1 Empathy, Sympathy, Concern and Moral Agency

Neil Roughley and Thomas Schramme

Humans frequently act and react in line with what they take to be *morally* right or wrong. Humans are also animals with strong *prosocial* behavioural tendencies, that is, tendencies to benefit others. Both facts cry out for explanation. Historically, the first kind of explanation has been seen as the province of philosophy, whereas the second has primarily been the object of research in social, developmental and – more recently – neuropsychology. This division of labour is unfortunate, not least because neither question is likely to be answered satisfactorily without consideration of the relationship between morality and prosociality. Moreover, the tools of both disciplines – conceptual analysis, stringent argument and controlled experiment – are equally germane to both questions. Finally, and key for the topic of this volume, there are good reasons to think that certain kinds of emotional connection between human agents are crucial for the explanation of both forms of behaviour. Here, we group together the kinds of emotional connection that appear the best candidates for such a role – empathy, sympathy and concern – under the label *forms of fellow feeling*. The question uniting the contributions to this volume concerns the roles that these forms of feeling have in explaining moral thought, moral action and even moral norms, roles they may in part have as a result of their contribution to prosociality.

Forms of fellow feeling have been of central interest for both psychology and philosophy. One of the most prominent psychological claims in this area has been Dan Batson's 'empathy-altruism hypothesis'. According to the hypothesis, where an agent's welfare is threatened or impeded, an observer's coming to feel a *valence-congruent emotion* for the agent generates motivation in the observer, the ultimate goal of which is the removal of that threat or impediment (Batson 2011, 29). If the hypothesis were correct, then valence-congruent affect would be a highly important motivator of prosocial behaviour. It would be particularly important because the motivation it tends to generate is intrinsic motivation to contribute to another's welfare. Such motivation is in turn plausibly at work in a fairly

3

large set of cases that fulfil a condition frequently taken to be decisive for moral motivation by moral philosophers: that is, action for others out of intrinsic reasons.

Psychologists have further made the case that comparable forms of prosocial motivation may also be generated in an agent by their coming to feel emotions as a result of *adopting another person's emotional perspective* (Hoffman 2000, 30ff.). Negative emotional reactions grounding both directly in concern for another's threatened welfare and in forms of perspective taking – generally distinguished as 'sympathy' and 'empathy' – thus look to be good candidates for a mediating, perhaps even constitutive role in moral agency. That is, it seems that the dispositions to sympathy and empathy might be essential features of any person who qualifies as a moral agent.

This hypothesis is advanced in no uncertain terms by the psychologist Frans de Waal (de Waal 2006) and the philosopher Michael Slote (Slote 2004; 2007; 2010), both of whom see empathy as decisive for morality in its entirety. Slote in particular delineates structures according to which empathy is not only central to forms of moral motivation, but also to the functioning of moral judgement and even as the criterion of morally right action. A less radical position has been advanced by Lawrence Blum, who takes it that central areas of the moral life, if not its entirety, ground in empathy and sympathy (Blum 1980; 1988). Nel Noddings has advanced a structurally similar position, according to which the key emotional dimension of morality is a feature of what she calls 'care' (Noddings 1984). Such conceptions are related to positions in the history of philosophy, broadly known as 'sentimentalist', in particular those of David Hume, Adam Smith and Arthur Schopenhauer. However, philosophers since Kant have repeatedly criticised conceptions of morality grounding in such emotional mechanisms, arguing that they are unreliable, biased towards friends, family and in-groups and unable to track important morally relevant properties other than welfare (Kant 1785, 398f.; Prinz 2011a; 2011b).

Present-day conceptions of morality as based in some form of fellow feeling or, more modestly, as requiring fellow feeling at some significant point in their structure are advanced in a context in which there is burgeoning social, developmental and neuropsychological work on empathy, sympathy and concern (Eisenberg and Strayer 1987; Zahn-Waxler et al. 1992; Preston and de Waal 2002; Singer and Lamm 2009; Bischof-Köhler 2012; Decety 2012; Maibom 2014). It seems that any assessment of the viability of moral conceptions that place such mechanisms at centre stage should do so on the basis of an understanding of what the empirical literature tells us about their functioning. This is particularly important in view

of the very specific philosophical frameworks within which the original 'sentimentalist' theories were developed: Hume's and Smith's discussions are framed by an empiricist philosophy of mind and Schopenhauer's by a rather extravagant metaphysics. Moral sentimentalism today, or any ethical theory in which sympathy, empathy or concern plays a significant role, ignores the exciting empirical work on these topics at its peril.

The following introductory discussion has two parts. In the first, we examine various phenomena of fellow feeling, from the point of view of their conceptual structures, the conditions of their development and their import in social interaction. In the second part, we turn to the question of the role of these phenomena within various conceptualisations of the moral life.

The *first part* of the introduction begins with a proposal as to how we should best narrow down the field marked by our somewhat archaic umbrella term *fellow feeling*. We use it to pick out above all positive affective relationships between the emoter and some other sentient being, in particular affective relationships that seem possible without presupposing the existence of moral norms. This specification raises the question of the status of what is often called 'cognitive empathy', which appears necessarily to involve neither a positive relationship nor an affective component. In discussing this question, we summarise Heidi L. Maibom's instructive contribution to the volume (1.1). We then go on to look in some detail at the variations in the parameters that allow the conceptualisation of the various emotional phenomena covered by the terms 'empathy', 'sympathy' and 'concern', as well as of closely related states and mechanisms. This is especially pertinent, as both the mechanisms involved in the relevant affective phenomena and their relationship to motivation differ in various ways. We focus on two decisive parameters. The first is the extent to which there is a correspondence between the valence of a fellow feeling and the valence of that feeling's target (1.2). The second is the question of whether playing host to the relevant fellow feeling depends on a mechanism of affect transfer (1.3). Both parameters help to draw a line between empathy and what is frequently labelled 'emotional contagion', on the one hand, and sympathy and 'personal distress', on the other. They also help us to provide a suggestion as to how concern and 'care' should be conceived.

With these distinctions in hand, we then turn to the two central affective phenomena at issue in the volume, *sympathy* and *empathy*. We argue that for an understanding of both, as of their relationship, it is essential to be clear on their intentional structures, that is, on the way in which they conceive, refer to, or are 'directed at' other agents and their mental states. To this end, we introduce a set of terms that pick

out different dimensions of the intentionality of the states and processes involved. On this basis, sympathy can be analysed as an emotion, or group of emotions, whose 'target' is another sentient being and whose 'formal content' concerns an impediment to the welfare of the being thus picked out, where the particular type of impediment is specified by the emotion's 'focus' (1.4). Whereas sympathy's structure, thus understood, is fairly transparent, the same cannot be said for empathy, which, as a phenomenon grounding essentially in an affect transfer mechanism, may appear to have as its target either another agent or merely her mental states. It is here that it seems imperative to distinguish different affective phenomena that may be called 'empathy', only some of which involve the positive take on the other that makes a form of feeling a feeling for one's fellow. Transfer or mirroring of emotional states of another is insufficient for such a take. This feature only enters where the empathiser 'goes along with' the relevant emotion, that is, where her emoting becomes genuinely vicarious. We offer a proposal as to how this should be understood (1.5).

The last section of part 1 of the introduction looks at the *causal environment* of the various feeling forms. The questions of both their evolutionary status and of their proximate causal conditions are discussed, as are their causal and conceptual relations to other morally relevant phenomena. Here, proposals developed in several contributions to the volume – those of Batson and Slote, as well as of Doris Bischof-Köhler and Norbert Bischof – are summarised and set into relation to one another (1.6).

The *second part* of the introduction goes on to ask what the import of the emotion concepts distinguished might be in differing conceptions of moral agency. The likelihood that they will be of moral import derives from the fact that moral action, which is at least primarily a kind of action for other-regarding reasons, is likely to ground in natural capacities to transcend one's own self-interest. Such capacities seem to be at work where someone sympathises with, empathises with, or is concerned about some other or others. Three empirically based contributions to the volume that discuss the relationship of fellow feeling and prosociality, particularly in ontogeny, are discussed at the beginning of part II of the introduction. Here, we present the findings and conclusions of Nancy Eisenberg, Amrisha Vaish and Carolyn Zahn-Waxler and her coauthors Andrew Schoen and Jean Decety (2.1).

From here, we move on to the question of the significance of fellow feeling for an understanding of morality, where morality is taken to substantially overlap, but not to be identical with, prosociality. If we take it that moral agents are particularly concerned with moral reasons, we need to hone in on competing conceptions as to how *moral reasons* might be generated or grasped. Section 2 of this second part of the introduction

therefore begins by outlining two distinctions among moral theories which provide some clarity as to the role forms of fellow feeling might be assigned. The first distinction is established on the basis of competing answers to either the metaphysical question as to how moral reasons are constituted or to the epistemic question as to how agents come to grasp moral reasons. In either case, the answer can refer essentially to emotions. Where emotions are taken to be essential either to the constitution or understanding of moral reasons, the corresponding theory can be labelled *sentimentalist*; where no such affective features are thought to be essentially involved, the theory can be labelled *rationalist*. Clearly, forms of fellow feeling can only be thought decisive for moral agency if the relevant moral theory is sentimentalist in at least this undemanding sense.

A second distinction divides theories along the lines of *how* emotions are thought to fulfil their constitutive or epistemic function. According to one kind of approach, moral values or reasons might be constituted through the adoption of a perspective on the weal and woe of others that takes them as the object of sympathy, concern or care. In this perspective, the fate of the relevant beings is registered in terms of their effects on an observer's own sympathetic dispositions. Such an approach can be labelled a third-personal sentimentalism. This can be contrasted with second-person approaches, according to which reasons are constituted in two-way interaction between moral agents, rather than in a one-way receptive process. In a second-personal sentimentalism, empathy, understood as a mode of adoption of others' affective perspectives, might be taken to be decisive for the constitution of moral reasons (2.2).

The paradigmatic construction of a *third-personal* sentimentalism is provided by the eighteenth-century philosopher Francis Hutcheson, for whom value is constructed from the point of view of a sympathetic or benevolent angel. Hutcheson takes this metaethical model to lead naturally to a utilitarian normative ethics, according to which the morally right is determined by the relative level of well-being that would result from the realisation of competing behavioural options. In contrast, a *second-personal* sentimentalism is naturally associated with the generation of what, in philosophical parlance, are known as agent-relative reasons. These are reasons for actions owed by agents to specific others, independently of the overall good, which is identifiable from a view from nowhere. Stephen Darwall's second-personal account of morality belongs here; accordingly, his contribution to this volume is presented at this point. There may appear, then, to be a close relation between empathy understood as vicarious emoting and a deontological ethics that specifies agent-relative duties, on the one hand, and sympathy, understood as a third-personal reaction to the well-being of others, and utilitarianism on

the other. The third section of part 2 of this introduction discusses the plausibility of such a mapping and defends a more differentiated understanding of the possible relations (2.3).

With these clarifications in place, we turn in the final section of the introduction to the question of the roles various normative ethical theories might assign to forms of fellow feeling in their conceptions of *moral agency*. Moral agency plausibly involves a number of key susceptibilities and capacities: a sensitivity to moral reasons; an ability to deliberate appropriately on their basis; a propensity to be motivated by the perception of or deliberation on moral reasons and a capacity to act correspondingly. If one were to take it that morally appropriate behaviour is necessarily mediated by deliberation, it may appear that the primary agential capacity required is that of the application of the criteria of right action. Such an intellectualist conception of moral agency would leave little room for fellow feeling. However, such a conception is implausible, as moral agency in many cases works spontaneously, generating right action without the need for deliberative mediation. This point motivates a distinction between criteria of the right and forms of morally appropriate moral motivation. One way to elaborate this point involves assigning a secondary role to abstract criteria of the right, instead focussing primarily on situation-specifically appropriate motivation. Such a theoretical move is characteristic of virtue ethics. A second, consequentialist way of fleshing the point out is to see those forms of everyday moral motivation as desirable that happen to contribute to the realisation of what is right all in all, even if everyday moral agents are completely unaware of the connection. Conceptions of this latter kind might see forms of fellow feeling as contingent components of moral agency, perhaps components whose role needs strengthening by susceptibility to sanctions.

More interesting for the purposes of this volume are conceptions for which fellow feeling is an essential feature of moral agency. This might be so if local forms of sympathetic motivation are taken to be necessary substitutes for the universal benevolence that structures the contours of the right, but which is beyond everyday agents (Hutcheson). Alternatively, it might be so if the expression of 'full empathy' is taken to be the criterion of the right (Slote). And it might be so if respect is taken to be dependent on vicarious emoting (Darwall). At the end of this section of the introduction, we present the views of those contributors to the volume that centrally involve claims as to the role of fellow feeling in moral agency, that is, the views of Thomas Schramme, Lawrence Blum, John Deigh and Neil Roughley (2.4).

As will have become clear even at this stage, many of the authors of this volume have already contributed significantly to the key issues to be

discussed here. Indications of such contributions will be woven into the following discussion. Where the articles the authors have penned for this volume are focussed on, the relevant paragraphs will be introduced by their names set in bold typeface.

1.   Forms of Fellow Feeling

There appear to be a number of different emotional mechanisms that relate moral agents and moral 'patients' and that can lay claim to a central role in moral agency.[1] The psychological and philosophical literature has produced a wide range of terms with overlapping meanings to pick out these and related mechanisms. As is to be expected, the parameters employed in different terminological coinages depend on the varying interests of the authors. As is also to be expected, there are a number of confusions that result both from a somewhat uneven use of the relevant parameters as well as from the collision of only slightly, but significantly different terminological stipulations. Here, we attempt to order the field and at the same time situate the contributions of the authors of this volume to the decisive conceptual questions. Obviously, it is not the words used that are important, but the phenomena we should be picking out. As we need to be able to pick them out reliably and unambiguously, conceptual clarification is of the utmost importance.

1.1   *Fellow Feeling*

Let us begin with the rather old-fashioned formulation that gives this volume its title: 'fellow feeling'. And let us begin with a wide explication, that is, by naming fairly broad necessary conditions, which we can narrow down as we progress. For a psychological state or mechanism to be a form of fellow feeling in the umbrella sense in which we are using it here, it has, firstly, to be or involve an affective component and it has, secondly, in some significant way to relate its bearer to some other sentient being. In order to further narrow the range of affective states covered by the term, two additional moves seem appropriate. The connection to the relevant other should, thirdly, involve some sort of positive take on the other. Anger, for instance, is an emotion involving a significant relation to another, but does not belong here. Fourthly, the emotion should not be one that can only be made sense of if one presupposes the existence of normative standards. Moral respect is a

---

[1] Moral patients are those affected or intended by moral actions. Patiency hence refers to 'the other side of agency' (Reader 2007).

matter of a positive relation to another or others, but seems to require the acceptance of moral standards. In contrast, the attitudes, mechanisms and practices picked out by terms such as 'empathy', 'sympathy', 'concern', 'compassion' and 'care' are generally thought to be features of human life that are natural in the sense that they have at least basic variants that do not presuppose morality, but to which they may contribute substantially.

Alongside the terms just mentioned, the psychological literature makes use of the related locutions 'emotional contagion' and 'distress' as well as working with various compounds such as 'empathic concern', 'empathic distress', 'sympathetic distress' and 'personal distress'. The latter three of these compounds tend to be thought of as designating phenomena that should not be subsumed under the umbrella term, as they entail no positive take on others. Finally, various authors distinguish different variants, particularly developmental variants of empathy, picked out by qualifying adjectives such as 'egoistic', 'quasi-egoistic', 'veridical' and 'full' or 'full-blown'.

Before embarking on a discussion of these varying faces of fellow feeling and their relatives,[2] a word is in order about a phenomenon that seems infelicitously grouped together with the affective states and processes just listed, although it is picked out by one of the fellow-feeling terms. 'Empathy' is sometimes used to refer to a specific variant of so-called theory of mind, mind reading or social cognition, the variant known as 'simulation'. As the procedure is carried out in order to achieve cognitive aims, it is sometimes known as 'cognitive empathy'.[3] In simulation, mental states of another are, in Goldman's terms, 'enacted' by the agent on the basis of the other's situation in order to enable ascription of those states or the products of decision making. Simulated beliefs and desires, for instance, generate 'pretend-intentions', which in turn cause genuine beliefs as to the contents of the other's intention. 'Simulated' or 'pretend' mental states are frequently characterised by simulation theorists as 'offline', that is, they are quarantined off from the mental states that might interfere with the agent's taking on the perspective of the other (Goldman 2006, 20). For these cases, characterising the process

[2] A number of the contributions to this volume contain suggestions as to how we should differentiate between various related concepts. Cf. Batson this volume, 60ff.; Blum this volume, 151ff.; Darwall this volume, 291ff.

[3] James Blair uses the compound expression as a synonym for 'Theory of Mind' (Blair 2005; 2007; 2009). Other authors, for instance Alvin Goldman and Karsten Stueber, simply talk of 'empathy', but make it clear that they are using the term to pick out simulationist variants of social cognition (Stueber 2006, 111ff.). Goldman specifies that mind reading as simulation is 'an extended form of empathy (where this term's emotive and caring connotation is bracketed)' (Goldman 2006, 4).

that generates the relevant beliefs as 'empathy', even of a cognitive kind, is potentially misleading. It is potentially misleading because the core of the term's etymology is the Greek 'pathos', that is, 'feeling' or 'affect'. Of course, semantics is not a slave to etymology. Nevertheless, there are two important questions regarding the precise relationships between affect and cognition here, and so it seems sensible to keep the relevant phenomena as terminologically distinct as possible.[4]

The first question concerns the role of 'mindreading' in (emotional) empathy. Some authors take it that the former is a necessary precondition of the latter (cf. Blair 2005, 700; Blair 2007, 5; Blair and Blair 2009, 146). Moreover, as Goldman has pointed out, certain conceptualisations of empathy or empathising – particularly that developed in Baron-Cohen's early theory of autism (Baron-Cohen 2003, 21ff.) – work with cognitive, affective and motivational components without differentiating their roles (Goldman 2006, 203f.). Doing so is problematic in view of the separability of the relevant phenomena, both conceptually and neurologically. The point appears particularly important in view of the disputed extent to which autism involves an impairment not just of mindreading, but also of the affective processes picked out by talk of empathy (Blair 2005, 706ff.; Blair and Blair 2009, 146ff.; Baron-Cohen 2011, 82f.; Schramme, this volume).

The second point at which distinguishing between cognitive and affective mechanisms is paramount here concerns a very specific kind of mind reading, one which, if it really is something humans do, would merit an etymologically accurate description as 'cognitive empathy': it would be a form of social cognition that is itself essentially affective. Certain authors take this to be the case where the mental phenomena generated in the mind reader are themselves affective, for instance, where the simulation of beliefs and desires leads the simulator to 'enact' the sadness of another. Here, it seems that the cognitive process of simulation must itself involve playing host to an affective state. This is plausibly what Batson is claiming when he says that taking on 'the imagine-other perspective' on another's emotional state involves 'not simply understanding, but *sensitive* understanding' (Batson 2009, 267, emphasis in original; cf. this volume, 63f.).

There is, however, a serious question here as to how we should precisely understand an agent's playing host to an affective state in a purely 'pretend' manner, that is, in a way that does not involve the agent having the emotion in a full sense. In his classic discussion, Max Scheler

4 Stueber, however, argues that there are decisive historical reasons why the term 'empathy' should primarily be used for the cognitive phenomenon (Stueber 2006, 26ff.).

attempts to provide a phenomenological description of such cases. According to Scheler, there is an affective feature involved that is absent in mere knowledge or in judgement, but which nevertheless does not amount to the experience of the 'real emotion' of the other. There is, he claims, a form of 'feeling the other's feeling, not just of knowing it', that is essentially cognitive, an affective grasping of a feeling's quality, which does not actually attain the status of an emotion of the agent herself (Scheler 1923a, 5). It is difficult to suppress doubts as to whether the idea of feeling the quality of an affect without feeling the affect itself is coherent.[5]

Scheler sees this affective-cognitive phenomenon as a capacity necessary for novelists, dramatists and historians – independently of whether they feel *for* their fellows, such that, for instance, thus 'enacting' the suffering of another leads to suffering on their own part. The simulationist explanation, according to which the affect in question is quarantined from other relevant mental states, may help to explain why such kinds of affective-cognitive experience need have no consequences for an agent's motivation. However, they do not in themselves explain why simulating the affect involved in suffering is not itself a form of suffering.

In her contribution to this volume, an extended discussion of the structure and status of simulationist accounts of perspective taking, **Heidi L. Maibom** points out that 'enacted' emotions are, unlike simulated beliefs, extremely difficult to keep 'quarantined' (Maibom this volume, 110f.). Such emotions, she says, have the tendency to linger. Indeed, she claims that feeling the relevant emotion for the person whose mental life was simulated ('affective empathy') is the default result of such affective-cognitive enactment. That is, according to Maibom, 'enacted affect' seems subject to three forms of spread: temporal spread, as it tends to linger; mental network spread, as it tends to overcome the mechanisms of mental quarantine; and personal spread, as it tends to focus on the person whose mental life was simulated. This raises the question of how these forms of spread might be explained. The first and second forms might be explained in terms of the difficulty of feeling an affect as one that is not

---

[5] Scheler uses two German compound terms to pick out this phenomenon: 'Nacherleben', which literally means 'experiencing-after', and 'Nachfühlen', literally 'feeling-after' (Scheler 1923a, 4f.). Although the latter term is an everyday German term used without any such necessary affective dimension, meaning simply 'understanding', Scheler is clearly employing it terminologically to pick out a form of affective simulation that he takes to have a purely cognitive function. He compares it to episodic visual or auditory memory, which involves episodes of seeing or hearing 'in one's mind's eye' or 'ear' and can thus be contrasted with forms of propositional memory. The comparison may have persuaded the translator of the 1959 English version of the book to render 'Nachfühlen' as 'visualized feeling' (Scheler 1923b, 9).