philosophy in the 20th century).

- Shaun Gallagher

The story so far...

The first two chapters of this book dealt with the internal problems of theory theory (TT) and simulation theory (ST) approaches to intersubjectivity. I have argued that, in the first place, both TT and ST fail to capture the interactive and relational *phenomenology* of our everyday encounters with other minds. Its proponents often parry this objection by going 'underground', arguing that the processes they postulate should be understood as being operative at the sub-personal (neurobiological or cognitive) level. In doing so, however, they implicitly seem to acknowledge that mindreading fails as an adequate characterization of intersubjectivity at the personal level. Moreover, it is questionable whether it makes sense to apply concepts at the sub-personal level that were originally coined at the personal level. Secondly, both TT and ST face serious difficulties when it comes to explaining how we are able to navigate our social environments *in the adaptive and context-sensitive way we do*. Instead, they tend to 'solve' this problem with an appeal to innateness. This, however, seems to be nothing more than an excuse for a lack of real understanding.

Chapter 3 further investigated the deeper assumptions that underlie TT and ST approaches to intersubjectivity. By accepting the problem of the other mind as a genuine

problem, both TT and ST buy into a questionable picture of intersubjectivity: one that suggests a conception of the mind as a passive spectator, and takes for granted a phenomenology of uncertainty. This fosters the idea that our interactions with others require some kind of intervention between our initial observation of others and our final reaction towards them. At the center of this picture is the assumption that we are normally at a theoretical remove from other minds, and have to adopt a third-person stance towards them for the purposes of prediction, explanation and control. It is in this sense that TT and ST promote a *theoretical* approach to intersubjectivity.

By contrast, I have proposed an account of intersubjectivity that is very much *practice-oriented* (chapter 4 and 5). According to my proposal, our common sense encounters with others can be explained as being facilitated by three types of second-person practices: (i) *embodied practices*, allowing us to employ various innate or early developing capacities that provide a basic form of social understanding; (ii) *embedded practices* of joint attention, enabling an understanding of others within a broader social and pragmatic context; and (iii) *narrative practices*, providing us with stories in order to further fine-tune and sophisticate our intersubjective interactions.

These second-person practices to a large extent *obviate* the cognitively and conceptually demanding mindreading procedures postulated by TT and ST, and severely restrict the scope of intersubjective understanding in terms of mental states such as beliefs and desires. They provide us with a satisfactory explanation of our social engagements that is at the same time far more parsimonious. From a pragmatic perspective, the problem with TT and ST explanations of social interaction is that they come with severe *developmental constraints*, such as mental concept mastery, inferential abilities and analogical reasoning. If we want to take these constraints seriously (and I have argued that we should), then we cannot but conclude that young children fail to meet the necessary requirements to pursue a career in mindreading.

Another advantage of my proposal is its ability to address the TT and ST troubles with context-sensitivity. It simply points to the strong orientation towards the *concrete* and the *particular* that is characteristic for most of our interactions with other minds. At the same time, however, this presupposes a radically different notion of the mind: not as a passive, static spectator, but instead as an enactive, embodied and embedded *participant.*

All of this results in an enactive approach to intersubjectivity that increasingly works towards a *trivialization* of the problem of the other mind. It does so by challenging the four

assumptions that are implicitly taken for granted by TT and ST approaches to intersubjectivity (prologue, p.15-16), arguing instead that: (i) our dealings with others are not essentially problematic; (ii) the conception of the mind that is at the basis of such a conviction should be rejected, (iii) our everyday social encounters do not *by default* require theoretical interventions, because (iv) they are firmly grounded in *second-person interactions* that can be understood in terms of direct perception-action couplings.

But does it make a difference?

Of course, the litmus test for a pragmatic second-person approach to intersubjectivity is whether it actually *makes a difference*, not only for our interpretation of the processes that facilitate social interaction, but also for empirical studies in this area. It should not come as a surprise that I think this is indeed the case. Although it is certainly not easy to investigate our social skills from a perspective that is truly second-person, doing so *will* and in fact already *has* paid off substantially. Let me give one example of a recent EEG experiment by Tognoli et al. (2007) to illustrate this claim.

Electroencephalography or EEG is a neuro-imaging technique in which a large number of electrodes are placed onto the head in order to record the electrical activity that is produced by the firings of neurons within the brain. In many EEG studies on (aspects of) intersubjectivity, lonely subjects are passively sitting in a chair while facing a monitor screen, and they are asked to perform all kinds of computer-based tasks by endlessly clicking yes or no buttons with their fingers. There is no genuine *second-person interaction* involved in these experiments. Tognoli et al. (2007), however, managed to drastically enhance and improve the set-up of their EEG experiment by placing two subjects over against each other and letting them *interact*. Initially, the subjects were asked to rhythmically wag their fingers at their own preferential pace, but they were prevented from seeing each others' hands. Then the barrier placed between them was removed, so they could see each other while continuing to wag their fingers. When subjects were allowed to see one each other's fingers moving, they sometimes adjusted their own movements and synchronized, and sometimes they did not, behaving in an independent manner. By recording, measuring and analyzing both behavior and brain activity in these interacting subjects simultaneously, the experimenters found a so-called 'phi complex', a brain rhythm

operating at 10 Hz and located above the right centro-parietal cortex. According to Kelso, one of the principal investigators, these findings suggested that a unique pattern can be seen in the brains of two people interacting and that these brain activities distinguish independence from cooperation: 'This new brain rhythm that we have discovered and termed the "phi complex" actually distinguishes when you're socially interacting and when you're not'.

This claim is probably highly exaggerated, and I certainly do not wish to defend my argument on the basis of the *specific* findings presented in this experiment. My only point is that this study shows that pursuing a second-person approach in scientific experimentation (and thus adopting a different conception of intersubjectivity) *does* make a difference to what we will find, even at the sub-personal level. And since such an approach does much more justice to the phenomenology of our everyday social interaction, I am convinced that it is worth pursuing.

There is another way in which my pragmatic approach makes a difference to scientific research. By stressing the irreducibility between the various practices presented in this book (the 'levels of explanation'), it aims to discourage an interpretation of our intersubjective skills solely in terms of neurobiological mechanisms. Instead, it suggests that each level of practice might contribute to a more complete understanding of intersubjectivity. Again, an example might be helpful.

Autism spectrum disorders (ASD) are characterized by various social and communicative deficits, such as problems with imitation, empathy and language use, but also by nonsocial symptoms, such as an obsessive concern for sameness, preoccupation with objects or parts of objects, echolalia, and a variety of sensory and motor behaviors such as oversensitivity to stimuli and repetitious and odd movements (see Happé 1995, 113ff). Elucidating the underlying neural bases of ASD has been a challenge because the manifestations of this disorder vary in severity (low and high-functioning) as well as expression (Autistic Disorder, Asperger's Disorder, and Pervasive Developmental Disorder-Not Otherwise Specified).

Nevertheless, it has been proposed that a dysfunctional mirror neuron system (MNS) early in development might be responsible for the cascade of impairments that fall under the heading of ASD (Williams et al. 2001). Despite the fact that the heterogeneity of the ASD condition seems to argue against a single cause, the idea behind this proposal is that ASD is *primarily* a failure of empathy, which in turn depends on the kind of inner imitation

that is generated by the MNS (Carr et al. 2003). Now, on the one hand, studies by Dapretto et al. (2005) and Obermann (2005) suggest that there is some evidence for abnormal MNS functioning during action and observation imitation in individuals with ASD. On the other hand, however, there are also critical voices arguing that the MNS approach to intersubjectivity is seriously *flawed* (cf. Hickok 2009). Lingnau et al. (2009) have even suggested that there might be no experimental evidence for the existence of a human MNS *whatsoever.*

Of course, such a dispute by itself is not an argument against reducibility or the role of the MNS in explanations of ASD. But it does indicate that there is a serious problem with the idea that intersubjectivity = empathy = imitation = MNS, and the subsequent argument that problems with intersubjectivity therefore have to be traced back to MNS dysfunction. The problem is that the search for so-called 'prime movers' at the sub-personal level often implicitly results in a very *impoverished* phenomenology at the personal level, and an unjustified *simplification* of something much more complex. When such an impoverished phenomenology is used as a starting point for scientific experimentation, it yields results that are rather different compared to a much richer phenomenology that tries to capture our natural, second-person ways of dealing with other people. Klin et al. (2003), for example, has pointed out that there are remarkable differences in findings between ASD studies in which the participants were presented with static pictures of faces (e.g., Van der Geest et al. 2002), and one in which they were shown much more dynamic depictions of social interactions (by means of video). They argued that in such more 'spontaneous' situations, the 'deviation from normative facescanning patterns in autism seems to be magnified' (p.346). In other words, the attempt to replicate a more naturalistic social situation eventually gave the investigators more insight in the severity of ASD. It is highly likely that more attention for the various second-person practices in which ASD symptoms manifest themselves in the end also provides us with a *fuller* explanation of what goes wrong in ASD. Gallagher (2004) has suggested that an integrative account of ASD therefore needs to take into consideration not only possible neurobiological problems, but also dysfunctional behavior at the level of primary and secondary intersubjectivity (fig. E.1)

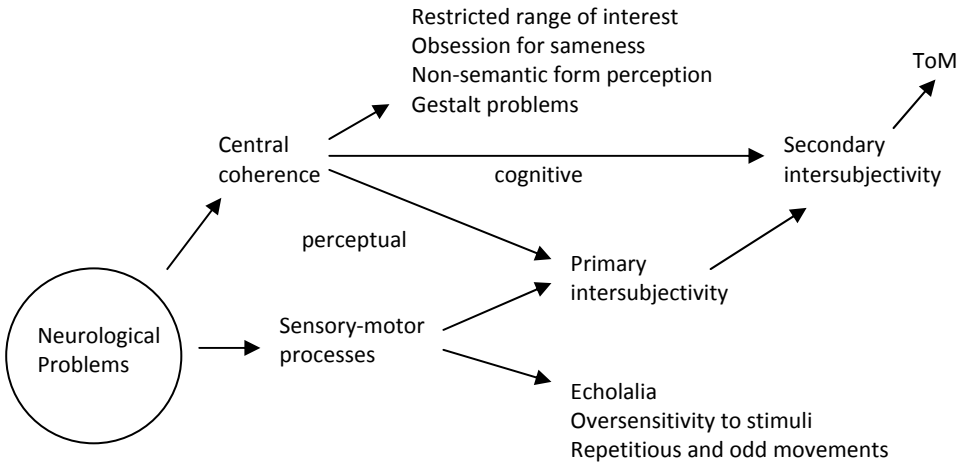

Fig. E.1 A fuller picture of what can go wrong in ASD (Gallagher 2004)

Pragmatism versus reductionism

The obsession with 'prime movers' or 'real causes' that is characteristic for certain scientists clashes with a second-person approach to intersubjectivity, but it fits very well with a particular philosophical (Cartesian) paradigm. In this final section, I want to briefly comment on this paradigm and propose a view that is more in line with the pragmatic view endorsed throughout this book.

In previous chapters I have introduced a conception of the mind as a coupled complex system of brain, body and environment - one that emerges as the result of continuous interactions with other minds. I have explained these interactions in terms of embodied, embedded and narrative practices, and argued that these practices are not reducible to each other or to the sub-personal (neurobiological) processes that structure and shape them.

The idea of an emerging mind and the assumption that the practices in which it participates have their own (relative) 'autonomy' and explanatory pay-off is very problematic according to some philosophers. The main problem is this: if we grant the emerging mind new causal powers at each stage of development in which it grows in complexity, then how can we explain mental or 'downward' causation - the causal influence of a whole on its own micro-constituents?

Kim (1999), for example, argues that this kind of downward causation is either *otiose* or violates the 'causal closure of the physical' when understood to happen diachronically' (pp.28-33). This is because it relies on three principles that are mutually incompatible:

i) The physical realization principle: every emergent event or property M must be realized by (or determined by, or supervenient on) some physical event or property P (its 'emergence base').
ii) The causal inheritance principle: If M is instantiated on a given occasion by being realized by P, then the causal powers of this instance of M are identical with (or a subset of) the causal powers of P.
iii) The principle of the causal closure of the physical domain: any physical event that has a cause at time t has a physical cause at t. Hence, 'if we trace the causal ancestry of a physical event, we need never to go outside the physical domain' (Kim 1993, p.280).

In combination, these principles confront the pragmatist who is committed to emergence and believes in mental causation with a pressing dilemma: either mental causation is otiose, because the putative causal power of the emergent is preempted by the causal power of the physical elements on which the emergent is based, or mental causation violates the principle that the physical domain is causally closed.

Recently, however, Thompson (2008) has advanced a number of arguments against some of the metaphysical assumptions that underlie Kim's picture of mental causation. In what follows, I briefly discuss these arguments to the extent that they provide support for my own pragmatic proposal.

In the first place, it is important to notice that Kim accepts a 'layered model of reality', according to which the world is composed as a hierarchically stratified structure of levels of physical entities or particulars and their characteristic properties. Its bottom level consists of whatever physics is going to tell us are the most basic physical particles out of which all matter is composed (e.g. electrons, neutrons or quarks). And these objects are in turn characterized by certain fundamental physical properties and relations (e.g. mass, spin, or charm). Against this background, the challenge has become to explain how, as Kim (2000) puts it, 'it is possible for the mind to exercise its causal powers in a world that is fundamentally physical' (p.30).

Thompson criticizes this worldview because a mereologically ordered hierarchy grounded on a base level of particulars is a metaphysical picture projected onto science, whereas the image science projects is of networks of processes at various spatiotemporal scales, with no base-level particulars that 'upwardly' determine everything. Contemporary science does not articulate a conception of nature as grounded in a basic set of *particulars*, but instead refers to *fields* and *processes*. There is no bottom level of basic particulars with intrinsic causal properties that upwardly determine everything else. Everything is process all the way 'down' and all the way 'up', and processes are irreducibly relational - they exist only in patterns, networks, organizations, configurations, or webs (cf. Campbell and Bickhard 2002, Hattiangadi 2005).

Thompson argues that Kim's picture of mental causation presupposes an 'elementary-particle-version of Cartesian substance metaphysics' that allows for part/whole reductionism. For the part/whole reductionist, 'down' and 'up' describe more and less fundamental levels of reality. Higher levels are realized by and determined by lower levels, in accordance with the layered model of reality as described in the previous section. This idea finds its expression in the principle of physical realization: every mental property M must be realized by a physical property P.

According to Thompson's 'process view' of the world, however, 'up and 'down' are context-relative terms used to describe phenomena of various scales and complexity. There is no base level of elementary entities to serve as the ultimate 'emergence base' on which to ground everything. As Thompson (2007) puts it, 'phenomena at all scales are not entities or substances but relatively stable processes, and since processes achieve stability at different levels of complexity, while still interacting with processes at other levels, all are equally real and none has absolute ontological primacy' (p.441). Such a process view obviously fits well with the pragmatic view propounded in this book.

What about the third principle, i.e. the assumption of the causal closure of the physical domain? Thompson observes that, in the first place, it is unclear what is precisely meant by 'physical' in this respect. Proponents of physicalism usually talk freely of 'mental' and 'physical' properties, as if these terms track two clearly contrasting classes of entities that can be compared experimentally (cf. Strawson 2006). However, the very idea that mental properties *qua* mental can be distinguished from and systematically contrasted to physical properties in a meaningful way, as for example Mclaughlin (1994) would have it, is deeply suspect. It is simply not clear what 'physical' includes and excludes, and it is also hard to

see how one could go about answering this question short of having a complete and true physics (cf. Montero 1999, 2001). Moreover, if we construe the principle of causal closure more narrowly as to mean the causal closure of the microphysical domain, then the principle is not obviously true and may even be false or incoherent (Dupre 1993, Hattiangadi 2005).

Although it is difficult to make intelligible the idea that complex systems are *causally* closed, Thomspon argues that there is a different way in which complex systems can said to be closed. Complex systems are closed in the sense that they are *autonomous*.[110] An autonomous system consists of a network of processes, in which (i) the processes recursively depend on each other for their generation and their realization as a network; and (ii) the processes constitute the system as a unity in whatever domain they exist.[111] According to Varela (1979), an autonomous system can be defined as a system that has *organizational closure* and *operational closure* (pp.55-60). The term 'closure' does not mean that the system is materially and energetically closed to the outside world (which of course is impossible). On the contrary, autonomous systems are thermodynamically far from equilibrium systems, which incessantly exchange matter and energy with their surroundings. 'Organizational closure' describes the self-referential (circular and recursive) network of relations that defines the system as a unity. At any given instant or moment, this self-referential network must be maintained, otherwise the system is no longer autonomous and no longer viable in whatever domain it exists. 'Operational closure' describes the recursive, re-entrant, and recurrent dynamics of the system. The system changes state on the basis of its self-organizing dynamics (in coupling with an environment), and the product of its activity is always further self-organized activity within the system (unless its operational closure is disrupted and it disintegrates).

What is important about complex systems is that they are also sufficiently open to allow for *emergent properties* with *new* causal powers. Emergent properties 'arise' out of more basic properties and yet they are 'novel' or 'irreducible' with respect to them.[112] Even

[110] 'Autonomous' literally means 'self-governing', or 'conforming to its own law'.

[111] The paradigmatic example of an emerging, self-organizing non-equilibrium system is a living cell. The constituent processes in this case are chemical; their recursive interdependence takes the form of a self-producing, metabolic network that also produces its own membrane; and this network constitutes the system as a unity in the biochemical domain. This kind of autonomy and self-production in the biochemical domain is known as *autopoiesis* (Maturana and Varela 1980).

[112] For example, Sperry (1969) writes that: 'First, conscious awareness [...] is interpreted to be a dynamic emergent property of cerebral excitation. As such conscious experience becomes

Kim (1999) does acknowledge that complex systems bring along new causal powers: 'Complex systems obviously bring new causal powers into the world, powers that cannot be identified with causal powers of the more basic simpler systems. Among them are the causal powers of microstructural, or micro-based properties of a complex system' (p.36). Strangely enough, however, he still claims that these properties are 'not themselves emergent properties; rather, they form the basal conditions from which further properties emerge (for example [...] consciousness is not itself a microstructural property of an organism, though it may emerge from one)' (ibid.) According to Kim, emergent properties such as mental properties can only be causal because they 'inherit' their causal powers from their 'emergence base' physical properties. This is what he calls the causal inheritance principle: if M is instantiated on a given occasion by being realized by P, then the causal powers of this instance of M are identical with (or a subset of) the causal powers of P (cf. Kim 1993). But Thompson argues that Kim's refusal to endow emergent properties with new causal powers is mainly due to his acceptance of part/whole reductionism, according to which micro-based properties are decomposable into the intrinsic causal properties of micro-level entities.

Of course, Thompson's story about 'mind in life' is not without problems. But it is helpful insofar it shows that some of the 'traditional' philosophical assumptions that might be in conflict with my pragmatic approach to intersubjectivity do not have to be taken for granted without questioning.[113] In this respect, the above considerations corroborate my own story about the 'mind in practice'.

---

inseparably tied to the material brain process with all its structural and physiological constraints. At the same time the conscious properties of brain excitation are conceived to be something distinct and special in their own right [...] Among other implications of the current view for brain research is the conclusion that a full explanation of the brain process at the conscious level will not be possible solely in terms of the biochemical and physiological data (pp. 533-5).

[113] Kim has always maintained that the problem of downward causation is primarily a *metaphysical* problem - of showing *how* mental causation is possible and not *whether* it is possible. But I think it is precisely Kim's metaphysics that is hard to swallow. Perhaps what we need is a notion of causation that is fundamentally *explanatory* (cf. Baker 1995). Instead of saying that explanation presupposes causation (as Kim does), we could say that the notion of causation presupposes a variety of explanatory practices. We do not necessarily need to motivate this skepticism about causality on Humean grounds. Norton (2003), for example, has argued that we can also justify our denial that the world is fundamentally causal by pointing at our 'enduring failures to find a contingent, universal principle of causality that holds true of our science' (p.2).