7 Moral Sense Theories and Other Sentimentalist Accounts of the Foundations of Morals

Christel Fricke

Interest in what has been called a 'moral sense' originated in the late seventeenth century, as part of a debate about the nature of human beings in general and humans' moral nature in particular. Participants in the debate collectively rejected four commonly held views of human morality. They found (1) the Cambridge Platonists' moral rationalism and (2) Gershom Carmichael's (and others') natural law theories of morality too remote from actual processes of moral judgement and decision making; (3) they rejected Thomas Hobbes' psychological egoism as excessively reductive; and (4) they found moral relativism – which Thomas Burnet erroneously attributed to John Locke – objectionable on normative grounds, since they were committed to the defence of moral universalism.

Their ambition was to defend a normative moral theory that was empirically informed, sentimentalist in kind and universalistic. Some of them, such as Shaftesbury and Hutcheson, claimed that this aim could only be reached by assuming that a moral sense was an intrinsic part of human nature; others, such as Hume and Smith, rejected this claim without however giving up the project of a sentimentalist and universalistic moral theory. The central questions of the debate were, What exactly is the nature of a moral sense? What impact does it have on humans' moral beliefs, desires and actions? And how can it provide the foundation for a universalistic account of human morality?[1]

SHAFTESBURY

Anthony Ashley-Cooper, Third Earl of Shaftesbury (1671–1713), is commonly considered to be the first philosopher to propose a moral sense theory. But he did not coin the term; neither was he the first to develop the idea that people are naturally provided with a sense for distinguishing between moral good and evil, a sense somewhat similar to perceptual

senses such as vision or smell. An early use of the term 'moral sense' can be found in the writings of Henry More (1614–87). However, More's usage of the term 'moral sense' relates to the 'meaning' of a text.[2]

Thomas Burnet (1635–1715), in his three *Remarks upon an Essay Concerning Humane Understanding* (1667–9), objects to various doctrines he believes he has found in Locke's *Essay*, including Locke's rejection of 'natural conscience' as the source of humans' moral understanding and his apparent endorsement of moral relativism.[3] Burnet claims the existence of a moral sense as an intrinsic part of human nature without actually calling it thus. But he comes very close: he ascribes to human beings a 'sense of those moral differences', namely the differences between 'Good and Evil', which we rely on for 'distinguishing one thing from another in Moral Cases' (*Remarks*, 65). This sense is both similar to and different from humans' perceptual senses: whereas the latter are 'outward' senses, the former is an 'inward sense' (*Remarks*, 64–5). Both types of sense, however, are sources of information. Outward senses tell us about the perceptual properties of external objects, while inward senses tell us about the evaluative or moral properties of actions. The 'sense of moral differences' functions as a 'Spring and Motive of our Actions' (*Remarks*, 66). Burnet also calls this sense 'natural conscience' (*Remarks*, 66), distinguishing it from Locke's idea of 'conscience' as 'nothing else but our Opinion of our own Actions' (*Remarks*, 58). According to Burnet, moral sentiments have a two-fold function: they inform moral beliefs and guide moral volition and action. Our moral beliefs are true in the same way in which perceptual beliefs are true. And the motivational authority of the sentiments informing these beliefs is universal; it originates in the inward sense for moral distinctions that all humans naturally possess.

We do not know for sure whether Shaftesbury was familiar with the debate between Locke and Burnet.[4] In his *Sensus Communis, an essay on the freedom of wit and humour in a letter to a friend* (1709),[5] he makes explicit that he means to reject Hobbes' motivational egoism.[6] He points to 'a thousand other springs, which are counter to self-interest', namely 'passion, humour, zeal and faction' as well as 'the better and enlarged affections', such as 'kindness or generosity', 'pure good nature and friendship' and 'any social or natural affection' (*Characteristics*, 53–4). These affections reveal a 'common sense or the love of mankind' (*Characteristics*, 57).

Ten years earlier, in his *Inquiry Concerning Virtue or Merit* (1699), Shaftesbury takes his starting point from the question of how religious

devotion and moral virtue are related. He denies that they are 'insepar-able companions', relying for the defence of this view on an open ques-tion argument: 'If we are told that a man is religious, we still ask, "What are his morals?"' (*Characteristics*, 163). In this *Inquiry*, he argues that moral virtue is a matter of having a 'natural moral sense' (*Characteristics*, 180); however, he does not make explicit use of this term more than once in the *Inquiry*. Instead, he uses a number of other terms, including that of the 'natural temper' of a 'good creature':

> *a good creature is such a one as by the natural temper or bent of his affections is carried primarily and immediately, and not secondarily and accidentally, to good and against evil.* (*Characteristics*, 171, Shaftesbury's italics)

The good and evil to which people respond with feelings of approval or disapproval by virtue of their moral sense is 'what is good or ill in the species or in society' (*Characteristics*, 177); it concerns the actual consequences of an action. The 'public interest' is by no means incompatible with an agent's own interest; both interests are 'con-sistent' and 'inseparable' (*Characteristics*, 193). By attributing a natural moral sense to all people, Shaftesbury implies that all people are virtuous by nature. While their natural virtuous disposi-tion can be corrupted by external influence, for example by 'super-stition' or 'ill custom' (*Characteristics*, 171), this influence cannot completely destroy their virtuous nature (*Characteristics*, 177). Nor can it count as an excuse from acting morally. Even though Shaftesbury does not see virtuous behaviour as opposed to an agent's pursuing an 'interest of the private nature' (*Characteristics*, 193), he speaks of an 'obligation' to be virtuous (*Characteristics*, 192). This view makes him vulnerable to the objection of not rejecting rational egoism after all:

> to be well affected towards the public interest and one's own is not only consistent but inseparable.[7] (*Characteristics*, 193)

Shaftesbury's moral sense has a complex function. It is a disposition to have first-order sentiments in response to the conse-quences of an action. These sentiments inform moral beliefs; they also motivate an agent to act according to her or his moral beliefs. But the moral sense is also a disposition to have second-order sentiments on which people rely for assessing first-order motivational sentiments, both their own and those of other people:

> In a creature capable of forming general notions of things, not only the outward beings which offer themselves to the sense are the objects of the affection, but the very actions themselves and the affections of pity, kindness, gratitude and their contraries, being brought into the mind by reflection, become objects. So that, by means of this *reflected sense*, there arises another kind of affection towards those very affections themselves, which have been already felt and have now become the subject of a new liking or dislike. (*Characteristics*, 172, my italics)

The moral sense allows humans to evaluate affections and actions and to endorse or reject them reflectively. It is this sense that distinguishes humans from other animals whose behaviour is guided by instinct (see also *Characteristics*, 175). As a 'reflected sense', the moral sense is essentially a sense of a 'rational creature' (*Characteristics*, 177, see also 182). As a 'natural and just sense of right and wrong' (*Characteristics*, 177), it informs true moral beliefs.[8] Shaftesbury compares the moral sense both to our external perceptual senses (*Characteristics*, see 58) and to our inner sense of beauty, a sense that allows us to evaluate aesthetically what we have first perceived through our external senses (*Characteristics*, 172). Since the moral sense allows people to assess their own affections and actions, Shaftesbury also calls it 'conscience', underlining its immediate and inevitable impact on the mind of a 'sensible creature' (*Characteristics*, 209).

There is debate about how to read Shaftesbury's moral sense theory. Are people's affective responses to the actions they perceive constitutive of the moral values of good and evil? Or do these second-order affections represent actions and the underlying affections as moral facts?[9] First-order moral responsive sentiments are responses to matters of fact, namely the consequences of actions which either promote the interests of the human species and those of a particular society or stand in the way of such promotion. These first-order sentiments give rise to moral beliefs and desires. They constitute a person's virtue. Second-order moral sentiments inform beliefs about an agent's virtue. These sentiments provide first-order moral sentiments with overriding authority: they make the individual agent feel good about her or his moral actions and bad about his or her non-moral actions. It seems that Shaftesbury holds a response-dependent account of human virtue. The virtuous person responds with certain affections to certain matters of fact; while these responses do not constitute facts about what promotes

people's and societies' interests, they do constitute a person's virtue. And when people morally judge an agent, they do not judge the consequences of her or his actions but the motivational attitude underlying it.

HUTCHESON

Francis Hutcheson (1694–1746) was the second major moral sense philosopher. His views on the moral sense are most clearly articulated in his early work, namely the second part of his *Inquiry into the Original of Our Ideas of Beauty and Virtue* (1725)[10] and his *Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense* (1728)[11]. In the latter work, he embeds his account of the moral sense in a complex account of the mind, developed along the lines of Locke's empiricism. Hutcheson believes that humans were created by God, but otherwise he does not attribute a major role to religion in his inquiry into moral matters. He sets out to defend Shaftesbury's view of the moral sense against Hobbesian objections which had been revived by Bernard Mandeville's *Fable of the Bees* (1714). But in defending Shaftesbury, Hutcheson also proposes modifications of his view. While he endorses Shaftesbury's distinction between first-order and second-order moral affections, the 'moral sense' as he conceives it is a disposition for second-order sentiments exclusively. Due to our moral sense we respond with affections of approval and disapproval to people's actions and the affective motivation on which they are based:

> If we observe the Sentiments of Men concerning Actions, we shall find, that it is always some really amiable and benevolent Appearance which engages their Approbation. (*Inquiry*, 136)

First-order approving affections are, according to Hutcheson, all of one kind; they have their source in benevolence (see *Inquiry*, 105–6). Benevolence is constitutive of virtue; it gives rise to sentiments that inform both moral beliefs and desires. The benevolent person approves what is morally good and acts accordingly. Hutcheson's defence of the existence of genuine benevolence against Hobbesian and Mandevillean psychological egoism implies a clarification – if not a criticism – of Shaftesbury's account of the morally relevant features of actions. The consequences of an action that promote selfish interests have to be distinguished from those consequences to which a virtuous person responds with benevolent sentiments. Hutcheson proposes a utilitarian account of these consequences:

that Action is best, which produces the greatest Happiness for the greatest Numbers; and that, worst, which, in like manner, occasions misery. (*Inquiry*, 125)

The benevolent agent chooses the action that brings about as much happiness as possible for as many people as possible, without making this choice dependent on what impact the action has on his or her own happiness. Any concerns of the latter kind would affect, if not undermine the 'moral goodness' of the choice (*Inquiry*, 85). In making a morally good, purely benevolent choice, the agent reveals her or his virtue and thereby deserves the approval of all people who are provided with a moral sense. One of the reasons for Hutcheson's endorsement of a utilitarian principle of virtuous choice is to avoid reintroducing innate ideas as providing guidance for the virtuous person (see *Inquiry*, 137–8). While Hutcheson is a proto-utilitarian, his theory of moral judgement is not strictly consequentialist. The benevolent agent makes choices guided by the utilitarian principle; but her action is not good merely by virtue of its consequences; what makes it virtuous is the underlying benevolent motivation (see *Inquiry*, 125–6 and *Essay*, 181). Otherwise, the moral sense would be superfluous.

People's benevolence and the moral sense are universal in two ways: they are part of human nature; and they have no limits, extending to all mankind, including 'the most distant part of the Species' (*Inquiry*, 147). The existence of cultural practices that are contrary to virtue and the moral sense cannot refute this claim. After all, virtue is not the only source of sentiments that can inform people's beliefs and desires. Selfish passions can get in the way, as can superstition or errors of computation in utilitarian thinking (see *Inquiry*, 136–45). Hutcheson admits that our benevolence and moral sense allow for degrees of concern; people are more sensitive to the prospects of happiness of those near and dear to them and less sensitive to those of other people. This allows him to observe in 'natural affections' such as parental and romantic love, manifestations of our moral dispositions including the recognition of 'merit', 'gratitude' (*Inquiry*, 147–8), and 'compassion' (*Inquiry*, 159–61). But these natural affections should not be mistaken as sources of our moral dispositions, and the same applies to our love of honour and our aversion to shame (see *Inquiry*, 154–5).

Hutcheson admits that benevolence and the moral sense are not the only sources of sentiments to guide our evaluative beliefs and

desires. What, then, provides them with more authority than competing sentiments that arise from our natural self-love? To answer this question, Hutcheson relies on a distinction between self-love and a reflected interest in happiness:

> Let the Obstacles from Self-love be remov'd, and Nature it self will incline us to Benevolence. (*Inquiry*, 178)

That benevolence rather than self-love will, through reflection, be recognised as the guide to proper happiness is manifest in feelings of uneasiness and dissatisfaction which commonly accompany those actions performed for non-benevolent reasons (see *Inquiry*, 176–7). But while 'reflection and reason' provide further evidence for what we feel by virtue of our moral sense, they do not make the latter superfluous:

> Notwithstanding the mighty Reason we boast of above other Animals, its Processes are too slow, too full of doubt and hesitation, to serve us in every Exigency, either for our own Preservation, without the external Senses, or to direct our Actions for the Good of the Whole, without this moral Sense. (*Inquiry*, 179)

Hutcheson's claims about the moral sense and the diminished role he attributes to reason in moral matters gave rise to objections from adherents both of moral rationalism and of psychological egoism. Thus, Gilbert Burnet questioned the authority of a moral sense, pointing out that, independent as it was from reason, it could easily err.[12] And John Clarke objected that, without the threat of sanctions, people could not be motivated to be moral.[13] In his *Essay*, Hutcheson responds to his critics. His rejection of moral rationalism anticipates the so-called naturalistic fallacy argument commonly attributed to Hume: reason alone cannot constitute ends for a rational agent, and true propositions cannot give rise to desires (see *Essay*, 143).[14] A rational agent's desires and intentions depend on his or her experience of pleasure and happiness. In his response to the latter objection, Hutcheson points out that every person trusts her or his moral sense and that the intersubjective uniformity of moral judgements based on moral sentiments is explicable by the fact that the moral sense is part of human nature and that 'the *Senses* of all Men are pretty *uniform*' – including in particular the moral sense (*Essay*, 176).[15] While social control can support the authority of the moral sense, it does not depend on it. Hutcheson even speaks of the moral sense as of a 'Sense of Obligation' – the obligation to let benevolence guide our

volitions and actions and override self-love (*Inquiry*, 178).[16] While Shaftesbury had associated the moral sense with conscience, this notion does not play an important part in Hutcheson's account of the moral mind.[17]

Shaftesbury's and Hutcheson's conceptions of the nature of human morality and the moral sense are similar insofar as they both distinguish between first-order and second-order moral sentiments. Furthermore, both make room for reflection and reason as informing moral sentiments. But whereas Shaftesbury underlines the reflective nature of the moral sense, Hutcheson attributes a reflective nature mainly to first-order moral sentiments arising from benevolence; second-order sentiments of approval or disapproval are immediate rather than reflected responses. Both Shaftesbury and Hutcheson agree that the moral sense is a source of internal sanctions for non-moral actions; it provides the first-order moral affections and the desires to which they give rise with additional author- ity. However, they do not agree on how to understand the challenge of moral obligation. According to Shaftesbury, self-love and moral motiva- tion are not essentially opposed, even though it requests a reflected moral sense to understand that. Hutcheson distinguishes the utilitarian reason- ing informing moral action from rational egoism as it underlies non-moral action. Both Shaftesbury and Hutcheson see the moral sense as a source of what Stephen Darwall has called an 'internal ought'. According to Darwall, Shaftesbury's particular conception of the moral sense as a reflected sense anticipates the Kantian doctrine of autonomy (see Darwall, *British Moralists*, 21). Kant, however, rejects the idea – promoted by all sentimentalist moral philosophers – that sentiments can provide a desire with the kind of absolute universal authority that morality requires, and this applies equally to spontaneous and reflectively informed sentiments as Shaftesbury described them.[18]

Joseph Butler (1692–1752) is not commonly considered to be a moral sense philosopher. There are straightforward reasons for this: as a bishop, he could hardly agree with the moral sense philosophers insofar as their proposals for naturalising morality diminished the role of reli- gious devotion as its intrinsic source. Butler himself was committed to a natural and divine law theory of morality; nevertheless, he understood his dealings with the subject as committed to the method of 'science'[19] (*Works*, 36), namely to 'matters of fact' rather than 'abstract relations of things' (*Works*, 37). Underlying is Butler's idea that moral laws were revealed to us through our moral nature as created by God. Butler was well acquainted with the debate about the nature and functions of the

moral sense. In his *Second Dissertation* entitled *Of the Nature of Virtue*, first published in 1736, he speaks of the 'moral nature' of people and attributes to them 'moral faculties of perception and action' (*Works*, 309). He mentions that one 'moral faculty' had been called 'conscience, moral reason, moral sense, or divine reason', and that it had been considered 'as a sentiment of the understanding, or as a perception of the heart' (*Works*, 309), but that it also serves as a 'rule of action' (*Works*, 310). He attributes to people a 'sense or discernment of actions, as morally good or evil' (*Works*, 310) and even 'moral understanding and moral sense' (*Works*, 312). Like Shaftesbury and Hutcheson, Butler sees human moral nature as a source of both moral beliefs and the desire to act accordingly. But he rejects both Shaftesbury's claim that the moral sense responds to what is in humans' interest and Hutcheson's claim that human virtue can in all its appearances be reduced to benevolence (see *Works*, 312).[20]

David Hume and Adam Smith take the inquiry into the nature of human morality one step further. Like their predecessors, they are committed to the project of moral sentimentalism and universalism. However, they are sceptical concerning the assumption of a natural moral sense as a part of human nature that allows all people to have moral sentiments, sentiments furthermore which provide the grounds for unanimous and universal moral beliefs and desires. This scepticism is based on three kinds of observations: (1) while people provided with a system of visual perception can see without learning to see, they have to learn to be moral – and will fail to do so if deprived of the opportunity; (2) shared standards for the morally good and bad are subject to change over time; and (3) different cultures can develop different standards for the morally good and bad. Their idea is that the universal morality to which people feel committed is an ideal; while they should try to reach it, they are unlikely to actualise it in full. The empirically minded moral philosopher is well advised to make room for a pluralism of moral cultures. Furthermore, he should provide an account of the processes of moral development and progress, both in individuals and in groups, and inquire into the question of what such an account can contribute to the justification of people's non-ideal moral beliefs and desires.[21] Hume and Smith, however, disagree on how to understand the natural psychological foundations of human morality, on the dynamics of individual and collective moral learning and on the sources of the normative authority which people's moral beliefs and desires can rightly request.

HUME

David Hume (1711–76) developed his views on the nature of human morality first in Books II and III of his *Treatise of Human Nature* (1739/40)[22] and then – in a more concise form – in his *Enquiry Concerning the Principles of Morals* (1751).[23] Methodologically speaking, he is committed to a Newton-inspired mechanistic account of the mind that takes all our beliefs and desires to be traceable to their original 'impressions' and 'ideas'. But this does not exclude, at times, his reliance on a more common-sensical method, namely on 'proof drawn from daily experience and observation' (*Treatise*, 225). As for our moral beliefs, evaluative in kind as they are, and the corresponding desires, they arise from external and internal sentiments such as bodily pleasures and pains. These sentiments are triggered by external objects and are first order in kind. They give rise to two different kinds of passions, states of the mind that inform evaluative beliefs and desires (see *Treatise*, 181). Passions are second-order sentiments, responses to first-order sentiments. The direct passions inform our evaluations of things as either good or bad (*Treatise*, 281); the indirect passions are either self-centred (for example pride and humility) or other-centred (for example love and hatred). Like Hutcheson, Hume thinks that an action is good or bad by virtue of the underlying affective motive, not by virtue of its consequences (*Treatise*, 307). Virtuous people perform virtuous actions and approve of them. But Hume rejects Hutcheson's account of virtue in terms of benevolence (*Treatise*, 310, 316).

Underlying our pleasures and pains and the moral passions triggered by them are two affective dispositions: selfishness (*Treatise*: 313, 333) and 'sympathy' (*Treatise*: 206). Sympathy is empathic in kind. It allows us to share the feelings of other people, in particular of those near and dear to us. This empathic sympathy also makes us care about what other people think about us and how they evaluate our behaviour. Sympathetic sentiments inform, along with indirect passions, our moral judgements; they allow us to track the objects of these judgements. Moral judgements are judgements about agents' characters as virtuous or vicious, and these traits are manifest in the actions the agents most commonly perform and the sentimental responses these actions commonly induce among the members of the agent's narrow circle. 'Vice and virtue … are not qualities of objects but perceptions in the mind'. Therefore, 'they may be compared to sounds, colours, heat or cold' (*Treatise*, 301).

When people engage in empathic sympathy, they silently assume that all people are similar to each other (*Treatise*, 208, 232, 238). However, similarity comes in degrees; we are more sensitive to sentiments of people whom we recognise as very similar to us and less sensitive to sentiments of other people, for example strangers with whom we are not acquainted or who belong to other cultures (*Treatise*, 228–9). While every human being can empathically sympathise with every other human being, those who live closely together and interact on a regular basis become increasingly similar to each other, and this facilitates their mutual empathic sympathy and approval (*Treatise*, 206). Put in more modern terms, Hume's claim is that specific groups of people, due to their constant interaction and the sentiments to which it gives rise, create shared standards of pro-social and anti-social behaviour and encourage each other to behave pro-socially. Actions that commonly induce positive sentiments are in conformity with silently shared standards of pro-social behaviour. The particular standards of behaviour may vary between different groups since people can be different and they may have to deal with different kinds of external challenges; Hume speaks of phenomena 'diversify'd by a variety of circumstances' (*Treatise*, 375; see also *Moral Enquiry*, 198).

Hume distinguishes between natural and artificial virtues (see *Treatise*, 305). It is natural that people, due to their natural selfishness and empathic sympathy, develop a disposition to approve of pro-social behaviour and to disapprove of anti-social behaviour in themselves and others. But when they institutionalise the respective social standards and legislate the principles of justice, variations between groups and cultures become apparent. According to Hume, 'justice' is an artificial virtue (*Treatise*, 311), in contrast to the 'natural' virtues as they shape inner-cultural motivational attitudes. The principles of justice reach beyond the confines of particular cultures. Hume mentions two principles of justice in particular: the rules that protect property rights and the rules that regulate promises and contracts (see *Treatise*, 335). These rules function as the backbone of every society, whatever its particular culture is. The institutions of property, promises and contracts depend on universal authority; they cannot be sufficiently protected if their authority is limited to a particular culture. All people have a duty to respect the rules of justice. Even though justice on the one hand and selfishness on the other can sometimes give rise to incompatible desires (see *Treatise*, 312–13, 333–4), they do

not inevitably do so: 'Our sense of duty always follows the common and natural course of our passions' (*Treatise*, 311).

Hume rejects rational egoism as an adequate account of people's motivational dispositions; this rejection is the central concern in his *Enquiry*. The 'common' disposition to engage in just behaviour is the result of an evolutionary process in the course of which people – infirm and needy by nature as they are – have learned to cooperate in order to satisfy their natural needs under conditions of harshness and scarcity (*Treatise*, 312; see also 317). Learning to be pro-social and to respect the rules of justice is both an evolutionary process of collective learning and a process of individual learning that every child has to live through:

> custom and habit operating on the tender minds of the children, makes them sensible of the advantages, which they may reap from society, as well as fashions them by degrees for it, by rubbing off those rough corners and untoward affections, which prevent their coalition.
> (*Treatise*, 312)

Hume is famous for denying that reason can underpin moral distinctions between vice and virtue. Indeed, as a moral sentimentalist, he claims that sentiments provide the basis for our moral judgements, sentiments as they arise in a complex interaction of selfishness and sympathy: 'Morality . . . is more properly felt than judg'd of' (*Treatise*, 302). These sentiments are themselves reflective, informed by thinking either about causal relations and how they contribute to personal utility (see *Treatise*, 295–6), or about agents' character traits.

Hume calls the faculty of the mind by which we perceive such character traits as either virtuous or vicious a 'moral sense' (*Treatise*, 302; see also 375) or a 'moral taste' (*Treatise*, 371). But he does not thereby attribute an additional sense to humans. Our disposition to love and approve of virtuous agents and to hate and disapprove of vicious agents is explicable in terms of our natural perceptual capacities, of our passions, and of the processes of collective and individual learning in which we inevitably engage. The assumption of a moral sense adds nothing to the previously given account of the sentiments and passions which inform our moral beliefs and desires (see *Treatise*, 303–4).

While the moral sense guides our interaction with those near and dear to us, our disposition to respect the rules of justice is manifest in 'a general sense of common interest . . . which . . . all the members of the society express to one another, and which induces them to regulate their conduct by certain rules' (*Treatise*, 315). Guidance provided by this sense

is not restricted to our interaction with those near and dear to us. Nevertheless, morality and justice are not substantially different; the former provides the foundations for the latter, and 'the first rudiments of justice must every day be improv'd, as the society enlarges' (*Treatise*, 316). Accordingly, we rely on the same standards when judging the virtue or vice either of strangers or of people with whom we are closely acquainted. Hume even claims that 'the sense of moral good and evil follows upon justice and injustice' (*Treatise*, 320) and he speaks of this sense also in terms of '*sympathy* with *public* interest' (*Treatise*, 321, Hume's italics). This acquired sense of justice is, however, not to be mistaken for the natural empathic sympathy which people have in addition to their natural selfishness and which provides the foundations for the development of pro-social attitudes within small social circles in the first place.

How, according to Hume, can the universal authority of justice develop on the basis of local moralities? In his answer to this question he explores the assumption that human beings are naturally similar to each other. They are in particular similar in their natural vulnerability and dependency on co-operation for surviving and flourishing in a world of scarcity. Since they are provided with the same natural dispositions, they will develop similar social conventions whatever the specific circumstances are in which they live. The 'circles of familiarity' as they are constituted within specific cultures do not override the natural disposition to sympathise empathically with every other human being. As much as close acquaintance facilitates empathic sympathy, it also gives rise to partialities. In order to prevent such partialities, judgements about the virtue or vice of an agent should be made from a 'steady and general point of view' (*Treatise*, 371–2), the point of view of a 'spectator' (*Treatise*, 374). Here, Hume endorses a point previously made by Hutcheson: that personal attachments and interests can stand in the way of making a proper moral judgement, a judgement that can rightly request universal authority (see *Treatise*, 377).

In the *Enquiry*, Hume admits that his functional account of the universal authority of justice is vulnerable to a free rider objection. While no society can exist without its members engaging in pro-social and virtuous behaviour and respecting the rules of justice, an individual agent may be tempted to violate these principles if he can thereby promote his own utility without at the same time risking the existence of the society as a whole (*Moral Enquiry*, 282–3).

Hume's evolutionary account of the origin and continuous development of morality and justice is more psychologically realistic than the moral sense based accounts provided by Shaftesbury and Hutcheson. His explanation of the universal authority of the principles of justice is both similar to and weaker than previous explanations which relied on a genuine moral sense. It is similar in that it is mainly based on the assumption that all people are similar enough in their natural vulnerabilities that they will rely on roughly similar strategies for surviving and satisfying their various needs; and it is weaker since Hume attributes to the principles of justice a merely functional and conventional kind of authority, an authority that depends on a sufficient number of people actually submitting to these principles.

KAMES

In his *Essays on the Principles of Morality and Natural Religion* (1751/1758/1779),[24] Henry Home Lord Kames (1696–1782) states that he is dedicated to the scientific method in moral theory, to 'facts' and 'experiments' (*Essays*, 24) as well as to the defence of a universal account of morality and justice, just as his predecessors were. He rejects Hobbes' and Mandeville's rational egoism as much as the particulars of Shaftesbury's and Hutcheson's accounts of a moral sense. Still, he too claims the existence of a moral sense as the origin of the sentiments which inform both our moral beliefs and desires. However, he argues that these sentiments are not responding either to the consequences of actions or to the sentimental attitudes underlying actions; instead, they are responding to principles. The moral sense, according to Kames, is 'an instructor to regulate our actions' and it makes us approve 'rules of conduct founded on natural principles' (*Essays*, 56). Kames rejects two aspects of Hume's account of justice: its merely functional and conventional authority seems to him inadequate; and the restriction of its content to the regulation of property, promises and contracts seems to him too narrow. Instead, he suggests going back to an older understanding of the principles of justice as divine laws, and he suggests they should include principles that forbid 'the injuring others in their persons, in their fame, or in their goods' (*Essays*, 32). Violating these principles would not only be strategically unwise but '*absolutely* wrong' (*Essays*, 32, Kame's italics).

SMITH

Adam Smith (1723–90), in his *Theory of Moral Sentiments* (1759[1sr]/ 1790[6th]),[25] develops a new version of moral sentimentalism. It is a complex response to Hume's moral theory, partly approving and partly critical. Underlying are three claims that had been developed in the earlier debate about the nature of human morality: (1) that conscience plays an important part in moral self-assessment, (2) that moral judgements should be made impartially and (3) that at least the principle of justice that forbids harming other people should have more than merely functional and conventional authority. Smith follows Hume mainly in his developmental account of humans' moral nature and in his assumption that humans are naturally provided with two emotional dispositions, namely self-love and sympathy. Like Hume, he rejects Shaftesbury's and Hutcheson's assumption of a moral sense, pointing out that 'the views of mankind with regard to the propriety of their own conduct' are too 'partial' for assumption of a moral sense to be plausible (*TMS* III.4.5, 158).[26] However, Smith's accounts of the nature of human sympathy and of the driving forces and aims of moral development differ substantively from those suggested by Hume. This is due to his underlying conviction that morality and justice require more than merely conventional authority.

Smith agrees with Hume and many others on the function of self-love; it is the origin of sentiments that inform evaluative judgements and desires which motivate and guide people to take care of their own survival and happiness. But sympathy as Smith understands it differs from both kinds of Humean sympathy. Hume understood sympathy (both the empathic sympathy and the sympathy underlying our disposition to submit to the rules of justice) as contagious in kind. Smith's sympathy is of a projective nature and it plays an important role in the shaping of the standards of moral judgements, judgements both about oneself and about other people.[27] Smith sees sympathy as the emotional disposition that allows us to become aware of the fact that different people experience the world, and in particular human actions and their consequences, in different ways, due not only to their different points of view (their being affected in different ways), but also to their different vulnerabilities. Since people make proto-moral judgements about agents on the basis of their responsive sentiments to their actions, these judgements are not unanimous. Proto-moral judgements are partial. But moral judgements are not judgements about personal preferences people

happen to have. Moral judgements make claims to universal authority.
The question is therefore how proto-moral judgements, informed by
essentially partial responsive sentiments, can acquire this universal
authority.

To answer this question, Smith introduces the 'impartial specta-
tor' (*TMS* I.i.5.4, 24). This spectator is not simply someone who takes
a 'steady and general point of view' as Hume had suggested. The chal-
lenge is not to make sure that partialities of people concerned do not
affect a moral judgement; the challenge is rather to make an impartial
and universally authoritative moral judgement about an agent and his
action that takes properly into account the ways it affects various people.
To meet this challenge, the person affected and the spectator have to
engage in a process of sympathetic interaction and communication. In
the course of this process they switch points of view; they engage in
a process of emotional learning, sharing relevant knowledge of the cir-
cumstances affecting the person, and of her particular vulnerabilities.
Ideally, the person affected and her spectator will then agree on what
kind of sentimental response would have been appropriate to the parti-
cular action under these particular circumstances for this particular
person. In the course of this process the person affected internalises the
role of her spectator, learning to 'look' at herself as an external spectator
would. This 'looking' is a matter of both first-order sympathy
and second-order sympathy or antipathy; that is, a matter of a person's
awareness and approval or disapproval and revision of the sentiments
she felt when she was affected by an action. This is what Adam Smith
calls 'conscience' (*TMS* III.3.1–10, 134–40). Conscience is the faculty of
impartial moral self-assessment.

What a particular person affected by another's action and her
spectator agree to be the morally proper sentimental response to
a particular action – and thus the proper moral judgement about the
agent – will be less partial than the original response of the person
concerned had been. But it cannot claim impartiality and universal
authority before all relevant people have been involved in sympathetic
interaction and communication, before all of them have reached an
agreement on what the proper response for the person concerned would
have been. Underlying this is the idea that the transition from partiality
to impartiality can only be achieved through a process of sympathetic
interaction and communication between as many people as possible,
including in particular an agent, the people affected by her action, and
many other people who have either witnessed the action and its

consequences or acquired knowledge of it in some more indirect way. Impartiality comes in degrees.

To guarantee that this process does indeed bring about a consensus, participants have to submit to what Smith calls 'the most sacred laws of justice', namely 'the laws which guard the life and person of our neighbour; ... those which guard his property and possessions; and ... those which guard what are called his personal rights, or what is due to him from the promises of others' (*TMS* II.ii.2.2, 84). While Smith does not deny the Humean claim that respect for these laws is necessary for the existence of society, he attributes another function to them. For the sympathetic process to succeed, for it to produce agreements on moral matters that can rightly claim to be impartial and universally authoritative, it needs to be inclusive. Everybody is equally welcome to participate, and all participants must respect each other as equal participants in the process. And if they do so, they will inevitably submit to the 'most sacred rules of justice'.

The sentiments that inform moral judgements about agents are sentiments of people affected by the consequences of the agents' actions. But before such sentiments can provide the grounds for an impartial or universally authoritative moral judgement, they have to be approved by impartial spectators. Impartial responsive sentiments will thus be reflected sentiments, felt by people who are well-informed about external factual matters as well as about agents and, last but not least, about their own vulnerabilities. People are under an obligation to exercise self-command over their responsive sentiments and to make sure that they are as impartial as possible, before they form moral beliefs and desires based on these sentiments. This obligation arises from the claims to universal authority they make for their moral judgements, claims that can be more or less justified.

The philosophers of the moral sense as well as the moral sentimentalists who rejected the assumption of a moral sense equally tried to provide an account of moral belief and desire and the universality of moral obligation in the framework of an empirical science of man. What they have in common is the conviction that moral beliefs and the corresponding desires have to be grounded in sentiments and that moral relativism and subjectivism needs to be avoided nevertheless. Darwall has suggested conceptualising this concern as their search for an 'internal ought'. Indeed, in a naturalistic and sentimentalist framework, the source of this obligation can only be internal; it has to be traced to the sentiments of the agent. Several philosophers, first and foremost

Shaftesbury and Hutcheson, have suggested solving this problem by attributing a moral sense to people. Hume and Smith rejected this strategy as unconvincing in psychological terms; but they remained committed to the underlying naturalistic and sentimental project of moral philosophy. According to Darwall, the interest in the 'internal ought' that these philosophers share makes them forerunners of Kant's moral theory and the Kantian doctrine of moral autonomy in particular. Kant, however, rejected all their suggestions as 'tenuous' (Kant *AA* 6:215); according to him, a sentimentalist account of the morally good and its authority cannot provide ultimate justification for our moral beliefs and the corresponding duties. Nevertheless, the moral sense philosophers and the moral sentimentalists remain the first to try and provide an empirically informed, psychologically realistic account of human morality without giving up its claims to universal authority. While Hume has been the most influential, Smith is the one who anticipated most of the later psychological insights into the way empathy and sympathy contribute to the development of a moral character.

NOTES

1. On the philosophical debate about the moral sense, see D. Daiches Raphael, *The Moral Sense* (Oxford: Oxford University Press, 1947). A comprehensive study of the moral theories of Hutcheson, Hume and Smith with a focus on their respective accounts of virtue can be found in Vincent Hope, *Virtue by Consensus: The Moral Philosophy of Hutcheson, Hume and Adam Smith* (Oxford: Clarendon Press, 1989). Gill has provided a detailed analysis of the moral views of Shaftesbury, Hutcheson and Hume and of their predecessors Whichcote and Cudworth. He reads the moral sense debate as the birth of secular ethics; see Michael B. Gill, *The British Moralists on Human Nature and the Birth of Secular Ethics* (Cambridge: Cambridge University Press, 2006). Darwall reads the debate as an anticipation of the Kantian doctrine of autonomy; see Stephen Darwall, *The British Moralists and the Internal 'Ought': 1640–1740* (Cambridge: Cambridge University Press, 1995). For a comprehensive account of the debate and its historical context, see J. B. Schneewind, *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge: Cambridge University Press, 1998).
2. See 'Brief Discourse on Enthusiasm' 22 and 'In Defence of the Moral Cabbala' 232, in More More, Henry (MDCCXII), *A Collection of Several Philosophical Writings of Dr. Henry More*, 4th edn. (London: Joseph Downing).
3. See Burnet, Thomas (1697/1699), *Remarks on John Locke. With Locke's Replies*, ed. with an introduction and notes by G. Watson (Doncaster: Brynmill, 1989), 66–7.

Burnet's reading of Locke's views on morality is inadequate since Locke was himself an adherent of the natural law theory of morality and defended moral universalism. But Locke's hedonistic account of moral obligation induces such a misunderstanding. See on Locke's moral theory and its controversial interpretation Patricia Sheridan, 'Locke's Moral Philosophy', in The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta, https://plato.stanford.edu/archives/sum2016/entries/locke-moral/. And on Burnet's reading of Locke, Ernest Tuveson, *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge: Cambridge University Press, 1998).

4. See Tuveson, *Invention of Autonomy*, 255, 259.

5. In 1711, Shaftesbury published an anthology of his previous writings under the title *Characteristics of Men, Manners, Opinions, Times*; both the *Inquiry* and the *Sensus Communis* were part of this anthology. Anthony Ashley Cooper, Third Earl of Shaftesbury, *Characteristics of Men, Manners, Opinions, Times* (Cambridge: Cambridge University Press, 1999).

6. See Shaftesbury, *Sensus Communis*, Part III.

7. On Shaftesbury's account of self-love and moral obligation, see Darwall, *British Moralists*, 193–6 and 203–4.

8. See *Characteristics*, 171–7.

9. See Michael B. Gill, 'Lord Shaftesbury [Anthony Ashley Cooper, 3rd Earl of Shaftesbury]', in The Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta, https://plato.stanford.edu/archives/fall2016/entries/shaftesbury/.

10. Francis Hutcheson, *An Inquiry into the Original of Our Ideas of Beauty and Virtue*, ed. Wolfgang Leidhold (Indianapolis, IN: Liberty Fund, 2004) (quoted as '*Enquiry*').

11. Francis Hutcheson, *An Essay on the Nature and Conduct of the Passions and Affections: With Illustrations on the Moral Sense*, ed. and with an introduction by Aaron Garrett (Indianapolis, IN: Liberty Fund, 2002) (quoted as '*Essay*').

12. See Aaron Garrett, 'Introduction', in Hutcheson, *An Essay on the Nature and Conduct of the Passions*, xiii.

13. Ibid., xiv.

14. See Darwall, *British Moralists*, 224–6.

15. Another critical response to Hutcheson's theory of the moral sense came from Richard Price. Price raised several important issues concerning the metaphysical and epistemological implications of Hutcheson's theory. On Price's response to Hutcheson, see Raphael, *Moral Sense*, 99–145.

16. On Hutcheson's understanding of 'moral obligation', see Darwall, *British Moralists*, 218–43.

17. Hutcheson rarely uses the notion of conscience. See *Inquiry*, 189 and *Essay*, 65.

150 CHRISTEL FRICKE

18. See Immanuel Kant, 'The Metaphysics of Morals', in Immanuel Kant, *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1797/1996), 353–603 (quoted as *AA* 6: 215–16).

19. Joseph Butler, *The Works of Bishop Butler*, ed. and with an introduction and notes by David E. While (Rochester, NY: University of Rochester Press, 2006).

20. On the relation of Butler's moral thought to that of Shaftesbury and Hutcheson, see Darwall, *British Moralists*, 244–5.

21. On the dynamic and progressive account of human nature as Hume held it, see also Gill, *British Moralists on Human Nature*, 237–40.

22. David Hume, *A Treatise of Human Nature*, eds. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2006) (quoted as '*Treatise*').

23. David Hume, 'An Enquiry Concerning the Principles of Morals', in David Hume: Enquiries Concerning Human Understanding and Concerning the Principles of Morals, reprinted from the 1777 edition with introduction and analytical index by L. A. Selby-Bigge, 3rd edn. with text revised and notes by P. H. Nidditch (Oxford: Clarendon Press, 1975) (quoted as '*Moral Enquiry*').

24. Henry Home, Lord Kames, *Essays on the Principles of Morality and Natural Religion*, ed. and with an introduction by Mary Catherine Moran (Indianapolis, IN: Liberty Fund, 2005) (quoted as '*Essays*').

25. Adam Smith, *The Theory of Moral Sentiments*, eds. D. D. Raphael and A. L. Macfie (Indianapolis, IN: Liberty Fund, 1759/1790/1984) (quoted as '*TMS*').

26. See also *TMS* VII.iii.3.15, 326.

27. On the distinction between Hume's and Smith's understanding of sympathy, see Samuel Fleischacker, 'Sympathy in Hume and Smith: A Contrast, Critique, and Reconstruction', in *Intersubjectivity and Objectivity in Adam Smith and Edmund Husserl*, eds. Christel Fricke and Dagfinn Føllesdal (Frankfurt: Ontos, 2012), 273–311.