On *Moral Sentimentalism*

On *Moral Sentimentalism*

Edited by

Neil Roughley and Thomas Schramme

On *Moral Sentimentalism*

Edited by Neil Roughley and Thomas Schramme

TABLE OF CONTENTS

Part III: On Normative Sentimentalism

Part IV

INTRODUCTION

MORAL SENTIMENTALISM: CONTEXT AND CRITIQUE

NEIL ROUGHLEY AND THOMAS SCHRAMME

Michael Slote has long been one of the foremost contributors to discussions in moral theory. Both his work on consequentialism and his particular version of virtue ethics have been highly influential. In recent years – beginning with *The Ethics of* Care *and Empathy* (2007) and continuing up to his *Sentimentalist Theory of the Mind* (2014) – he has placed his various theoretical endeavours under the title of "sentimentalism". His key ethical work in this context is *Moral Sentimentalism*, published in 2010.

The present volume is an extended discussion of that work.[1] It begins with a summary by Slote of his book. There follow critical discussions of his meta-ethical and normative or aretaic sentimentalism. It concludes with Slote's response to his commentators. This introductory chapter provides a brief overview over sentimentalism in ethical theory (1) and of the central claims of Slote's version of sentimentalism (2), before outlining the main arguments developed in the commentaries and in Slote's replies (3).

(1)

A moral or ethical sentimentalist maintains that emotions play an essential role at some point in an adequate moral theory. That role might be epistemic, semantic, metaphysical, axiological or normative. A mild sentimentalism assigns emotions one such role, a strong sentimentalist

---

[1] The first four comments were presented and discussed together with Slote's replies at a workshop at Duisburg-Essen University in March 2013. The editors would like to thank Karina Derpmann for her help in organizing the workshop and for her invaluable support in the editing of this volume.

conception sees emotions as essentially at work in several of these areas, a maximally sentimentalist moral theory assigns them an essential role in all five. Within each area, the emotional component can in turn be seen as more or less decisive, maximal decisiveness being sufficiency. Metaphysical, epistemic and semantic sentimentalists claim that moral properties, moral judgements or the meanings of moral terms depend in some way on emotions. According to the axiological or normative sentimentalist, what is morally good or morally required is, or essentially involves emotions or emotional mechanisms.

Metaphysical sentimentalists generally see some kind of analogy, however loose, between secondary qualities and moral properties, taking the latter either to be response-dependent in a narrow sense (Prinz 2007, pp. 100f.) or to be properties that merit certain responses (McDowell 1985; Wiggins 1987; D'Arms/Jacobson 2000). Sentimentalism takes the responses relevant for the constitution of moral properties to involve emotional experiences or dispositions, as the secondary quality conception of colours takes the relevant responses to be qualitative perceptual experiences. Here, sentimentalism is one road to a form of internal realism about moral properties, according to which nothing would be morally good or morally wrong if there were no agents with dispositions to respond to certain occurrences in certain ways.

Epistemic and *semantic* sentimentalism were fused in emotivism, according to which (what are thought of as) moral judgements are in fact simply emotional occurrences, and the sentences that look to be making moral claims are merely expressions of the relevant emotions (Stevenson 1937). There is, however, no necessity that these two forms of sentimentalism go together. There are a number of epistemic sentimentalists for whom emotions can play the role that was traditionally played by 'intuition' in providing access to moral or other evaluative facts, a role that is analogous to perception's role in providing access to empirical facts (Brentano 1889; Tappolet 2000; Roeser 2011). Just as it doesn't follow that these facts need themselves be constituted by relations to emotions (epistemic without metaphysical sentimentalism), there is also no implication that sentences about such facts need express or refer to the emotions that may be needed in order to recognise them (epistemic without semantic sentimentalism). Contemporary expressivist conceptions may or may not take the non-cognitive attitude or attitudes expressed by moral sentences to be affective. Schroeder, for instance, gives no indication that the special attitude of 'being for' need be understood as involving affect (Schroeder 2008). In this respect, Gibbard's 'norm

acceptance' seems similarly agnostic, although Gibbard is an epistemic sentimentalist in taking the norms thus accepted to involve warrants for guilt and resentment (Gibbard 1990, pp. 47ff.). Blackburn has claimed that emotions are centrally, but not exclusively involved in our morally normative judgements (Blackburn 1998, pp. 8ff.). Whether or not emotions are involved in the attitudes taken to be expressed by ethical sentences, the doctrine has no metaphysical implications: even emotion-expressivists may well be, and expressivists generally are metaphysical non-sentimentalists (Gibbard 2003, p. 181).

Finally, according to *axiological* or *normative* sentimentalism, the reason why certain actions, persons or states of affairs are good or required lies in their relation to certain emotions. There are three notable such theories. First, hedonism is the classical position for which that relation locates the relevant emotions causally downstream from the action being judged: what makes an action good is its proclivity to bring about certain affective states. Hedonistic utilitarianism is the classical moral offshoot (Hutcheson 1725-38, II. III. VIII; Bentham 1789, I. III-VI). At least hedonism is such a position if either the 'pleasures' and 'pains' that are evaluatively or normatively relevant are all forms of joy and grief, for instance, the grief at one's physical pains rather than those pains themselves, or if one, somewhat idiosyncratically, conceives physical pains and pleasures as themselves emotions. Second, an ethics of compassion, according to which the moral value of an action is determined by the compassion of the agent expressed in their action (cf. Schopenhauer 1840, §16), is another historical variant, one which locates the relevant emotional mechanism causally upstream from action. Third, virtue ethics can be thought of as entailing a sentimentalist axiological dimension. As a theory of excellence of character, virtue ethics picks out not merely dispositions to act, but dispositions to react emotionally to events and the behaviour both of oneself and of others (Hursthouse 2001, pp. 108ff.). Whereas certain variants of normative ethics may be indifferent as to the affective states of those who conform to norms in their actions, virtue ethics can hardly avoid taking emotions as central objects of its aretaic evaluations. Whereas metaethical sentimentalism in its various forms is the subject of intense debate at present, strong forms of axiological or normative sentimentalism have been much more rarely advocated in recent philosophy.

(2)

Michael Slote is one of the few philosophers who have bucked the latter trend. In his *Ethics of Care and Empathy* (2007), he proposes "a general criterion of right and wrong action", according to which actions are morally impermissible if and only if "they reflect or exhibit or express an absence (or lack) of fully developed empathic concern for (or caring about) others on the part of the agent" (Slote 2007, p. 31). This proposal, which modifies the one advanced in *Morals from Motives* (2001), where "benevolence or caring" is named as the criterially decisive motivational feature (Slote 2001, p.38), recurs in *Moral Sentimentalism* (Slote 2010, p. 93). In each version, right-making mechanisms are emotional features of the agent located causally upstream relative to her actions—assuming, that is, that "reflect", "exhibit" and "express" are causal notions. Both *Morals from Motives* and *The Ethics of Care and Empathy* are works of axiological and normative sentimentalism, where the criteria for the latter types of judgement are derived from the criteria for the former. More precisely, Slote's "agent-based" conception makes aretaic evaluations, that is, value judgements concerning persons, primary and sees deontic judgements as derivative (Slote 2001, p. 4). The terminological slide between the two works corresponds to the strengthening of a dimension of partiality in the criterially decisive attitude. The claim that emotional sensitivity to proximity of various kinds is evaluatively relevant is "tentative" at first, seeming expressible by means of Hutcheson's term "benevolence" (Slote 2001, p. 137), but becomes a central claim in *The Ethics of Care and Empathy* and *Moral Sentimentalism*, in both of which a quasi-Humean notion of empathy takes centre stage.

Slote's *normative* project works with a combination of reference points that is unusual within the landscape of contemporary moral philosophy. First, whilst the project belongs under the rubric of virtue ethics, it connects not primarily with the Aristotelian tradition, but with that of the eighteenth-century British moralists. Second, it draws strongly on work done in the much more recent tradition of feminist care ethics. Slote interprets the notion of care in terms of empathy or empathic concern, which he connects with features of Hume's grounding of the 'natural virtues' in what in the eighteenth century was called 'sympathy'. Common to both strands of thought is not only the emphasis on the criterial importance of emotional reactions generated by feeling with, or for the other, but also the claim that the relevant emotional mechanisms involve a level of partiality towards particular others.

One corresponding noteworthy feature of Slote's normative view, according to which the constitutive structure of empathy provides a blueprint for the content of moral obligations, is the claim that we have stronger duties towards persons who are in one way or another closer to us than to those who are in the relevant respects more distant. Perhaps the view's most significant feature is the claim that the relevant structures of our emotional dispositions can account for the entirety of our true moral judgements (Slote 2007, p. 2; 2010, p. 9). This distinguishes it from other uses of care ethics in normative moral theory such as that of Lawrence Blum, for whom morality is not unitary, but consists of relatively independent partial and impartial spheres (Blum 1980, p. 9). In contrast, Slote argues that evaluations of the phenomena traditionally seen as the province of impartialist morality, including those with political dimensions, should also be grounded in empathy. The topics taken in *The Ethics of Care and Empathy* and *Moral Sentimentalism* to be amenable to such an empathy-based normative theory include promises, lies, respect, autonomy, property, liberalism, distributive and global justice. Respect, for instance, is shown, Slote claims, "if and only if one exhibits appropriate empathic concern" (Slote 2007, p. 57; 2010, p. 111); autonomy is defined causally as just what such respect tends to produce (Slote 2007, 61; 2010, p. 107) and a law is said to be just "if it reflects or expresses empathically caring motivation towards their compatriots on the part of the legislative group that is responsible for passing it" (Slote 2007, p. 95; 2010, p. 126).

Finally, Slote doesn't only argue that the *extension* of the objects of empathy-based normative judgement can be comprehensively expanded relative to the standard care-ethical conception. He also claims that the *contents* of many of our moral judgements turn out to be quite different to those traditionally associated with non-rationalist moral philosophy. Where a series of authors from Hutcheson to Rawls have assumed that non-rationalist positions must be consequentialist, Slote argues that an empathy-based sentimentalism is the best candidate to make sense of everyday deontological intuitions (Slote 2007, pp. 42ff.; 2010, pp. 22ff.).

If Michael Slote is one of a small contemporary minority of strong normative sentimentalists, his *metaethical* aims fit into a large and variegated landscape of conceptions that establish significant relations between moral properties, moral judgement, moral meaning on the one hand and emotional dispositions or episodes on the other. However, within this landscape Slote's views, delineated in *Moral Sentimentalism*, are

again characterised by an originality that sets them apart from the rest of the field.

The key metaethical claim advanced by Slote concerns the metaphysics of axiological and normative moral properties (between which he sees no substantial distinction). According to this claim, "moral goodness (or rightness) is whatever feelings of warmth directed at agents and delivered by mechanisms of empathy are caused by", where such feelings are constitutive of moral approval (Slote 2010, p. 61). The mechanisms that fix the reference of either axiological or normative moral predicates are at work causally downstream relative to the properties they pick out. For Slote, this quasi-Kripkean conception is superior to the response-dependent or ideal observer-based conceptions that are more usual within sentimentalism. His main reason for this claim is that these rival theories provide an insufficiently objective conception of the object of moral judgements, as such conceptions, in his view, make moral claims appear to be about the reactions of observers as much as about the actions observed (Slote 2010, p. 67). There is a sense, then, in which Slote's metaphysical view is not sentimentalist at all, as the reference to emotions in his definition is only a reference to the mechanism that picks out the metaphysically decisive properties, rather than a reference to constituents of the properties themselves. There is nothing "secondary" about the qualities picked out here. As, however, the relevant properties could not play the role they play *as ethical properties* without the empathic mechanisms Slote names, emotions do remain a feature of their metaphysics.

In contrast with much work in analytical metaethics, Slote explicitly side-steps the task of providing a semantic analysis of moral terms (Slote 2010, 51, 70). Although he also offers no analysis of moral judgement, he is concerned to argue that making moral judgements necessarily requires empathy, a claim that at present has more detractors (Prinz 2011a, 2011b; Goldie 2011; Maibom 2014b) than advocates (cf. Roughley forthcoming) within sentimentalist metaethics.[2] This necessity derives, Slote thinks, from the fact that fully understanding a claim about moral rightness or wrongness is only possible for an agent who has experienced feelings of the kind generated by the empathic mechanisms at work in approval or disapproval. A decisive component of his argument is that the type of affect involved in approval and disapproval has a unique phenomenology,

---

[2] On the various relationships between empathy and morality, cf. Maibom 2014a and Roughley/Schramme forthcoming.

which he picks out by means of the terms "warmth" and "chill" (Slote 2010, pp. 34ff.). Slote assumes that the amorality of psychopaths grounds in their inability to make moral judgements and claims that this inability derives in turn from their inability to experience the kind of affect constitutive of approval and disapproval (Slote 2010, pp. 54f.).[3] In a further twist of what one might think of as a sentimentalist hermeneutic spiral, Slote then argues that a correct moral epistemology, involving insight into the experiential precondition of moral judgement, is itself dependent on the epistemologist herself having had the relevant kinds of experiences (Slote 2010, p. 76). Without such experience the moral epistemologist cannot understand the key notions of moral goodness or rightness. Slote's epistemic sentimentalism thus extends to the preconditions of moral epistemology itself.

Slote's comprehensive sentimentalism is rounded off by the claim that there is an intimate – indeed, a priori – connection between moral metaphysics and normative moral theory. The only candidate that could do the causal work demanded by the metaphysical model is, he says, "agential warmth", which is presumably equivalent to the presence of fully developed empathic concern (Slote 2010, p. 61).

(3)

After a summary by **Michael Slote** of the main claims made in *Moral Sentimentalism*, there follow discussions divided into two groups. The first set discusses Slote's metaethics, the second his normative ethics. Where the former contributions concentrate primarily on the work that gives this collection its title, the latter also draw on *The Ethics of Care and Empathy*.

(3.1)

In the first metaethical contribution, **Neil Roughley** discusses Slote's claim to build up an understanding of morality from non-moral components, specifically from the concepts of approval and disapproval, themselves constituted by the empathic interaction of — "warm" or "chilled" — affective states of observers, agents and persons affected by actions. Roughley casts doubt on whether approval involves second-order empathy, as Slote claims. First, it seems unclear why what is approved of

---

[3] Detailed discussions of the concept of psychopathy and its relevance for understanding morality are collected in Schramme 2014.

is necessarily empathically derived motivation. Whilst there is certainly an argument that concern for others is the kind of motivation we essentially approve of where agents' behaviour affects others, there appears to be no argument why such concern should necessarily have its source in empathy. Roughley wonders whether the connection might be conceptual or causal, and rejects both possibilities.

Roughley then goes on to consider what the mechanics of approval might precisely be if the core of approval is taken to be empathically constituted. He argues that, if approval is essentially empathy with empathy, it looks like it is impossible for an observer to approve of an action that doesn't generate "warmth" — presumably relief or gratitude — on the part of the person affected. In order to solve this problem, Slote could, first, give up on the idea that the motivation approved of in the agent has to involve an affective experience of "warmth" on the agent's part. If the heart of an observer is warmed by the sight of the action, his characterisation of the agent as "warm-hearted" would simply be a metaphor, or perhaps projection of his own affective state. As this move stands in explicit contradiction to Slote's theory, a second option seems more likely: that the concept of empathy at work does not require that empathisee and empathiser feel the same emotion. It might suffice for empathy that the latter feels what the former would feel under certain circumstances. However, such a counterfactual conception threatens indeterminacy. It could perhaps be repaired by stipulating that what is felt by the empathiser is appropriate to the situation of the empathisee, but this move introduces a normative standard, which threatens Slote's naturalist sentimentalism.

In the last part of his article, Roughley argues that there are perhaps even more serious problems for Slote's conception of disapproval, as it is unclear how a parallel with the idea of second-order empathy might be spelled out. He discusses the ideas that disapproval might be empathy with partial empathy, projection of chill, empathy with cold-bloodedness or lack of empathy with lack of empathy. He rejects each of these possibilities. The article concludes by proposing that the sentimentalist should focus on the Strawsonian reactive attitudes, understanding indignation as vicarious resentment, that is resentment transferred by empathy. Roughley suggests that such a move might allow an understanding of empathy's role in what one might think of as a specific form of disapproval, the type that appears to be involved in our specifically deontic judgements.

Thomas Schramme addresses Slote's understanding of empathy as the "cement of the moral universe". Schramme agrees that empathy has a foundational role in morality. However, the role he is concerned to assign it lies on a somewhat different level to the metaphysical and conceptual level seen by Slote as decisive. This alternative function consists in enabling people to become moral persons, i.e. persons able to see and endorse the normative pull of morality. Schramme first analyses the concept of empathy, distinguishing between cognitive and emotional variants, or dimensions, of the concept, before going on to introduce the idea of empathy as a skill. Being able to emphasize with others, Schramme claims, brings with it the general moral capacity to care about others.

Schramme then pursues this thesis by discussing the problem of amorality. Psychopaths seem to be real-life exemplars of amoral persons, so it is interesting to learn from empirical findings that a distinctive deficit of psychopaths is indeed that they lack empathic concern for others. This establishes the case for seeing a foundational role for empathy in morality. Yet, the very fact that empathy allows for an emotional and a cognitive element arguably undermines any rigid distinction between moral sentimentalism and moral rationalism.

Where Schramme argues that a more precise conceptualisation of empathy causes problems for an ethical theory that claims to be sentimentalist in any emphatic sense of the term, **Karina Derpmann** argues that there is a better candidate for an empathy-based sentimentalism than Slote's modified Humean proposal. According to Derpmann, Slote underestimates the epistemic resources of Adam Smith's *Theory of Moral Sentiments*. Her strategy in the article is to discuss Slote's objections to Smith's conception of approval, in order to show, first that the objections can be rebutted and, second that Slote's theory has problems of its own in the relevant areas. According to Slote, an empathically generated concord of feelings between an observer and an agent is, pace Smith, neither necessary nor sufficient for the former's moral approval of the latter. Similarly, affective discord resulting from a striving for empathy is neither necessary nor sufficient for disapproval. Self-disapproval in the face of agreement in sentiments shows both that concord is insufficient for approval and that discord is unnecessary for disapproval. Judgements that another agent's behaviour is supererogatory show that affective discord is insufficient for disapproval and that concord is unnecessary for approval. Or so Slote claims.

Derpmann's response begins with a discussion of how an empathy-based conception can make sense of self-approval and self-disapproval. A Smithian theory, she argues, has no problems with such self-directed evaluation, as it sees all approval and disapproval as necessarily mediated by adopting the perspective of an impartial empathic observer. Slote's theory, in contrast, runs into serious difficulties with such self-evaluations because it needs a phenomenological distinction between empathic warmth and empathically generated empathic warmth, a distinction that phenomenology cannot deliver. Moreover, once the mediating role of the impartial empathiser is factored into Smith's theory, it has no problem with self-disapproval in the face of affective concord that is unfiltered by impartial empathy. In discussing Slote's example of supererogatory action, Derpmann first points out that in such a case there may well be affective concord in spite of the fact that the approving person might, for one reason or another, not perform the action they deem supererogatory. Further, the role of the impartial empathiser is here again decisive, as attitudes mediated via the perspective thus constituted can diverge from those that are unfiltered.

The possibility of supererogatory action is also a central concern in the contribution by **Matthias Gottschalk**, who discusses problems relating the metaethical and normative components of Slote's theory. The source of Gottschalk's worries is the fact that empathy comes in differing degrees. As Slote explicitly rejects ideal observer theories, this raises the question of how such a quantitatively variable phenomenon can be used to fix the reference of the key moral terms. The suspicion expressed by Gottschalk is that Slote's quasi-Kripkean model of reference-fixing entails a normative ethics that is either over-demanding or relativistic, although Slote seems unprepared to accept either alternative.

On the one hand, the normative consequences would appear to be over-demanding if we take it that more empathic persons are able to attain awareness of obligations that remain hidden to less empathic agents. If this is a feature of empathy's role, Slote may have to concede that expansive normative demands, such as those advanced by Peter Singer, follow from his own moral metaphysics. Thus understood, Slote's theory may, again in contrast to its explicit aims, not be able to make room for supererogatory action. The latter would only be possible where particularly empathic persons both feel warmed by an action under certain circumstances, but would not feel chill if it were omitted. Gottschalk questions whether such empathisers could rationally be the bearers of both these dispositions.

On the other hand, if the theory doesn't dignify particular empathisers as epistemically privileged relative to our moral obligations, the alternative appears to be a relativity of obligations to the empathic capacities of particular individuals or at least to those cultivated in particular cultures. Gottschalk takes Slote's repeated use of the expression "fully developed empathy" to be an attempt to solve these problems. He distinguishes five possible interpretations of the expression, which might pick out cases of empathy that is statistically normal; potentially realisable under given psychological and cultural conditions; potentially realisable for fully developed human persons; realised by persons with the correct measure of empathy; or cases that cross a minimal threshold, beyond which differences are irrelevant. Each of these interpretations is rejected. Finally, Gottschalk argues that a theory such as Slote's needs to pay more attention to the differing skills that contribute to, or count as empathy. He mentions the difference between the imagine-self and imagine-other perspective, as well as that between Hoffman's associative empathy and mediated associative empathy. Empathy in any of these modes, Gottschalk claims, may provide a model for different capacities for moral judgement that yield different material judgements. Slote owes us justificatory criteria for privileging one of these forms of empathy or a particular model of their interaction if this is to be the mechanism that fixes our moral obligations.

In contrast to the four previous contributors, who proceed from the assumption that empathy may have a significant function in metaethics, **Herlinde Pauer-Studer** casts doubt on this basic assumption, questioning two specific roles she sees it as playing in Slote's theory. The first concerns the status of the emotion-transferring mechanisms at work in the observer who is empathically "warmed" or "chilled" by agential empathy. Pauer-Studer takes it that such mechanisms are intended to deliver both normative and motivating reasons. However, she questions whether it can deliver either because Slote's focus on the feelings of observers appears to commit him to a third-person conception of morality, according to which reasons are agent-neutral. Such a conception, she argues, rules out an understanding of 'oughts' as normatively binding and as motivating their addressees, as it can make no sense of the first-person deliberative perspective. Only such a perspective, she claims, can make sense of moral injunctions as binding the will, whereas a third-person conception makes our relation to reasons purely epistemic.

The second role of empathy in Slote's theory rejected by Pauer-Studer is that of the reference fixer for terms such as 'morally good' and 'morally right'. She argues that Slote provides no justification for his central claim

that moral rightness is whatever it is that causes feelings of warmth through empathic mechanisms in observers of human action. Where Slote sees the plausibility of the claim as lying in the plausibility of the overall theory, Pauer-Studer argues that this is the decisive premise on which far too much in the theory depends. Moreover, she claims that there are good normative reasons for caution here, as empathy may motivate infringements of rights, be blind to justified retaliation or tend to cement traditional gender roles.

(3.2)

The normative topics dealt with in *Moral Sentimentalism* are drawn from interpersonal morality – for instance, respect for autonomy and paternalism – and social or political morality – specifically, social and distributive justice. The three comments in the present book focus on the interpersonal sphere of morality.

Monika Betzler queries whether empathy can ground respect for other persons' autonomy. There are many examples in which an empathetic perspective might lead us to focus mainly on the welfare of other people, but not necessarily on their own choices. Paternalism indeed relies on such an attitude towards protecting the well-being of other people, if necessary even against their own will. Respect for autonomy seems to take us away from the concrete person we encounter in morally salient situations to the generalized other as a specimen of creatures making their own choices. In order to evaluate Slote's project of basing normative ethics on empathy, and more specifically to ground respect in empathic concern, Betzler makes two distinctions that structure her discussion. First she distinguishes cognitive and affective empathy, before, second, distinguishing the claims that either form of empathy may be conceptually or causally related to respect.

These claims are discussed as applied to an example of paternalism discussed by Slote, in which a parent takes a child to the doctor against its will, because it needs treatment. Slote claims that this kind of paternalism could exemplify respect for the child's nascent autonomy. Betzler objects that affective empathy alone cannot account conceptually for respect for autonomy, because it does not seem to track what is good for another person. Obviously, people can be wrong in what they feel is good for them and affective empathy might not help us find out what is most in their interest. Cognitive empathy, according to Betzler, cannot conceptually ground respect for autonomy either, because it does not in itself involve

caring for the other. Moreover, the combination of both aspects of empathy, cognitive and affective, seems unstable.

Betzler discusses what she sees as an alternative: there might be no conceptual connection but a causal contribution of empathy to autonomy, a possibility that seems especially pertinent when we focus on the relational aspects of autonomy that have been emphasised in recent years. Autonomy is not something we have devoid of social relations. Yet Betzler again argues that it remains unclear what exactly empathy contributes in developing and maintaining personal autonomy. Indeed, she claims that empathy is developmentally neither necessary nor sufficient. Finally, she objects that Slote's conjoining of both conceptual and causal claims appears tainted by circularity.

In a paper that canvasses the possibility of extending Slote's sentimentalist framework, **Christian Budnik** focuses on the phenomenon of trust, which he takes to be an important ingredient of any account of morality. Budnik begins by adducing reasons why trust should be understood as an affective, rather than as a doxastic phenomenon. Two central reasons are, first, that trust involves a level of resistance to counter-evidence that goes beyond what would be expected if trust were a kind of belief and, second, that trust is peculiarly likely to be undermined by the investigation of the reasons that support it, even where the conclusion is that they do indeed provide such support.

In going on to propose a sentimentalist analysis of trust, Budnik distinguishes the questions as to the kind of affective attitude and as to its content. Both, Budnik claims, involve empathy: trust, he think, is a form of second-order empathy, where the truster has empathic access to the empathic concern of the trustee for the truster's own feelings. This conception both permits a distinction between trust and reliance and accounts for the fact that psychopaths seem as unfit for trust as they seem unfit for morality. Trust is, however, to be distinguished from Slotean moral approval by means of the distinction between third-person and second-person second-order empathy. Whereas approval consists in empathy for an agent who is empathic in her action relative to a third party, trust consists in empathy for an agent who is empathic in relation to oneself. Moreover, Budnik suggests, trust also appears ontogenetically prior to moral approval.

In the volume's final critical contribution, **Jörg Löschke** focuses on one of the key structural features of Slote's normative view: the centrality of

partial duties, a feature that an empathy-based moral theory can account for better than the standard normative theories – or so Slote claims. It is this claim that Löschke sets out to test. After querying the precise role Slote assigns to impartiality, suggesting that it may have at most a heuristic value, Löschke goes on to set up his investigation as a discussion of our reasons for partiality. He takes it that Slote understands moral reasons as facts that would bring about relevant empathic reactions in persons with fully developed empathy. Löschke spells this out as the claim that for Slote empathic reactions "constitute" or "ground" the relevant facts as moral reasons and that the strength of the relevant reaction also grounds the strength of the relevant reason.

The investigation then proceeds on the basis of a typology of reasons for partiality to be found in Slote's work. Löschke distinguishes six types: partiality deriving from spatial or temporal contiguity; from immediacy of danger; from shared lives; from family connections; from solidarity; and from shared nationality or cultural similarity. He then argues that, among the six types of partial duties, two appear justified within Slote's empathy-based framework, two appear equally well justified within an impartialist conception and two look not to be justified at all. If this is right, Slotean sentimentalism may be the best account of some of our partial duties, may be faced with equally strong explanatory candidates for some others and may even end up postulating further partial duties where there are in fact none. In last critical step, Löschke questions the plausibility of an empathy-based understanding of negative duties, according to which it would appear that we have weaker duties to refrain from harming outsiders than persons with whom we have a special relationship.

(3.3)

The final contribution to this volume is **Michael Slote**'s detailed and substantial response to the commentators. Slote both rejects certain criticisms and makes a number of new moves designed to accommodate other objections. Slote takes his reply to strengthen and extend the scope of moral sentimentalism.

In response to *Roughley*'s worries about the relationship between empathy and altruistic motivation, Slote alters part of the view developed in *Moral Sentimentalism*. In his response here, Slote now argues that the relation should be understood as conceptual, as the adoption of another agent's emotion necessarily requires taking in that emotion's intentionality, not merely the relevant affective experience. And the adoption of such

emotional intentionality, Slote assumes, brings with it corresponding forms of motivation. Slote also explicitly develops his conception of agential warmth further in response to Roughley's objections that agential empathy cannot plausibly involve picking up on warmth felt by an action's recipient. Agential warmth is now characterised as drawing on the warmth experienced in nurture. Slote further rejects both Roughley's various candidates for an understanding of empathy-based disapproval and his claim that Slote's sentimentalism cannot deal adequately with the deontic.

Slote agrees with *Schramme* that various dimensions of the concept of empathy require a more explicit working out than he undertook in *Moral Sentimentalism*. He agrees that the distinction between emotional and cognitive empathy ("simulation") is important, but insists pace Schramme that it is the former that is decisive for morality. The latter without the former may be possessed by psychopaths and, unlike what Schramme may be claiming, is, if not backed by emotional empathy, unable to explain the commitment to all humans counting morally. Slote rejects Schramme's attempt to undermine the distinction between sentimentalism and rationalism, arguing that the latter, certainly in its Kantian incarnation, claims to derive moral norms entirely from practical reason. However, Slote agrees with Schramme both that emotional empathy is decisive for being moral and that mature empathy and sympathy are conceptually inseparable – a point on which both authors agree in the face of Roughley's sceptical challenge as to the kind of connection between empathy and altruistic motivation.

Slote greets *Derpmann*'s claim that Smithian sentimentalism has farther-reaching resources than Slote acknowledges and than Slote's own sentimentalism possesses with a somewhat sceptical profession of interest. He encourages Smithians to elaborate their conception more systematically than has been done so far to make good Derpmann's claim. The central move canvassed by Derpmann, the mediation of empathic reactions through the unbiased empathiser may, Slote thinks, be promising, although he suspects that the move may not be genuinely Smithian and may, moreover, attenuate the actual connections between people, the importance of which Derpmann emphasises. Slote claims, further, that Smithian empathy is not emotional, but cognitive or "projective" – characterisations he tends to use synonymously – which may undermine the theory's claim to be genuinely sentimentalist. Finally, he disagrees with Derpmann's contention that a sentimentalism needs to make sense of self-approval and -disapproval, because, he argues,

approval involves a distance to the self that is only present in the evaluation of second or third parties.

Gottschalk's central claim, that Slote has serious difficulties reconstructing the distinction between obligatory and supererogatory action, is rejected by Slote. In contrast to what Gottschalk thinks, Slote sees no reason why especially empathic persons shouldn't be more lax in their dispositions to react emotionally to the actions of others than they are in their dispositions to empathise with those potentially affected by their own actions. Unlike what Kravinsky and Singer assume, Slote sees no morally problematic inconsistency in the possession of differing empathic dispositions on the levels of agential and second-order empathy. Slote does, however, concede, that his talk of "fully developed empathy" may have been misleading, as it may carry genetic implications unnecessary for his purpose. Nevertheless, the fact that there are difficulties specifying whether certain actions express full empathy is, he insists, not a weakness, but a strength of the theory, as this would explain the genuine problems we have judging certain actions.

In his fairly direct confrontation with the Kantian objections brought forward by *Pauer-Studer*, Slote sees the main bone of contention in the conception of moral agency: whereas, for the Kantian, moral normativity grounds in reflection on reasons on the part of the agent, Slote maintains that much of our moral agency can be covered by a quasi-Humean notion of natural virtue, the possession of which enables unreflective action that may be morally permissible or praiseworthy. In such cases, moral judgement will involve the third-person assessment of the agent, but, unlike what Pauer-Studer suggests, Slote makes no claims that such judgements constitute the whole of morality. Moreover, their motivationally internalist character results, according to Slote, from the fact that the second-order empathy that necessarily feeds into them in turn depends on the first-order empathy with those affected by the actions of the agent who is the object of the judgement. Slote also rejects the second strand of Pauer-Studer's objection to his empathy-based sentimentalism, which, he argues, both mistakes the object of the reference fixed by second-order empathy and misconceives the abductive character of the argument for his central metaphysical claim.

In his exchange with *Betzler*, Slote defends his claim that autonomy is both conceptually and causally related to empathy, placing his conception in a line with causal analyses of various concepts such as meaning, perception and justification. He also rejects Betzler's contention that

affective and cognitive empathy should be treated as separate candidates for these roles, pointing to the notion of mediated associated empathy, which he takes on from Hoffman. He denies Betzler's claim that the mother's motivation in taking the unwilling child to the doctor must be generated by sympathy, not empathy. The empathic motivation that brings the mother to act will, on the contrary, be derived from her vividly picturing what life may turn out to be like for her child in the future if she doesn't receive treatment. In spite of Slote's claims for the importance of proximity in determining strength of empathy, it can, he insists, be outweighed by seriousness of welfare issues. Finally, Slote responds to Betzler's bemusement that the parent who refuses to allow her son to ride a motorbike without a helmet should qualify as respecting his autonomy – rather than as caring more about his welfare than his autonomy. In reply, Slote seems to insist that, from the viewpoint of empathy, concern for serious welfare issues and respect for autonomy coincide.

Slote welcomes *Budnik*'s proposal to supplement his theory, agreeing that trust involves a form of second-order empathy with a structure close to the one Budnik delineates. However, he questions the conception that at least one of Budnik's formulations suggest, according to which trust is a matter of empathy with future empathy of the trusted. Slote, in contrast, claims that trust involves an agent "registering" that the other already has empathised with her. Slote also points out that it is possible to approve of someone's empathic concern for oneself, but remarks that in such a case, approval and trust would be mixed – a remark that raises the question of whether the two attitudes would then coincide or be separable components in an attitudinal mix. In the second part of his response, Slote brings further attitudes into play in sketching what he takes to be the empathy-mediated emotional constellation that is decisive for children's inculcation into morality, a constellation that also includes love and gratitude. He characterises the latter as the sympathy felt by the child for the sympathy or love it empathically registers in its parents or caregivers. Gratitude thus understood, Slote concludes, amounts to trust as analysed by Budnik.

Finally, in response to *Löschke*, Slote corrects the claim that he sees moral impartiality as developing out of partiality: in contrast to what most moral theorists claim, moral judgement is never, he thinks, strictly impartial. He welcomes Löschke's challenge to connect empathy and talk of moral reasons, claiming that empathy should be seen as reasons' conduit, providing access for an observer to the reasons an agent herself has. Slote then turns to the four types of partiality Löschke identifies as problematic for his theory. In response to the claim that shared lives and solidarity may

only provide epistemic reasons for partiality, Slote claims that the impartialist view behind such claims both appears procrustean and, moreover, needs to prove its plausibility as a whole theory when faced with the advantages of sentimentalism – a move Slote already makes in his response to Pauer-Studer. Slote also rejects the claims that blood ties and shared nationality don't provide reasons for empathy-based partiality. In both kinds of case, Slote names further considerations that may explain the reluctance to assign the connections in question normative force: in the first case, the love children need and, for example the undertaking of adults who adopt children, in the second case, the undertaking of a state that grants immigrants a new nationality. In both cases, Slote insists, there is nevertheless a normatively significant connection, one that is tracked by empathy.

PART I

CHAPTER ONE

SUMMARY OF *MORAL SENTIMENTALISM*

MICHAEL SLOTE

Moral Sentimentalism (MS)[1] attempts to revive eighteenth-century moral sentimentalism in a more systematic way than has recently been attempted. Sentimentalists have lately concentrated exclusively either on metaethics or on normative issues, but MS deals with both sorts of questions in terms of the basic unifying notion of empathy. Making use of recent work on moral development, it sees empathy as supporting sympathy and altruism, and it also details ways in which parents and teachers can help strengthen empathy and altruism in children.

MS claims that our empathic dispositions and our moral judgments are aligned in a way that is best explained by assuming that empathy enters into our moral concepts. Psychological studies indicate that empathy and altruism tend to be more strongly elicited by perceived danger than by danger one merely knows about; but we also think it is morally worse (e.g.) not to save a child drowning in shallow water right in front of us than not to make a contribution to Oxfam that would save some distant child we only know *about*. In addition, we don't feel empathy for our ongoing selves the way we can and do for other people, and that fact corresponds to our tendency to believe, for example, that it is wrong negligently to harm another person, but not morally wrong to do this sort of thing to oneself. In fact, there seems to be a pretty general correlation between our intuitive moral judgments and our familiar empathic tendencies; and MS argues that the best explanation of this correlation is that the very same empathy that makes us, say, more empathically responsive to perceived danger also enters into our moral concepts and judgments, so that we judge it to be *ceteris paribus* worse not to respond to such danger than not to respond to danger we merely know about. And

---

[1] All references in the text abbreviated with 'MS' refer to Slote 2010.

MS then homes in on the question of *how* empathy enters into those concepts/judgments.

Hume thought that moral judgments were based in prior sentimental attitudes of approval and disapproval, and MS agrees with this. But the book takes issue with Hume's particular theory of what approval and disapproval consist in. Hume's official view makes it hard to understand why we can't approve and disapprove of inanimate objects. But Hume also casually mentions another idea that I believe helps us toward a better understanding of the moral role of approval and disapproval. Hume notes that we can feel warmed by the warmth displayed in the actions of one friend toward another, and MS suggests that such warmth is an elementary form of moral approval: that approval is a kind of second-order empathy with empathy, a feeling of warmth felt in response to agential warmth. Similarly, though Hume doesn't mention this, moral disapproval can be identified with the empathic chill one feels at the cold-heartedness that someone displays in their actions or attitudes. We can then begin to explain how empathy enters into our moral judgments if we assume that approval and disapproval can be understood in the just-mentioned way and that they also enter into the making of moral judgments.

Now Hume offered a number of (mutually inconsistent) (proto-)theories of how approval and disapproval enter into moral judgments or utterances. But none of them allows moral judgments to have the kind of objectivity that most non-philosophers think they have and that ethical rationalism considers them to have. And MS seeks to offer an account of moral terms that allows moral judgments to be fully objective. It does so by making use of Saul Kripke's idea of reference-fixing. To understand what we mean by a term like "(objectively) red" we have to see that its reference is fixed in relation to the experience of redness: objective redness is whatever it is out there in the world that typically causes and is perceived via sensations of redness. Kripke's account of redness results in *a posteriori* but necessary truths about the physical character of red objects, and an exactly parallel theory of moral predicates would make basic moral judgments like "torturing babies for fun is wrong" also come out *a posteriori*.

But as rationalists going all the way back to Plato have noted, basic moral judgments *seem* to be *a priori*, and MS therefore seeks to find a way in which such claims can be both objective and *a priori*. It argues, in particular, that if we make a smallish modification in Kripke's reference-fixing approach (as applied to moral predicates), the sentimentalist can

obtain that desirable result. If we think of moral goodness or rightness, not as what (out there) accounts for most agent-directed feelings of warmth, but as what accounts for the *empathically derived* feelings of warmth we have toward moral agents, we can simply say that moral rightness/goodness is what accounts for approval (and that this reflects our understanding of "right" and "morally good"). And given what was said above about approval, it follows *a priori* that good actions are empathically warm and caring ones. The same point can be made about disapproval and the meaning of the term "wrong." So a semi-Kripkean sentimentalist theory of moral meaning allows one to accommodate and explain the a priori and objective character of moral judgments—and as a result some rather typical rationalist objections to sentimentalism don't seem to apply in the present instance. In addition, this approach allows one to understand the motivating force of moral judgments, since anyone who makes them has to be able to feel what others feel, and empathy, say, with the plight of another person, understandably does typically give us some motivation to help that other person. And if psychopaths lack the capacity for this kind of feelingful empathy, then on the present view they lack the capacity to understand or make genuine moral judgments (the way the congenitally blind don't fully understand judgments about redness).

MS anchors moral judgment in certain empathic reactions. The judgments aren't *about* such reactions, but the reactions *point us toward* what moral judgments are about. But one worry that can arise here is the possibility that, contrary to received understandings of Hume's views, sentimentalism allows for the possibility of deriving an "ought" from an "is". Despite my own extreme initial misgivings on the subject, MS uses the theory of moral language it offers to argue that judgments at least of right and wrong can be derived from "is"es describing certain kinds of approval or disapproval. But this may seem less problematic if one realizes that on the account offered, attributions of approval and disapproval depend on emotionally charged empathy just as much the making of moral judgments does.

MS goes on to explain how that account can help us understand the role that moral judgments or precepts play in the moral life. Sentimentally-understood moral claims express strong attitudes and feelings, so if someone tells us, for example, that what we are thinking of doing would in fact be wrong, we may well take in that opinion via a kind of empathic osmosis and be influenced not to do the act in question. (This draws on ideas of Hume and of Charles Stevenson.) In addition, a sentimentalist metaethics can help explain the role of moral precepts, rules, etc., in moral

education. If a child hears a parent call something wrong, they can empathically pick up on the negative attitude the parent is expressing and become more capable of being chilled or warmed by the kinds of actions that chill or warm their parent. This can lead them toward understanding moral judgments and reinforce their moral tendencies. Kantians, of course, have their own way of understanding moral education, but the most psychologically grounded Kantian approach to moral education, Lawrence Kohlberg's, has an impossible time dealing with the development of moral motivation without (in the end) bringing in empathy, and MS offers a picture of moral education as based in empathy that is more thoroughgoing and plausible than anything available from Kantianism or rationalism more generally.

However, Kantians and other rationalists have claimed that feelings are unreliable as a basis for morally acceptable action because they sometimes point us toward the wrong actions and because, even when they point in the right direction, their force often fades before they can efficaciously ensure that we do what is right. But if our sentimentalist metaethics is on the right track, then empathy, fully empathic concern for others, does systematically point us in the right direction, and MS argues that the criterion of right action it offers as *a priori* valid—namely, that actions are wrong iff they express or exhibit a lack of full empathic concern for others—generally yields intuitively acceptable normative judgments about particular cases. Moreover, full empathy doesn't tend to flag or fail before it can yield right action. MS describes a number of cases where weakened concern for another person tends to revive and restrengthen before one actually does anything wrong; and it argues that this will in fact happen with anyone who isn't lacking in empathic concern for others or for some particular other (like their own mother). It also argues that the kinds of negligent actions we tend to consider morally wrong can be systematically traced to motivational failures. This is virtue ethics, but of a kind that emphasizes our empathic connection with others as a criterion of normative morality.

But MS not only seeks to offer a picture or account of right and wrong action, but has significant things to say about such preferred rationalist/Kantian concepts as deontology, autonomy, respect, and justice. I argue, for example, that our deontological "dispositions" can be modeled on what psychologists and others have said about our "preference" for what we perceive. Just as the actual perception of danger arouses our empathic concern more readily or strongly than mere knowledge of such danger, so too does the fact that we might cause a certain harm have more

empathic immediacy for us than the mere fact that we might allow a certain harm to occur. We emotionally flinch from harming more than from allowing harm, and this tendency, like our empathic preference for what is perceptually immediate, antedates or is independent of the making of moral judgments. MS claims that sentimentalism offers a more articulate account and justification of deontology than anything one can find in Kant or in the writings of other rationalists.

Kantians emphasize human autonomy in a way that stresses our independence or separateness from others, but recent feminists have claimed that autonomy is best understood in relational and historical terms. And MS attempts to advance this discussion by arguing for its own distinctive notion of respect (for autonomy) and using it to clarify the idea of relational autonomy. Those who don't listen to their own children and who seek always to impose their own ideas of what is good for them (the children) show a lack of respect for the children, and the main criticism that Carol Gilligan famously made of patriarchy, namely, that it doesn't listen to women's "voices", also illustrates and points to the lack of respect with which patriarchy treats girls and women. So MS argues that empathy with the point of view, or ideas and aspirations, of the other constitutes or is a sufficient condition of respect for that other. This disagrees with traditional and Kantian ideas about respect, according to which respect (for individual autonomy) requires that one grant each individual various strong rights against interference by others. But the care-ethical sentimentalist view of respect that MS develops will insist—against these more traditional views—that we may justifiably intervene against certain individuals (e.g., people who abuse their wives or who seek to give expression to harmful forms of hate speech) if and when our doing so reflects empathy with the points of view of everyone involved in a given situation and a recognition that certain people have more at stake than others in that situation.

Autonomy can then be understood as what respectful treatment tends to produce. Women who lose their voices because patriarchy devalues their distinctive point(s) of view are deprived of autonomy *because* they haven't been treated respectfully. And MS argues that we should understand autonomy in terms of its historic and psychological underpinnings rather than in more metaphysical or individualistic/atomistic terms. It also seeks to give a normative account of justice in care-ethical terms. Justice is a moral concept, and a sentimentalist can use the same normative criterion for justice as s/he can for rightness or wrongness. Institutions, social practices, and laws can be considered just if they don't reflect a lack of