Conscience, guilt and shame[1]

1. Introduction

A great deal of morality is concerned with the public arena: examining what is required for a just and fair society, and trying to map out the nature and extent of our obligations to our fellow citizens and to the wider world. The concepts of conscience, guilt and shame, by contrast, seem primarily concerned with what might be called the interior dimension of morality: how each of us thinks and feels about our own conduct when we review it, or when we measure it against our sense of what is expected of us, or how we might have done better.

The idea of conscience carries, etymologically, something of this interior flavour. The Latin word *conscientia* – from *con* or *cum* ('with') and *scientia* ('knowledge') – originally meant knowledge shared with another (e.g. a fellow-conspirator); but it came to be used by extension of the private knowledge an individual shares (as it were) with him or herself, and hence the term is employed quite generally to refer to a person's inner mental awareness. Thus we find Descartes in the seventeenth century reported as saying that if we look within ourselves, we will see that we have 'intimate inner awareness' (Lat. *intime conscii sumus*) of the freedom of the will (Descartes (1648)). Elsewhere he defines a thought as 'that which is within us in such a way that we are immediately *aware* [Lat. *conscius*] of it' (Descartes (1641)). The thoughts of which we are 'conscious' in this sense need not have anything specifically to do with morality, though of course they may. This may partly explain how the term '*conscience*' in modern French does duty for what in English is conveyed by two separate words, 'consciousness' and 'conscience': it can mean either the direct inner awareness each of us supposedly has of our own mental states in general, or, in a more specifically moral sense, one's inner awareness of, for example, guilt at wrongdoing.

The connection between conscience and guilt is enshrined in common phrases like 'he must have a guilty conscience', used when someone is behaving in an embarrassed or furtive manner, or in a way that indicates awareness of wrongdoing. But conscience need not be guilty conscience. A commonplace book of John Marbeck (1581) defines conscience as 'the knowledge, judgement and reason of a man, whereby every man, in himself, and in his own mind, being privy to everything he either hath committed or not committed, does either condemn or acquit himself' (*Oxford English Dictionary*, s. v. 'conscience'). So conscience can be, as we say, a 'clear conscience', when someone's inner reflection leaves him in the happy position of finding nothing wrong with how he has behaved. Consistently with this, John Locke observes in the *Essay concerning Human Understanding* that 'conscience is … nothing else but our own judgement of the moral rectitude or pravity of our own actions' (Locke (1670), Bk I, Ch. 3, §8).

Because of its central importance in the moral life, one might suppose that conscience would be a primary concern of philosophers working in ethical theory. Certainly it bulked large in the writings of many moral philosophers in the seventeenth and eighteenth centuries; but it is surprising to find how comparatively little attention is paid to it in modern anglophone philosophy (see however Donagan (1977), Wallace (1978)). One possible reason for the relative neglect of the topic may lie in the increasing secularism of our philosophical culture, and the resulting mistrust of anything that smacks of a religious framework for understanding the human moral predicament. Notions like 'the examination of conscience' (found for example in the *Spiritual Exercises* of Ignatius of Loyola, *c.*1525) are associated in many people's minds, whether explicitly or implicitly, with the religious notion of sin, which many modern philosophers would probably say has no place in a rational theory of ethics. Yet while there have certainly been theological writers on ethics who have seemed preoccupied with the burden of human sinfulness, most of us, irrespective of religious allegiance or its lack, will recognize feelings of guilt and/or shame as having from time to time played a key role in our lives. At the very least they seem to have an indispensable role in the

---

[1] This is a draft of a paper the definitive version of which appeared in in R. Crisp (ed.), *The Oxford Handbook of the History of Ethics* (Oxford: Oxford University Press, 2013), pp. 729-43.

phenomenology and psychology of ethics; and arguably they are an integral part of our sense of ourselves as responsible moral agents.

The discussion that follows will first provide a short survey of the concept of conscience as it has evolved in the Western philosophical tradition. Certain questions arising from this about the supposed authority or normativity of conscience will then be examined. The relation between the idea of conscience and the notions of guilt and shame will be dealt with next, which will in turn lead on to the question (broached in the seminal work of Bernard Williams in this area) of whether the concepts of guilt and shame inhabit essentially different ethical landscapes. The chapter will conclude by looking at the contribution made by psychoanalytic thinking to our modern understanding of the phenomena of conscience, guilt and shame, and by asking why it is that the resulting insights have been so imperfectly assimilated into contemporary anglophone moral philosophy.

2. Conscience in the Western tradition

One of the earliest Western texts to prefigure the idea of conscience is Psalm 51 in the Hebrew Bible (or 50 in the Vulgate and Septuagint numbering), now generally known as the *Miserere*, and dating from several hundred years BC. Reflecting on his past conduct, the author utters a prolonged plea for mercy, linked to a frank acknowledgement of transgression: 'I acknowledge my guilt, and my sin is ever before me'. Traditionally, the composition of the psalm is attributed to King David, after he had been brought by the prophet Nathan to a keen sense of his wrongdoing in seducing Bathsheba and sending her husband to the front line to be killed. According to the story, what had made David feel the pangs of conscience was Nathan's parable about a rich man who was unwilling to draw on his own flocks and herds to feed a visitor, but instead took from a poor man his one ewe lamb, which he dearly loved. David's outrage at such behaviour was then turned in on himself, when Nathan told him 'thou are the man!' (2 Samuel 12: 1-23). The seemingly simple text presupposes an ethical framework of considerable subtlety. Three key points in particular emerge from the story about how the operation of conscience is envisaged. First, it involves a directing inwards by the subject of the kind of disapproval characteristically felt at the untoward behaviour of another. Second, it is linked to remorse and repentance, which is in turn made possible by a deepening both of self-awareness and of empathy: David's previously shallow grasp of the significance of his actions was altered under the imaginative stimulus of being presented with a vivid analogue of his own conduct, and thus starting to appreciate how being treated in such a way would feel for the victim. And third, the required response is not simply implanted from the outside by the prophet's condemnation, but is partly elicited from within. Once certain emotional and cognitive barriers are lifted, it is David's own conscience that convicts him.

The complex psycho-ethical framework presupposed here may be found in one form or another in many prophetic writings from the Old Testament, and it continues on through the parables of Jesus of Nazareth (such as that of the Prodigal Son (Luke 15: 11-32)), and into much subsequent Christian moral philosophy. If, by contrast, we turn to the other great source of Western ethics, namely the philosophy of the classical Greek world, we find a rather different picture. Although the workings of guilt and remorse are vividly present in much Greek tragedy (compare Sophocles' *Oedipus Tyrannus*), in the ethical writings of Aristotle there is no developed idea of conscience. Aristotle does discuss *aidos* ('modesty' or 'shame'), but he characterises it as a 'fear of disrepute' – something that may be useful in restraining the young, but which should be no part of the makeup of a person of confirmed virtue (Aristotle (325 BC), Book IV, Ch. 9). Aristotle's central ethical preoccupations are not with prophetic calls to righteousness or with self-scrutiny and repentance, but rather with mapping out the virtuous life of a successful and flourishing human being. Harmonious moral development, on the Aristotelian view, involves a happy match between inclination and right conduct, which arises from an individual's having been inducted as a child into a sound ethical culture. Aristotle does not describe this process in detail, but the main idea appears to be that even at a pre-rational level, the young child will be progressively encouraged to take pleasure in behaving virtuously. At this early stage, he or she need have no clear rational grasp

of exactly what makes the relevant actions good or bad; she is simply being habituated to act and feel appropriately. But by the time the child comes of age, habits of right behaviour will be supported by an understanding of what it is about the conduct in question that makes it praiseworthy or blameworthy. The outcome of this process is a person of mature ethical virtue: someone who has the right habits of feeling and action, but who also has the capacity to discern what should be done and why – someone who acts, as Aristotle puts it in the *Nicomachean Ethics*, 'knowing what he is doing and choosing it for its own sake', and who has the right feelings 'at the right times on the right grounds towards the right people, for the right motive and in the right way' (Aristotle (325 BC), Bk II, Ch. 4).

The Aristotelian conception of ethical virtue seems to lack that sense of humans as essentially conflicted beings which emerges so vividly in the Judaeo-Christian worldview – a worldview in which a sense of sin, and its corollary conscience, has a pivotal place. To be sure, Aristotle did give considerable attention to the phenomenon of *akrasia* – the failure through weakness to chose the best option – and this has some affinities to the disordered and conflicted state of which St Paul famously spoke when he said 'Woe is me! The good that I would, that I do not, and the evil that I would not, that I do' (Romans 7: 19). But Aristotle's model is not that of the transgressor tormented by conscience for his or her lapses of will, but rather of someone who desires the good but (under the influence of passion) is subject to cognitive error, mistaking the lesser good for the greater (Bk VII, Ch. 3).

In the letters of Paul and several other places in the New Testament, as well as in some non-Christian writers of roughly the same period (for example the Stoic philosopher Epictetus, though the interpretation of the relevant passage is disputed), we find the first occurrences of an explicit term for of the concept of conscience, namely the Greek word *syneidesis*. This is normally translated into Latin as *conscientia*, and appropriately so, since it is made up of exactly the same corresponding elements, coming from *syn* ('with') and *eidesis* ('knowledge'). In a frequently cited passage Paul remarks that the Gentiles, though not possessing the Law, 'do by nature the things contained in the law, and show the work of the law written in their hearts, their conscience (*syneidesis*) also bearing witness' (Romans 2: 15).

A corruption or copyist's misspelling of the New Testament Greek word appears to have led to the curious term 'synderesis', or 'synteresis', found in later patristic writers (starting with St Jerome in the early fifth century). Jerome speaks of synteresis as the 'spark of conscience [*scintilla conscientiae*] not extinguished in Cain when he was driven from paradise' (Jerome (414), at Ch.1, v. 7). In Thomas Aquinas, synderesis appears as the name for the innate, God-given cognitive disposition or 'rational power', which enables us to know right and wrong: 'Just as there is a natural habit of the human soul through which it knows principles of the speculative sciences … so too there is in the soul a natural habit of first principles of action, which are the universal principles of the natural law (Aquinas (1256-9), Qu. 16). The reference to 'habit' may seem reminiscent of Aristotle (whose account of virtue strongly influenced Aquinas elsewhere), but the emphasis here is not on childhood training but on rational knowledge of principles of action, implanted in the soul by God. Aquinas goes on to distinguish this general rational power of 'synderesis' from the more specific exercise of *conscientia*. The latter involves applying principles of practical knowledge to particular cases, so as to evaluate what one should do now, and what one has done in the past: 'in so far as knowledge is applied to an act as directive of it, conscience is said to prod or urge or bind. But, in so far as knowledge is applied to an act by way of examining what has already taken place, conscience is said to accuse or cause remorse, when what has been done is found to be out of harmony with the knowledge according to which it is examined' (Qu. 17, art. 1).

Although the main elements of this account are cognitive, conscience being seen primarily in terms of knowledge or its application, this last quotation shows that Aquinas also acknowledges the emotional components of conscience, such as feelings of remorse. This corresponds to our modern conception of conscience, which combines cognitive and affective elements (if I have a conscience about *X*, I *know* that I have done wrong, and I *feel* badly about it). The latter element becomes particularly important in the discussions of conscience found in many of the early-modern

philosophers, especially those whose 'empirical naturalist' programme (to use Stephen Darwall's label) led them to try to explain our moral capacities and judgments in terms of the natural sentiments, drives and impulses we find within us (see Darwall (1995), Ch. 1).

In the second of his *Fifteen Sermons* (1726), Joseph Butler, one of the most influential writers on conscience, at first appears to be taking an empirical psychologizing approach, talking of various 'natural principles' in man, including that whereby 'man approves or disapproves his heart, temper and actions'. But Butler then makes a crucial distinction between principles that are natural merely in the sense of being prevalent, or commonly occurring, and those which are natural in the sense that they carry an authoritative or (as philosophers now say) a 'normative' force. (For this latter sense of 'natural', compare the traditional philosophical use of phrases like 'natural law' or 'natural light'.) We may have natural dispositions to kindness and compassion, but 'since other passions [such as anger] … which lead us … astray, are themselves in a degree equally natural, and often most prevalent … it is plain the former considered merely as natural … can no more be a law to us than the latter.' But alongside such naturally occurring impulses, there is 'a superior principle of reflection or conscience in every man . . . which pronounces some actions to be in themselves just, right, good; others to be in themselves evil, wrong, unjust.' The deliverances of conscience, then, are not to be regarded as simply one group among the many competing internal principles which may motivate us, but have a special authoritative status, which enables them, in Butler's phrase, to 'be a law to us'. Butler concludes that 'it is by this faculty [of conscience], *natural* to man, that he is a moral agent … a faculty in kind and in nature supreme over all others, and which *bears its own authority* of being so' (Butler (1726), Sermon II, §8, emphasis supplied). This idea of the authoritative and quasi-legal status of conscience prefigures the conception of Immanuel Kant, later in the eighteenth century, according to which conscience [*Gewissen*] is the 'consciousness of an *internal court* in man, 'before which his thoughts accuse or excuse one another'. Every human being, Kant goes on to argue in the *Tugendlehre* or *Doctrine of Virtue* , is 'kept in awe an by internal judge', and 'this authority watching over the law in him is not something that he himself voluntarily makes, but something incorporated in his being' (Kant (1797), §13).

The strongly normative conception of conscience which we find, in different ways, in Butler and in Kant corresponds to an idea of conscience that is still widespread; but it raises the philosophical question of what might be the source of such an authoritative faculty, and of whether its supposed existence is compatible with the naturalistic framework for understanding human nature to which many present-day philosophers are drawn. To this cluster of questions we will now turn.

3. The authority of conscience

Kant's solution to the problem of the authority of conscience is too complex to unfold in detail here, but in essence it depends on his notion, found in the *Groundwork for the Metaphysic of Morals*, of the rational will as 'self-legislating' or 'giving the law to itself' (*selbstgesetzgebend*) (Kant (1785), Ch. 2). In the *Doctrine of Virtue*, Kant in effect links this conception to the idea of conscience. Within the 'internal forum' of each of us, my acts are brought before the tribunal of reason. But, Kant goes on to argue that it is absurd to think of a human being who is accused as one and the same person as the judge; and hence the subject must think of himself as being judged by another, who is 'an ideal person that reason creates for itself' (Kant (1797), §13). This looks like an internal analogue of the religious idea of a supreme authoritative judge; but on the question of whether it has any real external counterpart, Kant explicitly draws back from saying that a human being is 'entitled, through the idea to which his conscience unavoidably guides him, to assume that such a supreme being actually exists.' The idea of a supreme 'scrutinizer of all hearts', Kant insists, is given 'not objectively, but only subjectively.' At this point, however, the sceptical reader of Kant may be inclined to question how much the argument has achieved. For once the actual existence of a supreme external authority is put to one side, it may appear unclear how our rational will or rational choice, taken on its own, can be enough to supply the requisite authority (for a defence of Kant on this issue, see Korsgaard (1996)).

For Butler, following a long religious tradition, things are much simpler: the authority of the 'natural' inner voice of conscience derives from its being implanted in us by God. Conscience is our 'natural guide, the guide assigned to us by the Author of our nature' (Butler (1726), III, 5). Yet for all its apparent straightforwardness, this solution is not without its own problems. For if the moral authority of conscience derives from its supposedly divine source, a question may still arise as to the basis of the moral authority attaching to the source itself (a question that has roots in the so-called 'Euthyphro dilemma' first raised by Plato). The mere fact that a supremely powerful being issues commands, or implants them in our hearts, does not in itself seem enough to endow those commands with moral authority; if on the other hand we say we say we know the commands should be obeyed because we are told so by the authoritative voice of conscience, this simply takes us back to the original question of what endows our moral sense with the requisite normative status. Notwithstanding such worries, many theists have taken it as a primary datum that we do indeed have an authentically authoritative faculty of conscience, and have used this as a premise providing grounds to infer the existence of God. Thus in the latter nineteenth-century, Cardinal Newman wrote in his *Grammar of Assent* that 'in this special feeling of Conscience, which follows on the commission of what we call right or wrong, lie the materials for the real apprehension of a Divine Sovereign and Judge' (Newman (1870), Ch. 5).

Two decades earlier, John Stuart Mill, writing from an entirely secular perspective, in his essay *Utilitarianism*, had provided an analysis of conscience that put pressure on this type of argument from conscience. In line with the 'empirical naturalist' movement referred to at the end of section 2 above, he defined 'the essence of conscience' as 'a feeling in our own mind; a pain more or less intense, attendant on violation of duty' (Mill (1861), Ch. 3). He adds various qualifications – that the feeling must be 'disinterested', and connected with the 'pure idea of duty' – but the main effect of his account is a deflationary or demystifying one – to reduce the deliverances of conscience to nothing more than a set of psychological events or purely subjective feelings. The feelings, he observed, are typically 'encrusted over with collateral associations', derived from the 'recollections of childhood' and 'all the forms of religious feeling'; and this, he claims, is enough to explain away 'the sort of mystical character which … is apt to be attributed to the idea of moral obligation.'

Mill's account purports to be simply a piece of empirical psychology, but it clearly has serious implications for the normativity of conscience. Painful feelings linked to the violation of duty function as what Mill terms 'internal sanctions', and he wished to enlist these in the service of his own utilitarian ethics. But sanctions, as understood by Mill, are no more than causal motivators – means whereby a desired code may be inculcated into the population so as to reinforce allegiance; we are thus in the territory of inducements for compliance, not in the territory of authoritative reasons for action. Mill was sensitive to the objection that if what restrains me from wrongdoing is 'only a feeling in my own mind', one may be tempted to think that 'when the feeling ceases, the obligation ceases.' But he confines his reply to observing that those who believe in a more exalted and objective source of obligation are just as likely to transgress morally as those who think that what restrains them 'is always in the mind itself' (ibid.). Whether or not his account is ultimately satisfying, Mill's interesting observations about how the feelings of conscience are often 'encrusted' with associations formed in childhood anticipate key elements in later psychoanalytic approaches to conscience, to be discussed in section 5 below.

4. Guilt and shame

It appears to be part of the concept of conscience that it involves not just the evaluation of one's past conduct, but also, when the evaluation is negative, a characteristic sense of discomfort. When reviewing what we have done yields a satisfactory verdict, then the conscience is 'quiet'; but when, as often, there is much to regret, then we speak of 'the pangs' of conscience, or being 'pricked' by conscience. Just as imprudent eating produces subsequent indigestion, so immoral conduct generates an unquiet conscience.

Precisely what form this takes will no doubt vary, but one very characteristic human reaction is to feel *ashamed* at what one has done. Though shame is a complicated reaction, at the most basic level it seems to involve an awareness that one's faults have been, or may be, exposed to view. Thus in the ancient story of the Fall of Mankind, Adam, after his act of disobedience, becomes aware for the first time of his nakedness, and hides himself. 'Who told you that you were naked?', God asks; and then, in the archetypal stern voice of authority, 'Have you eaten from the tree from which I told you not to eat?' (Genesis 3:11). The association between shame and nakedness does not, as a widely parroted caricature might suggest, depend on any particular Judaic or Judaeo-Christian 'hang-up' about sexuality, but is found in many cultures. Thus even among the ancient Greeks, who were perfectly unembarrassed about exercising naked, the genitals were commonly referred to as *aidoia*, a derivative of *aidos*, 'shame' (cf. Bernard Williams (1993), Ch. 4). Shame is a matter of being 'embarrassed', in the widest sense of that term: being seen by others in a setting where your untoward behaviour is the object of a certain class of 'participant-reactive attitudes' – in this case negative ones such as scorn, contempt, ridicule, reproach, disapproval, blame, and so on (Strawson (1962); cf. Taylor (1985)).

This cluster of reactions appears to extend over wider territory than the domain of what we normally call 'morality'; and indeed shame may be, and often is, felt in situations (such as being accidentally locked out of one's house wearing only one's underwear) where moral censure is not in question. Reflecting on this kind of contrast has led some thinkers to subscribe to a distinction between the kinds of standard that hinge on a person's retaining the honour and esteem thought due to them – keeping their dignity, or not 'losing face' – and the standards that focus on guilt and individual responsibility. In line with this, some anthropologists have suggested a classification of cultures into *guilt cultures* and *shame cultures*, where the former type of society places great emphasis on ideas of conscience, personal accountability and liability to blame and punishment, while the latter emphasises personal status or standing, as measured in terms of public esteem or its forfeiture. (For such classifications, and problems with them, see Deigh (1996).)

This kind of distinction has sometimes been thought to imply a contrast between an ethically immature 'shame' system in which all the weight is placed on what others think, or find out, about me, and a more sophisticated 'guilt' system in which personal responsibility and inner moral integrity are paramount (Atkins (1960); cf. Dodds (1951). Against this, Bernard Williams has argued in his landmark study *Shame and Necessity* that a study of ancient texts such as the Homeric epics in which shame plays a major role shows that 'even if shame and its motivations always involve in some way … an idea of the gaze of another … for many of its operations the imagined gaze of an imagined other will do' (Williams (1993), Ch. 4, p. 82). Williams makes a persuasive case for supposing that what he calls the 'internalized other' is typically crucial to how shame works. If this is right, then since (as we have seen in connection with Kant's account) conscience and guilt equally appear to involve the idea of an 'internal forum' in which we are judged by an idealized other, it begins to look as if the contrast between shame and guilt systems is not nearly as sharp as is often claimed.

Nevertheless, Williams does manage to retain elements of the contrast between shame and guilt, in a way that furthers his own distinctive agenda in the philosophy of ethics, namely his critique of what he called the 'morality system' – that 'peculiar institution', with its associated idea of a special class of inescapable obligations (Williams (1985), Ch.10). Williams felt that this institution exerted a kind of tyranny over our thinking about ethics, and that we would be better off without it; and he further argued that reflecting on the different landscape of Greek ethical culture would help us to see how such an escape might be possible. Once we acquit ancient Greek ethics of the charge of being preoccupied with public esteem and the dangers of 'losing face', we start (on Williams's view) to see that their judgements about shame were reaching towards a much richer conception of how ethical evaluation might work, and one that is arguably in better shape than our own. So instead of the presumptuous 'progressivism' which looks back on the ethical outlook of the Classical world as falling short of our modern Western understanding of morality, we might learn from seeing how the mechanisms of shame are displayed as motivating the characters found

in (for example) Sophoclean tragedy: '[these] characters are represented … as experiencing a necessity to act in certain ways, a conviction that they must do certain things … The source of the necessity is in the agent, an internalised other whose view the agent can respect … The sense of this necessity lies in the thought that one could not live and look others in the eye if one did certain things.' For agents to have this kind of conception of what must be done does not, on Williams's argument, require them to acknowledge moral imperatives grounded in 'external reasons'. Rather, 'these necessities are internal, grounded in the *êthos,* the projects, the individual nature of the agent …' (Williams (1993), Ch. 5, p. 103).

Complex issues are raised in this passage, whose full explication would require us to move farther afield than is possible here – for example into Williams's interest in the 'genealogy of morals', and his associated attraction for aspects of the philosophy of Nietzsche. But what I think one may nevertheless discern in these arguments is a programme for the radical deconstruction of the Christian concept of conscience, with its vision of the authoritative divine voice (or its Kantian analogue) calling each of us to account for the moral quality of our actions. We are offered instead a Nietzschean vision of an ethical world in which values are grounded in no more than the 'individual nature of the agent' and his or her chosen 'projects' (Nietzsche (1887), First Essay; Nietzsche (1886) §203). Whatever one makes of this (and many will feel serious misgivings about the attempt to establish ethics on such an individualistic base), there is one general lesson on which both supporters and critics of Williams may agree, namely that philosophical discussion of concepts such as conscience, guilt and shame cannot fruitfully be conducted in abstraction from the history of how those concepts developed, or in isolation from the underlying worldview in terms of which they were shaped. Williams's own resistance to the 'morality system', as he implicitly makes clear at the close of *Shame and Necessity*, stems from a worldview that wholly rejects any absolute moral framework for understanding the human predicament: 'our ethical condition … lies not only beyond Christianity, but beyond its Kantian … legacies … We know that the world was not made for us, or we for the world, that our history tells no purposive story, and that there is no position outside the world or outside history from which we might hope to authenticate our activities' (Williams (1993), p. 166). The authoritative status of conscience, at any rate as traditionally conceived, does not look likely to survive in such a world; but if Williams' arguments are accepted, it is a world in which agents will still be able to measure themselves against powerful standards they will be ashamed to fall short of, even if those standards depend in the end on no more than the individual projects they have chosen to make their own.

5. Conscience and the Unconscious

The concepts of the 'internal forum', in Kant, or of the 'internalized other', as deployed by Williams, will for many modern readers have unmistakeable psychoanalytic overtones. Irrespective of these specific resonances, however, any attempt to provide a philosophical introduction to the topics of conscience, guilt and shame would be seriously incomplete if it failed to include some reference to how the ideas of Sigmund Freud and his successors have influenced our understanding of the concepts in question.

In a lecture expounding his views on the sources of psychological conflict, Freud discusses some of his patients who were in the grip of a persistent delusion that they were being observed; and he describes how such cases gave him the idea that that the *separation* of the observing 'agency' from the rest of the ego might be a regular feature of the ego's structure. 'There is scarcely anything', he observes, 'that we so regularly separate from our ego, and so regularly set over against it, as our conscience'. In the typical case, I want to do something that gives me pleasure, but my conscience makes me abandon it; or, if I go ahead 'my conscience punishes me with distressing reproaches'. This kind of familiar experience, Freud argues, makes it appropriate to talk of an independent agency in the ego, with a separate existence – what he labels the *Über-ich* – the "Over-I" or (in the Latinized equivalent used in the standard English edition of Freud) the *Superego* (Freud (1933), pp. 485-7 (Lecture 31, para. 6)).

At first sight this might look like a straightforwardly deflationary view of the conscience, of the kind found in Mill (see above): in place of the authoritative voice of God implanted in the soul, there is simply an 'internal sanction', that is so say a painful feeling, produced by childhood conditioning, that motivates me to conform to prevailing rules, and punishes me when I fail to do so. Freud would certainly have rejected any theistic view of the origins of conscience – indeed, he took a generally sceptical view of religious belief, arguing that it was an illusion generated by the infantile need for security in a hostile world (Freud (1929) p. 260). But there is much more to Freud's view of conscience than a denial of its divine origins. The crucial thing to note is that the Superego, for Freud, exerts its power precisely in virtue of not being directly accessible to consciousness. On the Freudian analysis, then, the operation of conscience is often more complicated than could be explained by a simple internalizing of a moral rule (such as 'thou shalt not steal!), backed by feelings that inhibit me from stealing and generate remorse should I transgress. What happens alongside this is often something potentially traumatic, a kind of 'dissection of personality', in which the Superego 'seems to have picked out the parents' strictness and severity, their prohibiting and punitive function, whereas their loving care seems not to have been maintained'. The result is a hapless state in which the individual, without fully consciously appreciating what is happening to him, is locked into an aggressive cycle of self-criticism, as he strives to fulfil ever greater demands for perfection. '[The child's] aggressiveness [towards the heavily controlling parent] is introjected, internalized, sent back whence it came, that is, it is directed towards his own ego. There it is taken over by a portion of the ego, which sets itself over against the rest of the ego as Superego, and which now, in the form of 'conscience' is ready to put into action against the ego the same harsh aggressiveness that the ego would have liked to satisfy upon others' (Freud (1929), p. 315).

Successors of Freud such as Jacques Lacan, have laid great stress on this characteristic harshness of conscience, and its damaging effects in the psyche. Lacan in a memorable phrase talks of the *gourmandisme* of the Superego: the more cruel the conscience becomes, the more inescapably it commands obedience, and yet the very obedience only serves to generate an ever more stringent and searching inspection of the guilty secrets of the ego. Like a sinister parasite, the more you feed it, the more it wants (Lacan (1986); cf. Rajchman (1991), p. 58). Such vivid characterizations of psychic conflict may well strike a chord in many people's minds, but they have possibly had the effect of alienating mainstream philosophical writers on ethics, many of whom are suspicious of psychoanalytic ideas, and tend to approach the arena of morality using only the more straightforward vocabulary of reasons for action, practical deliberation, and an agent's ordinary and relatively transparent projects, desires and beliefs (interesting exceptions include Scheffler (1992); Lear (2000)).

One reaction to this apparent stand-off between psychoanalytic and 'mainstream' approaches to ethics would be for moral philosophers to acquiesce in a compartmentalizing of the subject – a phenomenon already increasingly prevalent in many other areas of philosophical inquiry. Analytic ethicists, on this scenario, would continue largely to ignore the work of Freud and his successors, while writers sympathetic to psychoanalytic ideas would gravitate towards those specializing in 'continental' styles of philosophizing. Such an outcome, however, seems likely to lead to a serious impoverishment of philosophical understandings of concepts such as conscience, guilt and shame, which, as we have already seen, appear very hard to explicate without something like the idea of an internal forum, where the subject experiences some kind of division or conflict (however these notions are eventually to be unravelled). It may therefore be useful to conclude this part of our discussion by asking how, if at all, psychoanalytic insights into the workings of conscience might be integrated into the rest of philosophical ethics.

The starting point for such a reconciling project might be to note that although many of Freud's theories were developed in the context of his treatment of patients with severe psychological disorders, certain lessons may nonetheless be drawn that have a wider application to the ordinary human ethical predicament – a project that is prominent in much of the work of Carl Jung (cf. Jung (1931), p. 40). The result which both Freud and Jung envisaged as the goal of

therapy was an integrated condition in which the split-off parts of the self were reincorporated into consciousness. The precise details of this envisaged process are explained differently in different psychoanalytic theories, but the central idea is that a psychologically healthy person is one who is *maximally self-aware* – who is able to confront the various desires, inclinations, fears, aspirations and so on that arise within him, but in their true colours, freed from distortions and projections (Jung (1931); cf. Cottingham (1998), Ch. 4). Applying this to the specific case of conscience, and to the ordinary moral conflicts that most of us face, it seems clear that all but the most perfectly virtuous souls will need the support of feelings of guilt and shame to keep them on the right path. Things go awry not because such feelings exist, but when they are invested with an awesome and tyrannical power that distorts the agent's deliberations so that he is incapable of balanced and rational assessment of how he should act, or unable to reflect on his conduct except with paralysing shame or self-hatred. But self-awareness, self-scrutiny, and a preparedness to delve into the sources of our mistakes, seem to be necessary components in the ethical life of any human being who aspires to integrity and responsibility, as well as being preconditions for the kind of genuine remorse and repentance that allows for healing, moral improvement and growth.

This brings us back to our main theme of the origins of conscience, and the source of its authoritative power. It would take an absurdly blithe and blinkered view of the human condition to fail to acknowledge the moral flaws of our species – the violence and cruelty and the aggressiveness and selfishness of which we are capable. And it does not require a very sophisticated anthropology to realize that the survival of civilization depends on these dire tendencies being held in check, both by the 'external sanctions' of law and social pressure, and by the 'internal sanctions' of conscience, guilt and shame. But where does this internal, or 'internalized' voice come from? Simply to say that the relevant feelings were implanted in us as children by our parents is only a very short term explanation, since the question immediately arises as to how the relevant feelings were transmitted to them, when they were children. Freud was quick to grasp this point: 'parents and authorities follow the precepts of their own super-egos in educating children … Thus a child's super-ego is in fact constructed on the model not of its parents' [ego], but of its parents' superego [and thus] it becomes the vehicle for … the judgements of value which have propagated themselves from generation to generation' (Freud (1933), p. 493). Explaining away conscience in terms of a parental voice is, in itself, no more satisfying than explaining away the idea of God by saying that I received it from my parents, since it simply postpones the question of ultimate genesis. A systematic inquiry into the origins of conscience would raise many fascinating and controversial issues, but given the constraints of space a few brief reflections will have to suffice to bring our discussion to a close.

6. Conclusion: conscience and conflict

The persistence, down the generations, of feelings of conscience must surely come from roots that are deep in our human nature, namely that we aspire to the good, and yet are often drawn to evil. Given what seems indisputable, namely our inherent *conflictedness* in this respect, conscience can be seen, as it has been seen by many religious writers, as a necessary ally in the pursuit of the good – something, as traditional language has it, that speaks for our 'higher' nature, against the promptings of the lower. Casting conscience in this perhaps somewhat exalted but at any rate beneficial role turns out on reflection to be quite compatible with psychoanalytic insights about its harshness; for the harshness, as we have seen, stems from its being 'separated off' as an alien voice of punishment and control, whereas the idea of moral growth presupposes that this voice can be freed from its tyrannical overtones, once it is properly integrated into a healthy self-image, and seen as directing us to the good where our true fulfilment lies.

The final question this leaves us with is whether such a positive account presupposes, in the end, a theistic underpinning for the notion of conscience. It is a question that is by no means easy to answer. There seems on the face of it to be a perfectly coherent alternative, namely that the operation of conscience (maintained and transmitted down the generations, and no doubt given specific shape by the mores of society in any given epoch) derives ultimately from no more than

contingent facts about the way our species happens to have evolved. David Hume in the eighteenth century was perhaps the most eloquent spokesman for such a view, when he proposed that morality is founded ultimately on the natural feelings of benevolence we find within us. Moral virtues, he remarked, have a 'natural beauty and amiableness' which 'recommends them to the esteem of uninstructed mankind and engages their affections'. An integral part of this is the operation of conscience – although, consistently with his sunny vision of human nature, what Hume stresses is not the torments of guilt and shame, but the joys of a quiet conscience – 'inward peace of mind, consciousness of integrity, a satisfactory review of our own conduct' (Hume (1751) Sectn IX, part 2).

Yet the conflictedness of our nature cannot be ignored, even by someone who takes such a benign view as Hume. Along with the 'particle of the dove' which he maintained was 'kneaded into our frame', he was obliged to acknowledge 'elements of the wolf and serpent' (Hume (1751), Sectn IX, part 1). The dove, symbol of peace, is of course also the ancient Christian icon of the Holy Spirit (Matthew 3: 16), though one assumes this is not an echo that Hume can consciously have intended to evoke, given his atheism, or at least his resolute scepticism about the 'ultimate springs and principles' of the cosmos (Hume (1748), Sectn IV, part 1). The image of the wolf, by contrast, calls to mind the tag of the Roman poet Plautus, '*homo homini lupus*', recapitulated by Thomas Hobbes, when he observed in the *De Cive* 'that Man to Man is an arrant Wolfe' (Hobbes (1642), dedicatory letter) – a comment consistent with his famous account in the *Leviathan* of the natural state of humankind as the 'war of every man against every man' (Hobbes (1651), Ch. 13). Hume's other iconic image, that of the treacherous serpent, unavoidably conjures up the story of the Fall of Man, the archetypal narrative of human guilt and the unquiet conscience. At all events, the combined effect of these juxtaposed images might seem to be to underline the precariousness of the human condition, beset by contradictory impulses; and given this, it seems plain that the kind of guide to action that conscience has traditionally been supposed to provide cannot be underwritten simply by appealing, along Humean lines, to empirical facts about our nature. More is required; but what that 'more' amounts to is a matter of continuing philosophical debate. One thing at least should by now be clear from our discussion of conscience, guilt and shame, namely that any plausible philosophical account of their role in the good life will need to be able to draw on the resources of a moral psychology rich enough to do justice to the cognitive and affective complexity of our human makeup.

JC

References

Aquinas, Thomas (1256-9), *Disputed Questions about Truth* [*Quaestiones disputatae de veritate*]. English version, Washington DC: Regneri Publishing, 1954.

Aristotle (*c*. 325 BC), *Nicomachean Ethics*. Transl. by T. Irwin (Indianapolis, Ind.: Hackett, 1985). [Also in *The Ethics of Aristotle*, transl. J. Thomson, rev. H. Tredennick (Harmondsworth: Penguin, 1976).]

Atkins, A. H (1960), *Merit and Responsibility: A Study in Greek Values*. Oxford: Clarendon.

Butler, Joseph (1726), *Fifteen Sermons*. Extracts may be found in J. B. Schneewind (ed.), *Moral Philosophy from Montaigne to Kant* (Cambridge: Cambridge University Press, 1990), and also in D. D. Raphael (ed.), *British Moralists 1650-1800* (Oxford: Clarendon, 1969*).

Cottingham, John (1998), *Philosophy and the Good Life: Reason and the passions in Greek, Cartesian and psychoanalytic Ethics*. Cambridge: Cambridge University Press,

Darwall, Stephen (1995), *The British moralists and the internal 'ought' 1640-1740*. Cambridge: Cambridge University Press.

Deigh, John (1996), *The Sources of Moral Agency*. Cambridge: Cambridge University Press.

Descartes, René (1641), *Meditations: Second Set of Replies to Objections* [*Meditationes:* Secundae Reponsiones] in J. Cottingham, R. Stoothoff and D. Murdoch, *The Philosophical Writings of Descartes*, Vol. II (Cambridge: Cambridge University Press, 1985).

Descartes, René (1648), *Conversation with Burman*, transl. in J. Cottingham, R. Stoothoff, D. Murdoch and A. Kenny, *The Philosophical Writings of Descartes*, vol. III, The Correspondence, (Cambridge: Cambridge University Press, l991), pp. 332ff.

Dodds, E. R. (1951), *The Greeks and the Irrational*. Berkeley: University of California Press.

Donagan, Alan (1977), *The Theory of Conscience*. Chicago: University of Chicago Press.

Freud, Sigmund (1929), *Civilization and its Discontents* [*Das Unbehagen in der Kultur*], in *The Penguin Freud Library* (London: Penguin Books, 1985), Vol. 12.

Freud, Sigmund (1933), *New Lectures on Psychoanalysis* [*Neue Folge der Vorlesungen zur Einfürhung in die Psychoanalyse*], in S. Freud, *The Essentials of Psychoanalysis*, ed. Anna Freud, (London: Penguin, 1986), pp. 484ff. [Also in *The Penguin Freud Library* (London: Penguin Books, 1985), vol. 2, pp. 88ff.]

Hume, David (1648), *An Enquiry concerning Human Understanding*, ed. T. Beauchamp (Oxford: Oxford University Press, 1999).

Hume, David (1751), An *Enquiry concerning the Principles of Morals*, ed. T. L. Beauchamp (Oxford: Oxford University Press,1998).

Hobbes, Thomas (1642), *De Cive.* English version (possibly by Hobbes) published 1651. [Both the English and the Latin versions are available in the edition of H. Warrender (Oxford: Clarendon, 1984).]

Hobbes, Thomas (1651) *Leviathan*, ed. R. Tuck (Cambridge: Cambridge University Press, 1996).

Jerome, St., *Commentary on Ezekiel* [*Commentariorum In Ezechielem Prophetam Libri Quatuordecim*], in J. P. Migne (ed.), *Patrologia Latina* (Paris, 1857-66), Vol. 25.

Jung, Carl (1931), 'Problems of Modern Psychotherapy' [from *Seelenprobleme der Gegenwart*], in *Modern Man in Search of a Soul: Essays from the 1920s and 1930s,* transl. C. F. Baynes (London: Routledge, l933), Ch. 2.

Kant, Immanuel (1785), *Groundwork for the Metaphysic of Morals* [*Grundlegung zur Metaphysik der Sitten*], transl. and ed. T. E. Hill and A. Zwieg (Oxford: Oxford University Press, 2002).

Kant, Immanuel (1797), *The Metaphysics of Morals: Doctrine of Virtue* [*Metaphysik der Sitten: Tugendlehre*], ed. M. Gregor (Cambridge: Cambridge University Press, 1996), Part II.

Korsgaard, Christine (1996), *The Sources of Normativity*. Cambridge: Cambridge University Press.

Lacan, Jacques (1986), *Le Seminaire: Livre VII: L'éthique de la psychanalyse 1959-60*. Paris: Seuil, l986. [English version: *The Ethics of Psychoanalysis* transl. D. Porter (New York: Norton, 1992.]

Lear, Jonathan (2000), *Happiness, Death, and the Remainder of Life*. Cambridge, Mass: Harvard University Press.

Locke, John (1670), *An Essay* A*n Essay concerning Human Understanding*, [1689], ed. P. Nidditch. (Oxford: Clarendon, repr. 1984).

Mill, John Stuart (1861), *Utilitarianism*. Ed. R. Crisp (Oxford: Oxford University Press, 1998).

Newman, John Henry (1870), *An Essay in Aid of A Grammar of Assent.* London: Burns, Oates & Co. (Facsimile reprint, Elibron Classics: Adamant Media Corporation, 2007.)

Nietzsche, Friedrich (1886), *Jenseits von Gut und Böse*, transl. W. Kaufmann, *Beyond Good and Evil* (New York: Random House, 1966).

Nietzsche, Friedrich (1887), *Zur Genealogie der Moral*, transl. D. Smith, *On the Genealogy of Morals* (Oxford: Oxford University Press, 1996).

Rajchman, J. (1991), *Truth and Eros: Foucault, Lacan, and the question of ethics*. New York: Routledge.

Scheffler, Samuel (1992), *Human Morality*. New York: Oxford University Press.

Strawson, P. F. 'Freedom and Resentment' (1962), repr. in Strawson, *Freedom and Resentment* (London: Routledge, 2008), Ch. 1).

Taylor, Gabriele (1985), *Pride, Shame and Guilt*. Oxford: Clarendon.

Wallace, James D. (1978), *Virtues and Vices* (Ithaca, NY: Cornell University Press).

Williams, Bernard (1985), *Ethics and the Limits of Philosophy*. London: Collins.

Williams, Bernard (1993), *Shame and Necessity.* Berkeley: University of California Press.