Enquiries Concerning the Minds of Others

Matthew Parrott

(DRAFT: Forthcoming in *Knowing Other Minds*)

Human life is inescapably social. We spend much of our time conversing, coordinating with others, reacting to what they say or do, soliciting them for information or advice, and anticipating their future behaviours. These kinds of social interactions, and many more, presuppose that other human beings have minds. Furthermore, we ordinarily take ourselves to know quite a bit about the minds of other people, especially those who figure most prominently in our lives. As Paul Snowdon remarks in his opening chapter, 'no normal person seriously doubts that we know about the minds of others.' (pg. XXX) We normally take ourselves to know what other people think, or feel, or want and we rely on this knowledge in our actions and interactions.

Knowledge of the mental states of others thus figures prominently in the sort of interactive social life characteristic of human beings. Insofar as we are social beings, we need to figure out what others think or feel, in order to interact successfully or negotiate with them. But how exactly do we manage to achieve this? Once we start to reflect upon our everyday knowledge of the mental states of others, a number of questions arise. To take a simple example, I might take myself to know what my daughter wants for breakfast. How do I know what she wants? She told me. This seems like a perfectly good way of knowing what someone else wants to eat for breakfast. But, upon further reflection, we might wonder whether her telling me about her desires is enough for me to acquire knowledge of what she wants to eat. After all, couldn't she be lying? And even if she isn't, does what she says really give me good enough evidence to know what she truly wants?

Further questions can arise because, although we seem clearly to be committed to engaging in social interactions, it is far from clear precisely how this is psychologically accomplished. How exactly do we go about attributing mental states to others? What psychological processes or mechanisms are involved? Am I in a position to judge that my daughter wants cereal for breakfast only because, in some sense, I understand a kind of psychological theory, as advocates of 'theory-theory' maintain. That is, does my capacity to understand

what my daughter wants for breakfast depend in some way upon my knowing an abstract generalization (which would partly compose a psychological theory) concerning desires and verbal behaviour? If so, is that a theory I learned at some point in my life, or is it innate? To turn to another example, do I judge that may daughter is in pain only because I have imagined what I would feel if I were in her situation, as simulation theorists hold? That is, do I somehow exploit the very same cognitive resources and capacities that I use to understand my own feelings when I think about the feelings of others? If so, how exactly is this imaginative task performed?

Any of these questions about our social relations might prompt us to investigate whether, to what extent, and how we are able to know about the thoughts, feelings, desires, and intentions of others, and any of them might be associated with what philosophers have traditionally called 'the problem of other minds'. However, as Snowdon notes in his chapter, this phrase can easily mislead us into thinking that there is a single, well-defined philosophical question involving the minds of others (pg. XXX). Even if that were true at some point in the history of philosophy, in contemporary philosophy, the phrase 'the problem of other minds' can no longer be taken to denote a single problem or question. Rather, it should be taken to refer to a set of questions that naturally arise from sustained reflection on our everyday social interactions. Moreover, given the current widespread interest in social phenomena, it is not surprising that we find the same set of questions being pursued in many other fields beyond philosophy, including, to name a few, developmental psychology, cognitive neuroscience, sociology, and computational psychiatry.

In the remainder of this chapter, I shall distinguish three main problems that naturally arise when we reflect upon our everyday social interactions. Someone who is interested in exploring 'the problem of other minds' might aim to make headway on any or all of these three problems, or indeed on some other question concerning our social relations with others. The first of these problems is the epistemological question about how an individual is able to know anything at all about the mental states of others. Indeed, as we shall see in section two, reflection on this question can even suggest that one may not even be able to know whether there *are* any minds other than her own. The second problem I shall discuss,

which is closely related to the question of knowledge, focuses on the nature of mental state concepts like 'pain', which we employ in thinking about another person's mental states. These concepts appear to be generally applicable, in the sense that we appear to apply them unambiguously to both ourselves and to others. Yet, as we will see in section three, the generality of these concepts can begin to look problematic, if we think one acquires mental state concepts on the basis of her own experiences of, for example pain. Finally, the third problem concerns the psychological processes and mechanisms underpinning our ordinary attributions of mental states to others. As we will see in section four, theoreticians have different conceptions of the psychological structures or capacities needed to attribute mental states to other creatures, and also diverge over how we should best understand their ontogenesis. In the course of distinguishing these three problems, I shall also briefly introduce the different ways in which each of them is addressed in the chapters that follow.

Reading the essays collected together in this volume, one gets the sense that everyday social relations encompass an assortment of complex and interrelated phenomena, which present a wide range of problems. Thus, enquiries into the nature of our social relations could reasonably proceed along a variety of different paths. Nonetheless, the epistemological, conceptual, and processing problems that I shall focus on have been at the heart of an enormous body of literature, and are therefore unsurprisingly also central to the chapters comprising this volume. Yet, aside from the fact that all three problems concern some dimension of the social, very broadly understood, it is not entirely clear how they are related. Do theorists addressing one of these problems have anything to learn from those addressing the others? Or might one of the three problems simply collapse into one of the others? To put this another way, why should essays that focuses on the epistemological problem, the conceptual problem, and the processing problem be gathered into a single volume? In the final section of this essay, I shall address the question of how these three problems, and the different approaches to solving them represented within this volume, relate to one another.

The Epistemological Problem

Over the course of much of the 20th century, a number of philosophers were worried about the challenge that scepticism posed to our knowledge of others' minds.[1] For example, John Wisdom asks, 'what lies behind the smiling face, pleasure, pain, or nothing?' (Wisdom, pg. 374) In the terms of this disjunction, the sceptic raises the worry that for all I know nothing at all lies behind another human's smiling face. Appearances to the contrary, I have no basis for knowing (or justifiably believing) that others even have minds, much less am I able to know specific things about what others think, feel, or desire. But, in that case, according to the sceptic's argument, I am completely ignorant of others' minds.[2]

One way that philosophers have tried to motivate scepticism about our knowledge of others' minds is by appealing to cases of pretence or deception. Although I take myself to know that my daughter wants cereal for breakfast, I might have had the very same evidence for thinking this if she were merely pretending. But, since I cannot know what she wants in the latter case, I don't really know that she wants cereal for breakfast in the case where she does. The possibility of pretence means that whatever evidence I may have for thinking that may daughter wants cereal for breakfast simply isn't good enough to know what she really wants.

But even if appealing to pretence shows that I cannot know what my daughter wants to eat, such an appeal limits the reach of the sceptic's challenge (cf. Snowdon, pp. XXXX). My daughter can only pretend that she wants cereal if she is a creature with a mind. So, the possibility of pretence, even on a very wide scale, would not show that I completely lack knowledge of others' minds, for it would presuppose that others exemplify some type of mental states. At best, the possibility of pretence might show that I am never in a position to know what someone is really thinking or feeling, but that would not show that there

---

[1] This introductory chapter is intended as a historically-based overview of the contemporary philosophical landscape and does not attempt to sketch a complete history of philosophical work on 'the problem of other minds'. For a more comprehensive historical perspective, see Avramides (2001, Pt I.).

[2] Pyrrhonian sceptics would resist putting the conclusion of the sceptical argument this strongly. So, the sceptic here should be understood as more of an Academic or 'Cartesian' sceptic, rather than a Pyrrhonian. For further discussion of this distinction among types of scepticism, see Vogt, 2014.

could be nothing at all behind a smiling face. So, in order to motivate a stronger form of scepticism, one which claims that I am not in a position to know that another human even has a mind, the sceptic must introduce a scenario in which other humans were something like mindless zombies or androids. This is the kind of scenario envisioned by Wisdom when he asks whether 'nothing' lies behind the smiling face. The basic thought would be that all we have to go on in making attributions of thoughts and feelings to others is their publically observable behaviour; and crucially that behaviour could very well exist without any corresponding mental states. In our everyday social interactions, we certainly presume that another person's observable behaviour correlates with states of her mind. This is evident in everything we say and do. But, according to this stronger form of scepticism, we have no basis or warrant for that presumption. For this reason, one simply cannot know whether or not other human beings are completely mindless.

But even if this possibility of mindless zombies or androids can be used to motivate a strong sceptical conclusion, it isn't clear that we have isolated a problem involving anything that is distinctive to minds or mentality. The possibility that other human beings are mindless zombies is simply an instance of a more general sceptical scenario in which appearances are misleading. The reasoning that suggests no minds lie behind observable behaviour is structurally no different than the reasoning that suggests appearances of material objects might really be vivid hallucinations. In both cases, a sceptical conclusion is prompted by considering a scenario that is indistinguishable from what one purports to know. If this is correct, then Anil Gomes seems right to conclude that 'what we have here is a problem about knowledge, one which can be raised in the case of our knowledge of others' minds. And whilst this is a way of motivating scepticism, it doesn't motivate a distinctive scepticism about others' minds.' (Gomes, forthcoming, pg. XXXX; see also Snowdon, Chapter 2 and McDowell, 1982)

In order to refute scepticism, in either its weaker or stronger forms, one thing we must do is answer the following question: how are we able to acquire knowledge of another mind? (cf. Stroud, 2000) But this is an epistemological question that many philosophers have been interested in even though they are not especially concerned with rebuffing scepticism. It is a question that can arise from a simple desire to understand our ordinary ways of knowing

about the minds of others, or from reflection on apparent differences between ordinary self-knowledge and knowledge of others' minds. One may therefore simply seek to better understand of how we actually do acquire knowledge of other minds, and, as Snowdon emphasizes, recent discoveries in empirical cognitive science will obviously be relevant to this sort of intellectual pursuit (Snowdon, pg. xxx).

As I just mentioned, several philosophers take this epistemological problem to emerge because of an obvious asymmetry between our ways of knowing about minds. Philosophers tend to characterize this epistemic asymmetry in different ways. In her chapter, Asa Wikforss initially describes it in terms of 'directness', claiming that 'whereas we know our own minds directly, knowledge of other minds is always indirect.' (XXX; cf. Ayer, 1953; Davidson, 1991). For instance, it seems that when I know that I am in pain it is simply by virtue of my being in pain. By contrast, my knowledge that someone else is in pain seems to be based in some way on her observable behaviour and is, in this sense, indirect (cf. Wikforss, pg. xxx). As Wikforss notes, often this epistemic asymmetry between self-knowledge and knowledge of other minds is also characterized in terms of 'inference'. For instance, Paul Boghossian reports that 'in the case of others I have no choice but to infer what they think from observations about what they do or say. In my own case, by contrast, inference is neither required nor relevant.' (1989, pp. 151-52) However, Wikforss suggests the asymmetry can also be characterized in terms of 'evidence', in the sense that self-knowledge is not normally based on evidence whereas our knowledge of others' minds always is. In her chapter, Wikforss argues for a 'weak' interpretation of the asymmetry according to which the epistemic difference between ordinary self-knowledge and our knowledge of others' minds should be understood as 'a difference in the type of evidence available' (Wikforss, p. XXX) Drawing on a range of empirical research in cognitive and developmental psychology, Wikforss also argues that self-knowledge has both an experiential and an inferential component and is therefore similar in many respects to our knowledge of others' minds. Her chapter thereby illustrates how a useful way to approach the epistemological problem of our knowledge of other minds is by getting clearer on the epistemology of ordinary self-knowledge.

Historically, the most popular way of addressing the epistemological problem is to claim that that our knowledge of others minds is based on some kind of inference. Two different conceptions of the structure of the underlying inference have been most prominent. First, many philosophers have argued that we use an analogical inference to know about another's mind (e.g., Mill, 1865; Ayer, 1956; Russell, 1948; Hyslop and Jackson, 1972). For example, I recognize that my publically observable wincing is caused by my pain. Then, I observe my colleague wincing; so, by analogy, I conclude that she is also in pain. Second, some philosophers have argued that knowledge of others' minds is grounded in a kind of inference to the best explanation (e.g., Hyslop, 1995; Fodor, 1987; Pargetter, 1984). I see my colleague wince and then engage in a process of abductive reasoning in order to explain the occurrence of her wincing behaviour; although there may be various available hypotheses that could potentially explain my colleague's wincing, the best explanation of her behaviour is that she is in pain; so, I conclude that my colleague is in pain. One obvious reason that this might be selected as the best explanation is that it makes my colleague's wincing behaviour much more likely, although other considerations plausibly contribute to it being the best available explanation as well (cf. Lipton, 2004).

Even though an inferential approach of some kind seems intuitive, in recent years more and more of philosophers have come to endorse the provocative idea that our knowledge of others' minds is a form of perceptual knowledge (e.g., Cassam, 2007; Green 2007; McNeill, 2012; Smith, 2015; Spaulding, 2015).[3] According to the perceptual model, we literally perceive that another person is angry, or happy, or wants cereal for breakfast. A number of factors have contributed to the ascent of the perceptual model, of which I will briefly mention only three. First, philosophers influenced by ordinary language philosophy seem inclined toward the perceptual model because we regularly use perceptual locutions to report our knowledge of others' minds. Dretske, for instance, motivates his version of the perceptual model by noting that 'we commonly say we see that another person is angry or

---

[3] The perceptual view has had widespread influence and been defended both by philosophers in the analytic tradition (including, in addition to those mentioned above, Dretske, 1973; Duddington, 1918; McDowell, 1982; Stout, 2010) and also by philosophers who have been influenced by phenomenology (e.g., Krueger, 2012; Krueger and Overgaard, 2012; Gallagher and Varga, 2014; Scheler, 1954; Zahavi, 2011).

upset.' (1973, pg. 35; cf. Avramides XXXX). More recently, Mitchell Green remarks that 'we often say such things as that we could see the anxiety on someone's face, fell the trepidation in her handshake, and hear the exuberance in her voice.' (Green, 2010, pg. 45; cf. Scheler, 1954)[4] Second, as both Soren Overgaard and Anita Avramides notice in their chapters, the notion that mental states are perceivable was a central tenet of classical phenomenology. Contemporary philosophers who have been influenced by this tradition often take the perceptual model to give an accurate account of the phenomenal character of our experiences of social interactions (e.g., Gallagher, 2008; Smith, 2015). Finally, a number of advocates of the perceptual model have been heavily influenced by the rise of the embodied cognition research programme in cognitive science, which claims that cognitive processes depend directly upon aspects of an individual's body other than her brain (Clark, 2008; Wilson, 2015). In his chapter, Overgaard describes 'embodiment' as the thesis that 'at least some mental states extend all the way to the available surface behaviour.' (pg. XXXX). If this is right, if proper parts of our mental states are identical to perceptible surface behaviours then perhaps, contrary to what many have thought, another person's mental states can be directly perceived. This last idea raises an important question as to whether perceiving the fact that, for instance, my colleague is currently angry requires me to be able to directly perceive my colleague's anger. This relation between perception of a fact involving a mental and the perceptibility of mental states is the topic of Overgaard's chapter. He argues that, contrary to what many have thought, the latter does not lend any greater plausibility to the former.[5]

---

[4] McDowell seems to be influenced by these sorts of considerations as well, but his own reasons for the perceptual model are certainly not restricted to them. One reason that clearly sways McDowell is that he thinks the perceptual model looks far more plausible once we reject a certain conception of 'traditional epistemology' which he is keen to reject. This is a conception that holds that 'the basis for a judgment must be something on which we have firmer cognitive purchase than the judgment itself.' (1982, pg. 471)

[5] It is worth noting that Overgaard takes his chapter to be focused on how this relation bears on the processing problem, which I shall discuss in section four of this chapter. Nevertheless, the notion he discusses, that direct perceptibility of mental states lends support to the claim that mental state attributions to others are a 'perceptual achievement' has also figured prominently in accounts that appeal to perception as an answer to the epistemological question (e.g., Gallagher, 2008; Gallagher and Zahavi, 2008; Krueger and Overgaard, 2012).

The perceptual model is rapidly gaining influence among philosophers and several of the leading figures working on it have contributed chapters to this collection (i.e., Avramides, Gomes, McNeill, and Overgaard). Defenders of the perceptual model tend to emphasize the degree to which others' minds are transparent to us, or how our knowledge of what someone is thinking or feeling is what Will McNeill calls 'spontaneous'. However, in her chapter, Avramides draws on the work of Stanley Cavell to highlight cases where 'others are very often an enigma to me.' (pg. XXXX). This, in her estimation, is largely because other humans are 'individuals who have voluntary control over the expression of their emotions,' which, she concludes, 'gets in the way of at least straightforward perception of their feelings and emotions.' (pg. XXXX) Avramides stops short of denouncing the perceptual model completely, and instead offers some intriguing suggestions for developing a more nuanced version of the model.

Do perception and inference exhaust the possibilities for answering the epistemological problem? The two options have certainly framed discussion for the last century. But is not clear that these options are exhaustive and it would be worth exploring whether there may be other ways of dealing with the epistemological problem. In his chapter, Anil Gomes aims to present just such an alternative by arguing that our knowledge of others' minds is based on a person's expressive behaviour (cf. Green, 2007; Parrott, 2017). Gomes reports that 'it seems to me that our ordinary ways of thinking take our knowledge of others' minds to be mediated by other people's expressive behaviour in a way that we do not take paradigmatic cases of perceptual knowledge to be mediated by the distinctive appearances of the objects of perceptual knowledge.' (pg. xxx; cf. McNeill, pp. xxxx) Thus, on his view, expressive behaviour plays an ineliminable role in securing knowledge of others' minds. This idea is plausible enough, but Gomes also claims that knowledge based on expressive behaviour is 'a non-evidential and non-perceptual way of knowing about other minds.' (pg. XXX)

Naturally, the possibility of a distinctive mode of 'expressive knowledge' would certainly open up exciting new options for addressing the epistemological problem of others' minds. But is expressive knowledge really distinct from perceptual knowledge? This question is the focus of McNeill's chapter. Although he agrees that expressive behaviour is the basis for our knowledge of others' minds, McNeill raises difficulties for proposals, like Gomes's, that

attempt to distinguish expressive behaviour from other sorts of observable behaviour. Without such a distinction, however, it is difficult to see how basing our knowledge on expressive behaviour would amount to a distinctive way of knowing. McNeill thus makes a case for sticking with a perceptual account of our knowledge of others' minds.

How then should we answer the epistemological problem of others' minds? One thing the contributions to this volume seem to agree upon is that an adequate answer to this question, whether a more carefully developed perceptual account or an account that carves out a distinctive epistemic role for expressive behaviour, must pay attention to the ways in which our knowledge of others is different from knowledge of ordinary material objects. That is to say it will need to appreciate the fact that other persons are agents with whom we have sophisticated social interactions. Only once we grasp the full implications of this fact will we be in a position to adequately understand how we are in a position to know about others' minds.

The Conceptual Problem

Around the middle of the twentieth century, a handful of philosophers, influenced by Wittgenstein, became fascinated with a problem concerning our very capacity to think about others' minds. This came to be known as the conceptual problem of other minds.

Wittgenstein makes several remarks that seem to suggest there is a distinct problem involving our ability to conceive of another person's mental states. For instance, he famously claims that it is 'none too easy' a task to 'imagine pain which I *do not feel* on the model of pain which I *do feel*.' (Wittgenstein, Sec 302) Because the conceptual problem originates in passages like this, it is closely connected to questions of Wittgenstein interpretation, which is a complex issue beyond the scope of this chapter (see Avramides, 2001, Ch. X; Child, 2011). Nevertheless, it is partly because Wittgenstein's remarks are so cryptic, that they allow for different ways of understanding the conceptual problem.

In the following passage, Nagel clearly contrasts the conceptual problem with the epistemological problem:

'The interesting problem of other minds is not the epistemological problem, how can I know that other people are not zombies. It is the conceptual problem, how I can understand the attribution of mental states to others. And this in turn is really the problem, how I can conceive of my own mind as merely one of many examples of mental phenomena contain in the world.' (1986, pp. 19-20)

The last sentence of this passage suggests that the conceptual problem concerns the generality of mental state concepts. Take my concept of pain. Nagel's question seems to be this: how am I able to have a sufficiently general, unambiguous, concept of pain that is applicable to both myself and to others? The conceptual problem of other minds asks us for an explanation of how one is able to acquire such a general and univocal mental state concept (cf. Avramides 2001, Ch. 8; Gomes 2011). For this reason, the conceptual problem seems to be prior to the epistemological problem, in the sense that the latter is only intelligible on the presupposition that we do have general mental state concepts.

The conceptual problem seems to suggest that there could be some obstacle preventing me from acquiring such a concept. Indeed, if there were no *prima facie* obstacle to acquiring a general mental state concept, then the demand for an explanation of how such a concept is acquired would pose no great challenge. The force and interest of the conceptual problem derives from the intuition that there may be some insurmountable obstacle to acquiring a sufficiently general mental state concept (cf. Gomes, 2011; McGinn, 1984). But, when it comes to articulating precisely what that obstacle might be, things are not so clear.

The problem seems to emerge from the very natural thought that one acquires mental state concepts by virtue of experiencing one's own mental states. For example, many have found it obvious that one acquires the concept of pain on the basis of one's own experiences of pain. The conceptual problem would arise if there were something about acquiring a concept in this way that prohibited extending or applying it to others. Thus, Avramides claims that 'a problem can be raised for experiential concepts that makes it difficult to see how a child who has come to have the concept of pain as the result of feeling pain in her own case is able to extend this concept to others' (2001, pg. 221). But what exactly is the difficulty here? Let's assume that one does acquire the concept of pain by virtue of

introspective reflection on one's own experience of pain. Why should a concept acquired in this way not be something that one can legitimately apply to others?

One thought is that an obstacle arises because of the previously mentioned epistemic asymmetry between ways of knowing about mental states. Whereas our own mental states are directly presented to our conscious awareness, we are never directly presented with another individual's mental states. Only a person's observable behaviour can serve as the basis for attributing a mental state concept to her. However, if we become aware of pains under radically different first- and third-person modes of presentation, this suggests that our concept of pain is really ambiguous, that we really have two distinct concepts, one which we apply on the base of first-personal experience and another which we apply on the basis of publically observable behaviour. As Davidson argues, 'if we are given no explanation of this striking asymmetry, we ought to conclude that there are really two kinds of concepts: mental concepts that apply to others, and mental concepts that apply to ourselves.' (1991, pg. 207; cf. Strawson, 1959) But if we really have 'two kinds of concepts' then we would lack general mental state concepts that can be applied to both ourselves and others.

However, why exactly should the existence of this epistemic asymmetry stand in the way of our acquiring unambiguous, general mental state concepts? As Colin McGinn notes, we seem to be able to acquire other general concepts, like the concept of square, by virtue of abstracting away from the 'sensory modes in which the property of squareness is presented.' (1984, pg. 136) That is, we do not have one concept of square that applies to visually perceived shapes and another that applies to felt shapes. So, it isn't clear that the mere existence of two distinct modes of presentation would itself be a sufficient obstacle to acquiring a general mental state concept.

McGinn suggests that the conceptual problem arises because it is hard to see how one could possibly abstract from the first-person perspective on one's own mental states. That is, he seems to think that if one acquires the concept of pain on the basis of first-person experience, then the first-person mode of presentation becomes part of the content of one's concept. It would follow from this that there would be an insurmountable difficulty when it comes to applying one's concept to others.

But is there any reason to think this is true? Why can't we abstract from the first-person perspective? It doesn't seem generally true that a concept acquired through a first-personal mode of presentation thereby becomes a first-personal concept. For example, it is plausible that I acquired my concept of walking by experiencing myself walking. But there seems to be no reason to think there is an analogous problem about how I am able to have a general concept of walking.[6] So why is the case of pain so problematic?

Perhaps the conceptual problem arises not because of the epistemic asymmetry but because of the nature of our mental states. Thus, one might naturally think that my pains are *essentially* mine. So, if I acquire the concept of pain by virtue of introspective reflection on my pains, then one might think that I will have a concept of a property that only I can exemplify. And how could a concept of something that is essentially mine be intelligibly applied to someone else? But it is not clear that one cannot abstract from essential features of the paradigm objects or properties that one exploits to learn a concept. To use Kripke's well-known example, learning the concept of duck by seeing a group of ducks in the park, does not entail that my concept of duck includes any of the essential properties of those ducks (cf. Kripke 1982) So it does not seem to be generally true that the essential properties of an object or property automatically become part of one's concept of that object or property. Perhaps there is something special about the essences of mental states that generates this sort of entailment, but it not yet clear what that special something is.

Let us briefly consider one final possibility for why there may be an obstacle to acquiring general mental state concepts. Regardless of their nature or essence, one might think that my mental states are presented to me in a manner that also *presents* them as properties that one could not be aware of in *any* other way. In other words, one might think that part of what it is to take a first-personal perspective on one's pains is to have those pains presented as entities that no one else could possibly be aware of, as exclusively first-personal. Although this is a substantive claim about the way in which pains are *presented* to

---

[6] Consider also the possibility of perceiving a material object via an egocentric mode of awareness. There seems to be no reason to think that other perceivers couldn't also be aware of that object in the same egocentric way.

conscious experience (and thus not merely a claim about the nature or essence of mental states), it might generate an obstacle to acquiring a general concept of pain.

This is one of the issues at the heart of the third chapter of Strawson's *Individuals*:

'If, in identifying the things to which states of consciousness are to be ascribed, private experiences are to be all one has to go on, then, just for the very same reason as that for which there is, from one's own point of view, no question of telling that a private experience is one's own, there is also no question of telling that a private experience is another's. All private experience, all states of consciousness, will be mine, i.e. no one's.' (1959, pg. 100; cf. Kripke, 1982)

The question to which Strawson draws attention in this passage is how I could come to have a concept of a private entity that is not private to me, if I acquired the concept of such an entity on the basis of my own conscious experiences. To do so would require me to have some general conception of a first-person perspective, or of a subject of experience. But it is very unclear how I could get *that* sort of concept from something other than a general mental state concept. Naturally, if I already had the concept of a subject of experience, then I could abstract from 'an object that only I can be aware of' to 'an object that only a subject can be aware of'. But the question is how do I acquire this general concept. Strawson himself introduces the notion of 'person', which he describes as logically primitive, to play this role. Strawson's explication of the concept 'person' and his arguments for its introduction, are critically evaluated in the second section of Snowdon's chapter (cf. Davidson, 1991).

Aside from Snowdon's chapter, only Heal's contribution to this volume explicitly mentions the conceptual problem. Why has it fallen out of fashion? One possible explanation is that the problem seems to be connected to some kind of verificationism about meaning, which contemporary philosophers largely reject (cf. Snowdon). Another possibility is that the rise of various theories in cognitive science, which we shall turn to in the following section, may give the impression that the conceptual problem has been solved. For instance, some versions of theory-theory claim that all attributions of mentality, both to others and to

oneself, rely on tacit knowledge of a general psychological theory. As such, any mental state attribution would presuppose possession of general mental state concepts, concepts which could be either innate or learned in early development. In either case, there would be no question of how one acquires a general mental state concept simply from 'one's own case'. Whether this sort of theory would suffice for dissolving the conceptual problem is an important question to which we shall return in the final section. Yet, even though the conceptual problem receives comparatively less discussion in the chapters that follow, a presumption that mental state concepts are generally applicable does lie in the background of both the theories and specific explanations that are discussed. It is therefore helpful to understand the structure of the conceptual problem.

The Processing Problem

The second part of the twentieth century brought several key developments that have reshaped the landscape of how philosophers and psychologists think about the mind. It was during this time that cognitive science emerged. Among other things, this involved the rapid emergence of a wide range of empirically-based research programmes, each of which generated its own novel set of theoretical and philosophical questions.

One of the foundational principles of cognitive science was the idea that thinking is just a type of information-processing computational process operating on mental representations (for an early indication of this see McCulloch and Pitts, 1943). Thus, mental tasks, ranging from decision-making to perceptual discrimination, are traditionally viewed as computational processes which solved specific information-processing problems. Cognitive scientists aim to uncover the details of how exactly these processes work.

Related to the emergence of cognitive science, more and more philosophers became persuaded by functionalist theories of the nature of mind, and by reliabilist theories of knowledge. The basic idea behind functionalism is that a mental state just is a kind of function that mediates transitions between a certain set of sensory inputs and a certain set of behavioural outputs (Lewis, 1972). If functionalism is right, then exhibiting a specific set

of behavioural outputs is part of what it is to be in the state of pain.[7] According to reliabilist theories of knowledge, our actual belief forming processes reliably produce true beliefs, and, for this reason, they are epistemically justified or warranted. On this sort of view, understanding precisely how those reliable processes work just is understanding how we know about some domain. This is because epistemic warrant or justification 'is a function of the psychological (perhaps computational) processes that produce or preserve belief,' (Goldman, 1999, pg. 26; cf. Kim, 1988).[8]

These intellectual developments all strongly influenced the way theorists sought to understand social interactions. In particular, many philosophers and psychologists turned to investigating how we actually go about attributing mental states to others. Like all mental operations, attributions of beliefs, desires, or sensations, can be viewed as solutions to information-processing problems. For instance, we are aiming to acquire sufficient information to predict how another organism will behave in some future context from observations of the creature's current or past behaviour. Or, we might be aiming to coordinate some type of joint action. In either case, there is a question of how this is accomplished. What are the cognitive or psychological processes or systems which solve these problems? What sort of capacities, structures, or mechanisms are necessary for attributing mental states? And how exactly do they work? These sorts of questions present a processing problem insofar as they ask about the cognitive processes implicated in ordinary mental state attributions. The processing problem has become increasingly popular and is presently being addressed by a great deal of contemporary philosophical and empirical work.

Theoretical approaches to the processing problem have been dominated by two frameworks. First, there is 'theory-theory' which claims that individuals draw on an internally represented folk psychological theory to attribute mental states to others. To

---

[7] Functionalism thereby came to supplant both behaviourism and mind-brain identity theories. For an overview, see Levin, 2013.

[8] Reliabilist theories of knowledge are thus notably different from traditional 'internalist' conceptions of knowledge and justification, which typically require some sort of access or awareness on the part of the knower (cf. Chisholm, 1977).

possess such a 'theory of mind' is just to possess a body of general information about mentality (see Davies and Stone, 1995). This information would include generalizations linking different types of psychological states to various canonical input conditions and patterns of observable behaviour. When we attribute mental states, which theory-theorists think of as theoretical entities, we subsume some observed behaviour under some general psychological principle or law (Gopnik and Wellman, 1992). By contrast, simulation theory claims that we attribute mental states to others by virtue of simulating being in their situation (see Gallagher and Fiebich this volume). Robert Gordon illustrates this theory with an example of expert chess players who 'report that, playing against a human opponent or even against a computer, they visualize the board from the other side, taking the opposing pieces for their own and vice versa.' (Gordon, 1986, pg. XXX) Unlike theory-theory, which emphasizes the importance of general theoretical knowledge, simulation theory seems to privilege the first-person perspective in accounting for our capacity to attribute mental states to others. It is by knowing what about my own mental states in a certain counterfactual or imaginative situation that I am able to attribute mental states to someone else who is in that same type of situation.

A considerable number of objections have been leveled against each of these theoretical approaches and, many contemporary philosophers now opt for some type of hybrid view, according to which mental state attributions are sometimes the result of appealing to a theory and sometimes the result of simulation, or which aim to blur the distinction between simulation and theory. Moreover, as we saw previously, a number of researchers have also been influenced by recent developments in cognitive science, such as the embodied cognition research programme. One consequence of this has been the development of alternative theoretical frameworks like 'interaction theory', which is presented by Gallagher and Fiebich in their chapter. According to Gallagher and Fiebich, interaction theory diverges from theory-theory and simulation theory by virtue of focusing on 'understanding the meanings that emerge in our overt intersubjective actions, interactions, and social practices.' (Gallagher and Fiebich, pg. XXX)

Although Gallagher and Fiebich find much to value in interaction theory, their chapter aims to develop and defend a 'pluralist theory' of our capacity to attribute mental states to

others. In contrast to hybrid views, which they claim privilege one kind of cognitive process in explaining social understanding, pluralism 'neither favors one perspectival stance over another one (e.g., an interactive over an observational stance) nor any single social cognitive process over others.' (pg. XXXX) Thus, on their view, which cognitive processes are implicated in mental state attributions depends on a variety of different parameters, which vary from context to context. It is the goal of a fully developed pluralist theory to capture this complexity of social interaction and elucidate the precise roles played by contextual parameters in attributions of mental states to others.

Another significant development in cognitive science that has been very influential is the rise of two-systems models of cognition (for discussion see Jacob this volume). Traditionally, a single task like logical reasoning was thought to be sub-served by a single cognitive process or system. However, a two-systems model claims that some tasks are actually carried out by two distinct cognitive processes or systems, one of which is typically thought to be extremely fast and automatic and the other of which is typically thought to be slow and subject to voluntary control. (cf. Kahneman, 2003) The idea that mental state attributions are carried out by two systems has been defended by Ian Apperly and Stephen Butterfill. In their estimation, the hypothesis that mental state attributions are implemented by two distinct systems is the best way to account for the fact that belief ascription has competing demands of efficiency and flexibility (2009, pg. 953) It also explains fascinating but seemingly incompatible experimental results that use the paradigm of the false-belief task.

Many experimental studies seem to show that children do not development a capacity to attribute beliefs to others until around the age of four, but several other studies suggest that infants are already aware of another person's false beliefs. As Pierre Jacob characterizes it in his chapter, this presents 'the basic developmental puzzle…: why do 3-year olds fail explicit change-of-location or unexpected contents false-belief tasks if toddlers or even preverbal infants can represent the contents of others' false beliefs about either an object's location or unexpected contents?' (Jacob, pg. XXXX) Apperly and Butterfill's two-systems theory would offer a solution to this puzzle by insisting that there is one cognitive process that supports 'minimal' mental state attributions in infants and another, and a more

sophisticated system that underwrites the kind of flexible capacities to attribute mental states that we find in adults. However, Jacob's chapter raises a number of serious challenges for the two-systems framework. By drawing on a range of empirical work, Jacob carefully makes the case that the two-systems theory cannot actually resolve the developmental puzzle. He also canvasses a number of studies which he takes to suggest that a more sophisticated and flexible system may already be operating in children aged 18-months.

As empirically oriented study of the mind continues to expand in new directions, the number of questions concerning the cognitive processes underlying our ordinary mental state attributions will increase. New experimental discoveries will present opportunities for theorists to fine-tune existing cognitive theories or develop new theoretical paradigms, and each new development will likely present additional questions for philosophers to consider.

Philosophy, Psychology, and Morality
We have now seen that there are several distinct problems that fall within the domain of the 'problem of other minds'. But how exactly do these problems relate to each other? Might answers to the processing problem dissolve the epistemological or conceptual problems? Similarly, would a satisfying answer to the conceptual problem have any implications for what empirical science could discover about our psychological capacities to attribute mental states?

One possibility is that these three problems are completely autonomous from one another. On this view, the best cognitive theories of how mental state attributions are produced have absolutely no bearing on how we should think about the epistemological or conceptual problems. The processing question requests a description of how we actually attribute mental states. But learning descriptive facts does not imply anything about the epistemic properties that would be relevant to the epistemological or conceptual problems. The latter are questions that simply cannot be answered by giving a description of how we actually make mental state attributions. For example, the idea that the epistemological problem is strictly independent from the processing problem is suggested by Overgaard when he claims that 'it might be that the routines and procedures that people usually employ when they

attribute mental states to others fail to provide good justification for the resulting beliefs, epistemologically speaking.' (Overgaard. pg. xxx; cf. McNeill, 2015)

Similarly, if these three problems are substantially independent, it might be that our best theories for why mental state attributions count as genuine knowledge are completely disconnected from how those attributions are actually produced. For instance, suppose we are persuaded that, from a purely epistemic point of view, our knowledge of others' minds is a form of perceptual knowledge. One might think this would imply that the cognitive processes underlying our knowledge are non-inferential. But many think it is consistent to hold that non-conscious cognitive processes have inferential structure, while also thinking that, at the conscious level, our way of knowing is non-inferential (McNeill, 2015).[9]

Prima facie, strict independence might seem less plausible when it comes to the conceptual and processing problems, especially since prominent theoretical frameworks like theory-theory claim explicitly to address the conceptual problem. As I have mentioned, according to several theory-theorists, mental state concepts just are 'theoretical concepts', and they are either innate or are acquired like any other theoretical concept. Nevertheless, one might argue that the conceptual problem is independent from the processing problem by insisting that it is concerned primarily not with how we actually acquire or apply mental state concepts but with what entitles us to think that those concepts really are general and unambiguous. Even if the theory-theorist can describe how we apply concepts across both first-and third-personal contexts, this would not yet address the question of whether we are correct to do this (cf. Avramides, 2001; Davidson, 1991).[10] That is, the theory-theorist would

---

[9] This sort of view may be attractive to proponents of predictive processing models of perception and cognition. Andy Clark, for instance, claims that these models have no immediate implications for theories of conscious experience because 'what is on offer, after all, is a story about the brain's way of encoding information about the world. It is not directly a story about how things seem to agents deploying that means of encoding information.' (2013, pg. 196) If Clark is right, then one might argue that the epistemological and conceptual problems concern 'how things seem to agents' and so have no implications for how the brain encodes relevant 'information about the world'.

[10] Many theory-theorists deny that there is an asymmetry between first and third-person ascriptions, claiming instead that observable behaviour is the basis for all mental state attributions (Gopnik, 1993). But this is just to deny the phenomenon that generates the conceptual problem.

not have answered why one is entitled to attribute the same mental concept on the basis of ordinary self-knowledge and on the basis of observable behaviour. Similarly, one could argue that autonomy holds in the converse direction. If the conceptual question is concerned primarily or exclusively with what justifies the general applicability of mental state concepts, it might seem that any answer to this question would be compatible with a range of empirically-based accounts of how we acquire or apply those concepts.

If some form of autonomy thesis is correct, then the different problems we have been considering are not really connected. The best approaches to the epistemological problem or the conceptual problem would offer no clear insights into how we actually attribute mental states. Nor would the best account of how cognitive systems process information in order to produce mental state attributions have any implications for the questions traditionally addressed by philosophers. We could be in a position to answer any one of these questions without having to consider the leading approaches for addressing the others.

There are, however, more unified ways of conceiving of the relationship between these different problems. First, one might claim that both the epistemological and conceptual problems can simply be reduced to the processing problem. Quine famously proclaimed that 'epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science.' (1969, pg. 82) Philosophers who have been influenced by Quine's doctrine of naturalized epistemology, or by closely related reliabilist theories of epistemic warrant, will be inclined to think that the best way to answer the epistemological problem is just to give an accurate description of the psychological processes that produce mental state attributions. Such philosophers will take an analogously reductive approach to the conceptual problem and claim that grasping how it is possible for us to have general mental state concepts simply requires an empirically-informed understanding of the processes which actually generate our mental state concepts. In either case, the former questions can be reduced to questions about real psychological processes or capacities.[11]

---

[11] Naturally there are also many questions that can be about the reductive project, which I lack space to discuss here.

Autonomy and reduction are both extreme views about the relationship between the questions we have been considering. In between these two extremes, there is an intermediate position. It is a view according to which each of the problems provides constraints on the others. The problems are not substantially independent, as the autonomy thesis holds, but nor is it the case that we can reduce any one of them to any of the others.

Insofar as the problems we have been considering all concern our everyday social interactions, the view that the relationship between them is one of mutual constraint is an attractive option. It is implausible to think that epistemological and conceptual theorising are not answerable to advances or discoveries in cognitive science. To take just one example, we cannot have an epistemological theory explaining our knowledge of others' minds that is impossible for any human brain to implement. So, for instance, if, as Gomes argues, our knowledge is based on expressive behaviour, this suggests that humans must have the right sort of cognitive capacities required for discerning and recognizing expressive behaviours. It is therefore extremely plausible that developments in cognitive science that illuminate the structure of those capacities ought to constrain the ways we develop any epistemological theory of expressive knowledge of others' minds.

The converse is also extremely plausible. Reductionism is often motivated by the presumption that whatever processes generate our actual mental state attributions are reliable and thereby justified. But even if this is correct, it doesn't entail that epistemological theorizing or conceptual analysis has no bearing on how we ought to approach the processing problem. Suppose, for example, that an inferential account is the best way to answer the epistemological and conceptual questions. According to such a view, my knowledge that you are in pain is based on a kind of inference which takes me from a premiss about your wincing behaviour to a conclusion about your mental state. But not just any cognitive transition would yield knowledge. Rather, the transition must be mediated by my knowledge, most likely tacit, of the right sort of inference rule. This has implications for how we ought to approach the processing problem. Specifically, if it is correct, then it indicates that the cognitive processes implicated in my attribution of pain would need to include some component or mechanism that both mediated the cognitive transition in this

particular case and was also a common causal factor in any cognitive transition that instantiated a similar pattern of reasoning (Davies, 1995; 2000). This shows how an answer to the epistemological question can inform and constrain the way we theorise about the information-processing systems underlying mental state attributions.

The intermediate position I have just sketched views the 'the problem of other minds' as a set of interrelated questions, both philosophical and empirical, that arise from reflection upon our shared social life. It is worth noting that, although this set includes the three problems we have been discussing, its membership is not exhausted by them. For example, questions about how we relate to others have long been a central concern of moral philosophy. The connection between sharing a life with others and the nature of value is explored by Jane Heal in her chapter. As she describes it, her chapter is meant to remind us that 'our interactions with others sometimes take the form of reflecting on what is or may become valuable to us and on what we should do in the light of that.' (Heal, pg. XXX) This is a useful reminder because, according to a familiar sceptical picture, there 'is no such thing as objectively valuable'. (Heal, pg. XXX) Rather than denying the existence of value, Heal recommends a framework for thinking about value according to which what is valuable depends in various ways upon the groups to which we belong. For Heal, once we take the notion of 'collective intentionality' seriously we can see that we are part of an 'us', which means that there 'is already in place a standpoint other than mine and yours from which my ideas and your ideas about what is valuable for us can be appraised.' (Heal, pg. XXX) Heal's chapter thus reminds us that questions concerning our social relations with others are inevitably intertwined with questions of value and moral responsibility.

All the chapters collected in this volume could be said to embody the intermediate position that I have just sketched. The authors of chapters that focus explicitly on more traditional philosophical issues are clearly attentive to recent advances in cognitive science. Similarly, the authors of chapters that focus more attention on the latest developments in empirical research are quite alert to how experimental results might cohere with what they regard as the most credible frameworks for addressing the traditional epistemological or conceptual problems. Therefore, the essays in this volume can be auspiciously read as informing and

constraining one another, offering a range of perspectives on the phenomenon of our shared social life.