Moral Judgment, Action and Emotion

BERNARD HARRISON

> There is no general doctrine which is not capable of eating out our morality if unchecked by the deep-seated habit of direct fellow-feeling with individual fellow-men.
>
> George Eliot, *Middlemarch*, Chapter LXI

I

What makes us responsive, however occasionally, to moral demands? Why do people sometimes own up, go off to fight unwillingly in what they consider to be just wars, refrain from stealing a march on friends, and so on, even when they could by doing otherwise reap advantages far outweighing, in the scales of ordinary prudential rationality, any consequent disadvantage? Why has morality such a hold over us?

I want to begin by considering a set of supposedly complementary answers deriving in the first instance from Hume. Hume distinguished between 'natural' virtues such as clemency and 'artificial' virtues such as justice. He held that willingness to practise virtues of the first type, but not the second, could be explained by appeal to the operations of *sympathy*, that term being taken in a technical sense which, while etymologically exact, is not quite the sense which it bears in the ordinary English of our own, or, I suspect, of Hume's day. Our ordinary use of the term leaves it ambiguous whether sympathy is just community of feeling, as when we sympathize with someone who has suffered bereavement, or whether it also involves an element of judgment, as when we speak of sympathizing with a complaint made by one person against another. Hume's technical sense pushes the term towards the emotional end of its spectrum of ordinary meanings. Humean 'sympathy' is simply the capacity of one person's feelings to resonate in accord with those of another, and it is this mechanism that Hume takes to account for benevolence, as witness his occasional use of the phrase 'sympathy and benevolence' as a kind of compendium term.

'Artificial' virtues are those whose practice cannot be explained in this way because their exercise often compels us to restrain the natural flow of sympathy, as when justice requires us to restore 'a great fortune to a

miser, or seditious bigot'. Why, then, do people practise such virtues? The answer Hume offers starts from a roughly Hobbesian foundation. It is in everyone's interest that there should be some stable set of arrangements, framed by 'artifice and human convention' for regulating the acquisition and transfer of property, rather than a Hobbesian war of all against all. However, once a given system of regulative conventions has come to be generally accepted *de facto* in a society, several mechanisms operate to create sentiments of approval towards the system and its operations on the part of individual citizens.

Hume describes four such mechanisms. Firstly, there is sympathy itself, which is brought into play by the fact that the interests of others are now bound up *de facto* with the accepted systems of conventions. Secondly, there is the displacement of self-interest which arises from reflection upon the advantages of the system for each individual citizen. Thirdly, there is concern for reputation. Fourthly, and finally, there is the encouragement by 'politicians' of such sentiments on the ground of their social utility, and by parents on the ground that a man endowed with them 'is the more useful, both to himself and others'.[1]

Hume's mechanisms for mobilizing moral sentiment evidently can be, and often are, presented as ways of explaining the occasional readiness of people to respond to moral demands in defiance of their interests. So reinterpreted, Hume's discussion yields four types of explanation, proceeding respectively by appeal to the following notions:

1. Sympathy (in Hume's special sense)
2. Displaced self-interest
3. Fear of censure, ridicule, or other social sanctions
4. Socially and parentally conditioned approval and disapproval

J. L. Mackie, in recent writings,[2] maintains that these mechanisms form a connected system of mutually supporting elements. This can be taken in two ways, it seems to me, in one of which it is contestable. Certainly if we just take it for granted that the existence of responsiveness to moral demands must be explained in some such ways as Hume suggests, it seems perfectly reasonable to suppose that Hume's four mechanisms might form a *causally* connected system, with each reinforcing the effects of the others in directing the practical reasoning of individuals. But if what is at issue is the capacity of Hume's mechanisms to provide explanations of people's readiness to disregard their interests in response to moral claims, then it is less clear that the four explanations gain any cumulative force from their association with one another. In fact, as I shall argue, each is in-

---

[1] David Hume, *A Treatise of Human Nature*, III, ii, 2.

[2] J. L. Mackie, *Hume's Moral Theory* (London: Routledge and Kegan Paul, 1980), 87–88, 149.

sufficient in itself as an explanation of moral response, and gains nothing in this respect from its association with others.

Explanations of type 2 notoriously fail to deal with the numerous class of cases in which failure to comply with a generally accepted moral rule has no tendency to diminish the general level of acceptance of and compliance with that rule, and thus no tendency to diminish the advantages accruing to individuals from general acceptance and compliance. Some other explanation must therefore be found for responsiveness to moral demands in such cases; perhaps one of type 3.

Explanations of type 3, however, because they found moral response upon ordinary social prudence, fail to explain why people attach weight to moral considerations when there is no reason to fear any penalty for neglecting the counsels of that kind of prudence, or to hope for any reward for heeding them. The keeping of promises made in secret to dying friends and the doing of good by stealth are hoary examples of the sort of thing I have in mind. Still more embarrassing for type 3 explanation is the observable willingness of people to persevere in types of virtuous conduct which attract only contempt, misunderstanding and obloquy, conscientious objection being an obvious example.

There is also an interesting range of awkward cases which I shall lump under the heading *commonplace supererogatory heroism*. I have in mind the sort of case in which some ordinary citizen does something very brave which he was in no sense morally required to do: which was in no sense, as we say, 'expected of him'. A case of the kind, reported recently, involved a boy who steered his father's fishing boat close to a treacherous reef of rocks in a boiling sea in a successful attempt to rescue a drowning man. The boy knew the rocks and their reputation, and knew that the risk he was taking was so great that he was more likely to end by drowning himself than by saving the man in the water.

The first thing to be noticed about this kind of case is that to the extent that the boy's heroism is supererogatory he cannot be supposed to be under social pressure to act as he does. No one who knows the sea—no one in his immediate community, that is—will censure or ridicule him if he does nothing to save the man in the water, since those who know the sea know that, given the drowning man's proximity to the reef, there is nothing, rationally speaking, to be done. Conversely, if the boy does by some miracle succeed in saving the man in the water, his action in going to the rescue will not be universally applauded. Some will see it as a piece of schoolboy heroics; and it is not implausible that an actual boy in such a situation would be able to foresee this, having been warned often enough against foolhardy seamanship.

The distinction between *heroism* and *heroics* operating here is itself interesting. The case for regarding what the boy did as heroic is that he did what he did for moral reasons: in response to a moral claim. It would

have been *merely* foolhardy, that is, if he had steered close to the rocks to rescue a cat, or a floating oar. On the other hand what does 'for moral reasons' mean here? It cannot mean 'in obedience to a prescriptive principle accepted in his society', since it is precisely by appeal to the socially accepted principles governing conduct in such circumstances that we generate the case for regarding what he did as a piece of schoolboy heroics. Those principles tell us, for instance, that there are limits to what can reasonably be demanded of anyone in response to a moral claim of the sort made by the plight of the drowning man: limits which the boy should not have overstepped because in so doing he risked not only his own life but that of anybody else who might be foolhardy enough to come to his aid in turn. However, it seems that even general acceptance of this view of the matter need not render null and void, even in the minds of those who deprecate the foolhardiness of the boy's action, the moral claim made by the plight of the man in the water. That claim retains its force, and forms the basis of the opposing case: the case for regarding the boy's action as heroic despite its foolhardiness. It seems, in fact, that the status of certain states of affairs as presenting a moral claim is not simply a reflection of the acceptance within a community of certain corresponding prescriptive rules; and this makes it even harder to see, as it were, where the power which moral claims exercise over us *comes from*.

It is perhaps time to explore the resources of explanations of type 4, in terms of the social and parental conditioning of moral attitudes. Such explanations must, to succeed, live up to the demands of their own naturalism. If the appeal to conditioning is not to amount merely to hand-waving, that is, it must in principle be cashable out in terms of actual, temporally locatable processes extending over the considerable tracts of time that such processes require if they are to be effective.

Obvious difficulties for such a programme arise from the fact that individuals often arrive by independent moral reflection at conclusions at variance with the moral pieties current in the society in which they have been brought up. It is sometimes argued that moral heretics within one society always rely for emotional support and intellectual inspiration upon a secret or implicit adherence to some other society, real or imagined, which would give triumphal expression to the dissident values for which, it may be, the heretic suffers persecution in his own society. This is often true, but is beside the point. What we have to explain is not how the heretic is able to envisage dissident values, nor how he can find the emotional strength to defend them in the face of disapproval, but why he should attach moral weight to them in the first place. Explanations of type 4 tell us that to attach moral weight to some goal or principle of action is to feel a conditioned attitude of approval towards it, but that answer requires us to postulate an actual process of conditioning, which in the present case could have taken place neither in the heretic's own society

nor, evidently, in a society to which the heretic has never belonged, or which is merely an imaginary utopia.

One way of getting around this might be to argue that the heretic has done no more than modify values already recognized in his society, towards which he will already have acquired a conditioned attitude of approval. But then we have the problem of explaining the transfer of moral weight from the orthodox values to the heretic's modified versions. Such a transfer is easy enough to understand if all that the heretic has done is to bring orthodox values to bear in a new way upon some disputed case by a persuasive redescription. Thus somebody might argue that prostitutes deserve the respect accorded to anyone who undergoes unpleasantness in order to render a necessary service to society. But moral heresy can go beyond the mere redescription of cases. Suppose, as sometimes happens, that someone born and brought up in a staunchly working-class community comes to feel the moral weight of the kind of libertarianism which demands that individuals be left free to choose, if they wish, not to join a trade union. It seems quite clear that such a demand cannot be founded upon any factual redescription of what is involved in refusing to join a trade union which would bring such an act under the shelter of a conception of liberty which is already granted moral weight within the dissident's community. For the concept of liberty—of what a man may reasonably claim the right to do without hindrance from his fellows—as it is understood in the community expressly excludes acts held to be manifestly contrary to the common interests of working people as the community sees those interests. In the same way the supererogatory hero's sense of the overriding moral weight attaching to the plight of the man in the water could hardly have arisen through a descriptive assimilation of the drowning man's plight to a class of cases already invested with just such overriding weight by the common moral sentiment of the community. There is only one way of describing the man's plight which can give it moral weight—one which draws attention to the fact that he is not waving but drowning—and the verdict of common moral opinion upon the case viewed in this light is already known.

Someone might want to object at this point that it is absurd to claim that radical libertarianism, or self-sacrifice, are not values with a definite history and definite social constituencies. Of course such a claim would be absurd, but it is not one required by the present argument, since it is no part of that argument to claim that the values espoused by the supererogatory hero or the working-class liberal are ones which they have invented for themselves. There are no doubt plenty of channels through which they may have gained access to the liberal ideal of a right of conscientious objection or to the idea of a heroism which does not count costs. The problem is not to explain how people come to know about such ideals, but to explain why they should sometimes feel them to carry even

a *prima facie* moral weight, when the ideals in question are hostile to those instilled by social and family conditioning in the *milieux* in which they happen to live and to have been brought up. Of course the theorist of social conditioning could claim that moral points of view never do in fact leak under their own impetus across social and doctrinal boundaries; but that just seems factually implausible. There seems nothing for it but to withdraw to an explanation of type 1.

II

At first sight the appeal to sympathy might appear to solve most of the problems we have uncovered so far. By appeal to a Humean community of feeling between individuals we can explain how it is that we can sometimes take a direct interest in the sufferings and joys of others; and also, since on this account what is primarily at stake in moral response is *feeling*, which notoriously can sweep us off our feet, how it is that people can sometimes neglect their own interests in order to respond to moral demands. John McDowell, in a recent article, goes so far as to suggest that the existence of a direct and natural sympathy between persons constitutes a sort of explanatory gateway to the moral life: a way in which 'even the rankest outsider' might come to understand what morality is about, and how it is possible:

> . . . sheer bafflement at virtuous behaviour in general is very difficult to imagine. At some points even the rankest outsider would be able to attain a measure of comprehension of virtuous actions in terms of desires which people just naturally have: for instance the desire that people related to them in various ways should not suffer.[3]

My difficulty with both McDowell's position and Hume's, however, is that I seriously doubt whether any moral response, even those we most naturally ascribe to 'sympathy' and 'benevolence' in the moral senses of those terms, can plausibly be explained by appeal to any psychological disposition we 'just naturally have'.

The first difficulty with such an assimilation is that it requires us to equate, for instance, the moral weight of the drowning man's plight as felt by the boy with the strength of the emotional response which that plight evokes in the boy; to accept, more generally, the proposition that to attach moral weight to the plight of another person just *is* to experience sympathetic distress on his behalf.

[3] John McDowell, 'Are Moral Requirements Hypothetical Imperatives?', *Aristotelian Society Suppl. Vol.* 52 (1978), 19.

This seems altogether wrong. It is not difficult to think of examples in which the sympathy we feel for someone in a given plight and the moral weight we attach to the project of getting him out of that plight are altogether incommensurate. Suppose A, for example, has for years bullied his wife, not because of any viciousness or brutality in his nature, but simply because he is a big noisy man, clever in intellectual matters but stupid in human ones, who has generally been too absorbed in his work to notice what a domestic tyrant he has become. Finally his wife leaves him, and he is predictably distraught. He also constructs a moral case for himself, ascribing her departure to coldness and selfishness. His emotional distress resonates far more powerfully in me than does her relief at being free at last, because I like him, because I am aware of similar tendencies in myself, and because I am not the one who has to put up with him; whereas she is a reserved and private woman with whom I have never found it easy to get on. Nevertheless I attach no moral weight whatsoever to the project of persuading her to go back to him, which is the only thing that will cure his distress; nor do I feel the slightest temptation to accept his view of the moral rights and wrongs of the situation. Perhaps it will do him good to live through this passage of his life, but even if it will not, *he has no right to demand*, as we say, that his moods be coped with without limit. Far from its being one and the same thing to feel sympathy and to attach moral weight, sympathy here takes second place to moral considerations, which unsurprisingly (however puzzling the nature of the connection may be) are what we actually attach moral weight to in the situation.

This result will perhaps seem still less surprising if we turn to a second, and more traditionally canvassed, difficulty with the idea that readiness to respond to moral claims is the same thing as readiness to resonate emotionally in sympathy with the joys and distresses of others; the difficulty, namely, that the link which such a view forges between virtue and appetite is altogether too close. No doubt, of course, the appetite in question is a very delicate one. What delight could possibly be more innocent than the emotion which overcomes the possessor of a generous heart confronted with the commingled tears and joy of one whose life he has just saved? . . . and so on. The trouble is that the closer we push our description of heroic virtue towards the clichés of the Novel of Sentiment, the more we are forced to represent the hero as a gourmet of sentiment, whose response is less to the plight of the ostensible object of his virtue than to the opportunities which the situation affords for the enjoyment of certain refined gratifications about which he is crazy enough, to put it bluntly, to disregard his own safety. This consequence of investing the natural appetites of men with the power morally to validate their actions was apparent to Mandeville in 1714:

There is no merit in saving an innocent babe ready to drop into the fire:

the action is neither good nor bad, and what benefit soever the infant received, we only obliged ourselves; for to have seen it fall, and not strove to hinder it, would have caused a pain, which self-preservation compelled us to prevent . . . [4]

This passage strikes most modern readers as betraying a moral rigorism run mad, and in one sense it does. That people feel pain at the thought of children burning to death does not make the courage they show in preventing it any the less *courage*—provided, of course, that we can regard the act as a response to a moral demand, rather than as merely an attempt to avoid having to endure the pangs of sympathetic distress. If we are to escape Mandeville's conclusions, in short, we have to demonstrate the existence of a logical gap of some kind between the emotions and responses of morality and those of the natural man. Our ordinary conceptual scheme takes the existence of such a gap for granted. The moral theory of Sentimentalism, on the other hand, does its best to blur it, at the cost of opening the way to just the conclusions sought by Mandeville's curious philosophical Calvinism.

The logical gap which we ordinarily take to fall between the moral and the natural life shows, for instance, in the distinction to be drawn between benevolence in the moral sense and the casual generosity of the good-humoured man who, moral considerations aside, would always prefer to have happy faces about him. If there is an apparent literary *locus classicus* for the confusion between benevolence and joviality it is Dickens' *A Christmas Carol*, yet the distinction I have in mind is drawn there, explicitly or by implication, at every turn of the narrative. The party which Mr Fezziwig, that paradigm of the jovial host, gives for his employees is given, precisely, for people whom Fezziwig has a duty, as Dickens conceives it, to make as happy as he can. Had Fezziwig's generosity been exercised in favour of a party of cronies of his own age the act, however harmless and amiable in itself, would hardly have constituted an exercise of benevolence in the moral sense, and the moral point of the story would have been lost. Similarly, when Scrooge's heart has been broken by the Spirits of Christmas he does not rush into indiscriminate party-giving for any collection of people willing to come in off the street and help a crazy old man waste his substance; he begins to practise benevolence towards those who have a moral right to expect it of him: his clerk Bob Cratchit and his own neglected family. Had Dickens made Scrooge take the first of these possible choices, then once again the moral point of the story would have shifted. Scrooge's neglect of Cratchit and his family in favour of indiscriminate jollity would force us to conclude that the visitations of the Spirits had not

---

[4] Bernard Mandeville, *The Fable of the Bees*, Irwin Primer (ed.) (New York: Capricorn Books, 1962), 50.

after all left him with any clear sense of other people as presenting specific moral claims upon his benevolence, but merely with a desire to warm himself at the fire of boozy good-fellowship for the little time left to him. We should have a bleaker and more bitter tale, ending in a kind of dance of death.

Of course in one way Dickens' telling of the tale does betray contracted human sympathies: both the reclaimed Scrooge and his creator are more aware of moral claims originating within the circles of home and workplace than of no less pressing but more remote ones. But this distinction does not strengthen the case for equating moral impulse with sympathy in Hume's sense. Even the generalized benevolence which I owe, morally speaking, to those outside my immediate circle is owed, not just to anybody at all whose happiness might afford me a sympathetic pleasure, but to those in distress.

The logical connection between the notion of benevolence and that of distress offers an instance of the way in which moral concepts tend to form closed circles of mutually explicating notions. To the person 'inside' the circle of moral notions it is evident why the plight of the drought-stricken Turkana, say, creates a *prima facie* moral claim upon his generosity, while the fact that his colleague Soames would be pleased to receive an unexpected anonymous gift of £50 does not; but if we ask him to explain the difference he will have, from his 'inside' point of view, no option but to appeal to notions, such as that of *distress*, which to the moral outsider stand as much in need of explication as the concepts of *obligation* and (in the moral sense) *benevolence* themselves.[5] Looked at from this direction the various theories of the basis of moral responsiveness which we have considered so far all appear as attempts to understand 'from outside' how it is that to someone 'inside' the circle of moral concepts, moral considerations can sometimes appear to have greater weight than very powerful countervailing considerations of prudence and self-interest. Perhaps not very surprisingly, these attempts fail: but the question itself remains: what *does* give moral considerations weight in the practical reasoning of those for whom they are weighty? If an outsider's view will not help, perhaps we should try for a more careful articulation of what is involved in taking an insider's view of these things. We might ask, for instance, what role the circle of mutually explicating moral concepts plays in the moral insider's life, what is involved in entering that circle, and how entering it might change one's outlook on things. To these questions I shall now turn.

III

Suppose two people are playing a competitive game requiring manual

---

[5] Cf. David Wiggins, 'Truth, Invention and the Meaning of Life', *Proceedings of the British Academy*, **LXII** (1976), 343f.

skill, in which practice by either player during the game may reasonably be supposed to influence the outcome of the game. Player A is just about to take his shot when he has to leave the room. One of Player B's backers, C, says to him, 'Why don't you take a few practice shots while we're waiting for Arnie to come back?' Player B replies, 'No, I can't do that, it's A's turn', and on being pressed further repeats, 'No, I can't do it; it would be wrong. A trusts me.'

Plainly any of the standard sentimental explanations of moral scrupulousness could be pressed into service here to explain B's reluctance; equally plainly the circumstances surrounding such reluctances in real life are sufficiently complex and various to allow us plenty of room to formulate the example *ad hoc* in ways which would make such explanations look unconvincing when measured against 'ordinary moral intuition'. Perhaps ordinary moral intuition is confused and anthropomorphic; but, equally, perhaps there is some basis for those kinds of reluctance in real life which the standard explanations, invented to secure the coherence of a philosophical theory, help to obscure and obfuscate; and perhaps that explains why the language in which we ordinarily discuss moral questions obstinately continues to honour distinctions which evade philosophical formulation. On the surface, at any rate, it seems at least possible that the reasons for B's reluctance are just those he offers: (1) that it is A's turn; (2) that A trusts him; (3) that (1) and (2) entail (4) that it would be wrong to take practice shots in A's absence.

There are, of course, two familiar objections which are generally thought to obstruct such a reading. They are (a) the difficulty of explaining in what sense (1) and (2), which seem on the face of it to be statements of fact, can *entail* (4), which seems on the face of it to be a moral judgment; and (b) the difficulty of explaining how any of (1)–(4) can be *motivating* for B. It is important for the defence of sentiment theory, as of other forms of non-cognitivism, that these two questions be seen as essentially connected, as they are, for example, in Appendix 1 to the *Enquiry Concerning the Principles of Morals*. The reason there offered, as elsewhere in Hume, for doubting the possibility of relations of entailment between statements of fact and moral judgments is that the latter possess a power to motivate—an 'intrinsic action-guiding force' as recent discussion has it—which the former lack. This doubt serves to introduce the suggestion that the actual relationship between the two types of utterance is one of sentiment; and the essential correctness of this suggestion is then demonstrated by showing that it allows us to answer the question of how moral judgment, at least, can motivate us to action; the answer given by Hume and by modern non-cognitivists being that the function of a moral judgment is merely to give expression to some intrinsically motivating mental act, the act in question being conceived classically by Hume as a sentiment of approval or disapproval and by modern emotivists and prescrip-

tivists as respectively an attitude, conceived as 'a complicated conjunction of dispositional properties',[6] or an act of commitment or assent to some prescriptive principle. The intimate involvement of questions of logic with questions of motivation in the traditional arguments which found the non-cognitive position in itself suggests that we may begin to make progress in a contrary direction just by treating the question of entailment and the question of motivation as separate questions. Let us, then, begin with the first, and with turn-taking.

Turn-taking is a simple device, in the form of a Lewisian convention,[7] for producing a state of *de jure* equality between two or more participants. The kind of equality at stake can vary from case to case, and may, for example, be equality in the possession and use of some coveted object, or equality of access to a service, as in queueing, or, as in the present case, equality of opportunity to produce results affecting the outcome of a game. Given the nature of the convention, however, there is no difficulty about assigning truth-conditions to a statement of the form 'It's A's turn now': whatever the particular context in which the turn-taking convention is being operated, such a statement is true just in case the sequence of turns has in fact reached A, and false otherwise.

So 'It's A's turn' states a fact, but a fact which obtains only in virtue of the operation of a convention, the object of which is to equalize advantages between participants. Manifestly the convention cannot function properly to equalize advantages in the required manner unless each participant recognizes the right of every other participant to complain if, by accident or intent, his turn is usurped, and to claim restitution for such usurpation either in the shape of an extra turn inserted in the series to restore equality of turns up to a specified point T, or in the shape of some other advantage. It is to be noticed, now, that this right is not something which the participants could agree to recognize 'on moral grounds' subsequently to the institution of the turn-taking convention; for it would not be possible to institute the turn-taking convention without granting the general legitimacy of such claims. There is no effective difference between not allowing people to complain if they get fewer turns than others and not taking turns at all. Similarly it would be impossible in practice to operate the turn-taking convention to equalize the advantages of participants unless there were some way of deciding whether a claim for restitution of an allegedly usurped turn was justified or not. But the convention itself provides an empirical decision-procedure—that of checking back through the sequence of turns—which satisfies this condition. The same decision-

---

[6] Charles Stevenson, *Ethics and Language* (New Haven: Yale University Press, 1944), 60.

[7] David K. Lewis, *Convention* (Cambridge: Harvard University Press, 1969), 78–79.

procedure automatically establishes truth-conditions for statements of the form 'It's A's turn'. In fact, by reflecting on the reasons which require such statements to possess determinate truth-conditions if the turn-taking convention is to function in practice as a device for equalizing advantages, we can perhaps begin to see how it comes about that some statements manage to have both moral implications and straightforward truth-conditions.

I shall call the claim of a participant in a turn-taking convention for restitution of an usurped turn a *convention-backed claim*. A convention-backed claim is a demand by one person that others recognize and yield to some interest or purpose of his. It differs from other demands of this kind (demands for money backed by threats, for instance) in two respects:

(i) Acceptance of the *general* legitimacy of convention-backed claims of a given type is a necessary condition of participation in some Lewisian convention whose function in the relevant context is to equalize the advantages and/or expectations of participants, and in which both the person or persons who bring the claim, or on whose behalf it is brought, and the person or persons against whom it is brought, are participants.

(ii) The question of whether or not a particular convention-backed claim of a given type is justified can be settled empirically by the application to the circumstances of the particular case of the terms of the convention to which the claim in question appeals.

I want now to suggest that one common meaning of the adverb 'wrongly' —not by any means its only meaning but one which it bears in a great many unquestionably 'moral' contexts—is captured by the following definition:

X acted wrongly = X's action was such as to give rise to a justified convention-backed claim.

The interest of this definition depends, of course, on the extent to which one can regard the moral life as resting upon a foundation of conventions designed to equalize the advantages of participants, but such an interpretation covers at least a high proportion of ordinary moral claims. Promising is a case in point. The giver of a promise obtains the advantage of having his word relied on, while the person who accepts it obtains the correlative advantage of being able to rely on the performance of the promise. As in the case of turn-taking, the possibility of making convention-backed claims is built into the functional specification of the convention. If it were not in principle justifiable to complain about promise-breaking and promise-disregarding, and if there were no empirical decision-procedure, derivable from the form of the promising convention, for determining whether or not a promise has in fact been given, there would

be no effective convention of promising, since a promise would be effectively indistinguishable from a mere statement of intention.

Perhaps less obviously, it is not only possible but enlightening to regard the restricted notion of charity which has to do with the relief of distress as rooted in a general convention by which any human being suffering unmerited distress can appeal to the more fortunate for aid. Such a convention may or may not be regarded as cutting across the boundaries of social groupings (though it was so regarded, for example, by Aboriginal tribes who freely offered help in distress to early explorers of the Australian interior), but in either case the point of such a convention is clearly to equalize advantages by spreading the effects of misfortune through a community. The convention could not operate to equalize advantages in this way if it were not to be considered legitimate for those whose distress is neglected to demand the attention of the more fortunate if it is not freely forthcoming, since otherwise there would be no difference between an act of charity and an act of casual open-handedness; and no one would be able to place any more reliance upon the hope of receiving aid in distress than upon, say, the sporadically rewarded but entirely morally ungrounded hope of being treated to dinner by a more affluent friend. In short, what a person in distress appeals to is not the offhand chance of stimulating a brief flow of natural sympathy in potential benefactors, but his or her *rights* as specified by the terms of a convention in which such benefactors participate as much as he or she does. That the rights in question are limited in their extent by the function of the convention from which they arise, i.e. that of equalizing the effects of random misfortune by spreading them among the members of a community, is shown, for example, by the moral doubts entertained by trustees and the general public if a well-publicized disaster produces a flow of charitable donations so great that the money would, if fully disbursed to the survivors, not merely compensate their losses but make them rich.

Other conventions underlying various sorts of moral claim include the conventions of property (private or social as the case may be), confidentiality, the various conventions attaching reciprocal rights and duties to various social roles (parent, doctor, driver, for example) which give a sense to the concepts of negligence and betrayal of responsibility, truth-telling, and so on. To look at morality in this way, as a collection of conventional devices for equalizing advantage might appear at first sight to devalue the elements of spontaneity and warmth of feeling in moral response, but this appearance will diminish as the argument proceeds. At present we are dealing only with the conceptual structures which give sense to moral judgments and provide a foundation both for moral feeling and moral action.

The main object of the present section has been to explicate B's claim that the facts (1) that it is B's turn, and (2) that A trusts him, somehow entail that it would be wrong to take practice shots in A's absence. On

the analysis of 'X acted wrongly' suggested above the outline of the required explication seems clear enough. B is saying that his action in taking practice shots would be such as to give rise to a justified convention-backed claim against him by B, and that the conclusion that this is so can be seen to follow, given the nature of the convention-backed claim in question, from the terms of the turn-taking convention and from inspection of the sequence of turns since the start of the game. And he is clearly correct in this.

The effect of this analysis is, evidently, to turn 'wrong' into a descriptive term, wholly lacking in the 'intrinsic action-guiding force' popularly supposed to characterize the key terms of moral evaluation. And that seems to expose the project of taking B's reasons at face value to the second of the two difficulties we canvassed at the start of the section: that of saying how (1)–(4) can motivate. Once questions of motivation and questions of the logical status of moral discourse are prised apart, however, this difficulty too can be resolved, even if not altogether straightforwardly. The answer will require us to turn from 'It's A's turn' to look more closely at the logical grammar of 'A trusts me'.

IV

To trust somebody or something is at the most basic level simply to be willing to rely upon it, him or her. At this level nothing moral need be at stake: we can 'trust', in the sense of 'choose to rely on', a handrail or a footbridge as well as a human being.

In such trusting we very often ground our trust upon the presumed operation of some understood or partly understood mechanism. I trust the clock in the High Street, for example, because I presume that the position of its hands results from the working of an interior clock mechanism in good order, rather than, say, from their having been manually positioned by some joker. What prompts my trust is my ability to see that and (more or less) how the mechanism *must* produce certain results in certain specified circumstances of operation.

There is nothing to prevent us from regarding other human beings in this sort of light, and in fact we often do so regard one another and 'trust' one another accordingly. Sometimes the postulated mechanism is a matter of personality traits so ingrained as to operate more or less independently of the intentions or interests of the person concerned, as in the sort of case where we feel inclined to say things like, 'I'd trust John to lose his temper if Jim says things like that about foreign policy in his hearing'. Sometimes it is a matter of a postulated ordering of interests, as in the sort of case in which we say things like 'I think James can be trusted to react positively to anything which seems likely to keep the affair of the hospital tenders out of

the newspapers'. Either way, the principle of reposing trust upon the operation of a postulated interior mechanism which *must* produce certain results within a certain specifiable range of circumstances remains the same. I shall call this kind of trust *mechanically inferred trust*. It is perhaps intuitively evident that trust of this kind has an inherent tendency to exclude trust in the moral sense. To the extent, that is, that I trust John's ungovernable temper, or James's urgent need to keep the facts out of the newspapers, to produce certain results, it is John's temper or James's predicament that I trust, and not John or James. But this is to run ahead of the argument, whose immediate requirement is that we pin down what the moral sense, or use, of 'A trusts B' consists in.

Provided we can refer back to the notion of a convention whose function is to equalize the advantages of participants, and to the consequent notion of a convention-backed claim, we can define as follows a basic or logically minimal sense for 'A trusts B' as affirmed in moral contexts:

(1) A trusts B = A is willing to rely on B's unwillingness to allow a justified convention-backed claim of A's to be neglected or overridden.

Once again this might seem to give us an overly legalistic account of trust between persons. Certainly there is more to trust in the moral sense than this; but there is at least this. (1), that is, is what it claims to be: a statement of minimal conditions for the existence of trust in the moral sense between persons. It brings out the way in which the concept of trust, like other moral concepts, is rooted in certain specific kinds of convention.[8] An analysis of the concept of trust as deployed in moral contexts which made no reference to convention would presumably have to run along something like the following lines:

(2) A trusts B = A is willing to rely on B not to do anything which A would prefer him not to do.

But (2) is just wrong. To say that A trusts B is not to say, for instance, that A expects B always to be nice to him, or even that he expects him to refrain from pursuing his own interests if they happen to conflict with A's. It is clearly possible for two people to trust one another, for example, without either expecting the other to withdraw from competition for a job for which both have applied; while to suppose that one cannot trust a person as a friend without expecting him or her never to say what one does not wish to hear is just to confuse friendship with sycophancy. These and the like muddles arise when we try to treat even the most personal moral

[8] Cf. Julius Kovesi, *Moral Notions* (London: Routledge and Kegan Paul, 1967), 53–58.

relationships as 'natural' rather than 'artificial' in Hume's sense of those terms.[9]

Nevertheless, (1) is not adequate as it stands. It is manifestly possible for someone to be morally trustworthy in the bare or minimal sense captured by (1), either out of prudential calculation or, like Mr Casaubon in *Middlemarch*, out of egoism. Mr Casaubon, we may be sure, can be relied upon to discharge in a straightforward and honourable way any moral claim whose neglect, if discovered, would expose him to public censure, because in his intense and fastidious egoism he could not endure the thought that anyone might be in a position to bring a charge of dereliction of duty against him. George Eliot wants us to see among other things, I think, that there is nothing in this kind of rectitude capable of counteracting the effects of Casaubon's egosim in isolating him from relationships of genuine trust and openness with others, because there is nothing in it capable of preventing him from acting in a mean and unjust way when the possibility of public censure is absent. The episode of Casaubon's will, indeed, demonstrates the limits of a rectitude founded upon personal vanity.

What has to be noticed about this example is that anyone astute enough to grasp the connection between Casaubon's rectitude and his personal vanity has (i) good reason to repose a mechanically inferred trust in Casaubon's performance of any public duty, and (ii) equally good reason for not extending that trust beyond certain well-defined limits. And clearly the relationship between (i) and (ii) is a general, because a conceptual one. Any mechanically inferred trust is necessarily a trust which has limits which depend on the nature and circumstances of operation of the postulated mechanism to which it appeals.

By contrast, what I shall call *full-bloodedly moral* trust is neither mechanically inferred nor limited in extent. If I trust, in the full-blooded sense, an old friend not, let us say, to conceal documents whose absence exposes me to an unjust accusation, then (i) there are no circumstances in which I would expect him willingly and wittingly to do the opposite. (Of course, if he is drugged, or ignorant of the significance of the documents, or breaks down under torture, that is another matter. Full-blooded trust in my friend entails belief in the integrity of his will when he is in full command of himself and of the facts, not in the possession by him of superhuman fortitude or knowledge.) Moreover (ii) my reliance on him, despite its unconditional character, need not be justifiable in terms of inference from any postulated interior mechanism. I need not suppose, that is, either (a) that there is any permanent or transient trait of my friend's personality (lifelong commitment to a political cause, say, or domination by, or infatuation with, me) which can be relied upon to operate to my advantage in the present situation irrespective of his interests or, perhaps, his attempts

---

[9] Cf. J. L. Mackie, op. cit. note 2, Ch. vii.

to evade his own tendency to follow, in this case as in others, an ingrained and habitual pattern of behaviour; or (b) that his interests in the present situation compel him to consider mine. And, finally, (iii) it seems clear that if I did entertain beliefs about him of types (a) or (b), and if the beliefs in question were the basis of my belief that he would not willingly and wittingly conceal evidence favourable to my case, then my relationship with him would not be one of full-blooded and unconditional trust, but one of limited reliance upon his actions only to the extent to which these proved in fact to be reliably inferable from traits of his personality or from the current or long-term pattern of his interests.

We are thus led to the following summary characterization of full-blooded trust:

(3) A trusts B = A is willing to rely unconditionally on B's unwillingness to allow any justified convention-backed claim of A's to be neglected or overridden, because A believes that B's unwillingness derives neither (a) from B's calculation of his own interests, nor (b) from desires or character traits over which B has no control, but from a voluntary and settled commitment of will on B's part.

If we assume that when B says 'A trusts me' the kind of trust he has in mind is full-blooded trust, that A does in fact trust him in that way, and that A's trust is not misplaced, then clearly there is no problem about why B should offer that as a *reason* for not following C's suggestion. He is saying that he feels a *direct repugnance* towards doing anything which would constitute a betrayal of A's trust: a repugnance which has nothing to do, that is, with any intrinsic undesirability of the act in itself, or with any calculation of interest. And of course if he felt no direct repugnance of this kind towards breaking trust with A—if it seemed to him in general entirely immaterial whether or not an act could justifiably be represented as a breach of trust—then he would not *be* on a footing of full-blooded trust with A.

But perhaps we have made the nature of B's reasoning clear only at the cost of making the ultimate nature of his motives obscure. Given that full-blooded trust, as we have presented it, involves a commitment which by its nature cannot be justified in terms of the desires or interests of the person who makes it, why should B, or anyone else for that matter, ever be tempted to make such a commitment?

One thing that we might notice is that full-blooded trust between individuals, far from being the queer and recondite thing it might at first sight appear to be, is presupposed by a wide range of quite commonplace and familiar types of moral relationship. If we cannot full-bloodedly trust one another not to cheat at cards, then we are not related to one another as fellow-players, but as card-sharp and sucker, though for a time it may not

become clear which of us is which. If we cannot full-bloodedly trust one another not to wrong each other in pursuit of some personal advantage, then we are not related as friends, but as people who happen for the moment to have a use for one another. If we cannot full-bloodedly trust one another not to sell out a common cause or the national interest, then we are not related to one another as fellow-citizens or as co-workers for the same cause, but as opportunists or climbers on the same bandwaggon. When children see that they are at bottom not accepted unconditionally for what they are by their parents, but are loved only to the extent that they fulfil parental ambitions, then the relation between child and parent is apt to wear thin. And so on for every other kind of moral relationship: all require of us, at some point, an unconditional commitment to subordinate self-interest to the moral claims of the other party or parties to the relationship.

So one short answer to the puzzle of B's motivation might be this: he has made an absolute commitment not to cheat when A's back is turned, as a necessary condition of entering into a relationship of sporting fellowship with A, and he proposes to stick to it because he intends to remain in that relationship with A.

The trouble with this answer is that it makes it look as if, after all, B has an *interest* in not cheating; as if, that is, what primarily matters to him is not the injustice that would be done to A by his cheating, but his own concern to keep on friendly sporting terms with A. Worse still, it looks as if A's refusal to cheat yields him a *net benefit*, since *ex hypothesi* he wants to stay on terms of sporting friendship with A more than he wants to cheat.

However, it seems clear that to put matters in this way is, although unobviously, self-contradictory. If the upholding of A's rights in the game is not a primary concern of B's but is secondary to his concern to remain on terms of sporting friendship with A, then he is not, after all, really *on* terms of sporting friendship with A: he is just somebody who is prepared, for the sake of some ulterior purpose to act *as if* he were on terms of sporting friendship with A. Moreover it makes no difference whether the motive for which B wishes to feign concern for A's rights is simple self-interest (he wants to use A's friendship to advance his own career in professional sport, say) or a romantic infatuation with the idea of being A's friend. For no matter how keen B's desire to be on a footing of sporting friendship with A may be, if he feels no spontaneous indignation at the thought of A's being cheated out of his turn, but only, say, pleasure at being thus presented with an opportunity to behave as A's friend should behave, then he is not genuinely on any such footing with A and perhaps, unless he outgrows his combination of hero-worship and moral narcissism, never will be. For friendship, as for other kinds of moral relationship, to exist, the logical conditions of full-blooded trust must be met; and hence the process of 'making friends', whatever psychological account we give of it, has logically

to include the formation of commitments, together with the consequent patterns of experienced reluctance or concern, which cannot be justified in terms of interest or desire, precisely because such commitments are logically constitutive of the relationship.

Hume's distinction between natural and artificial virtues offers us a proto-Romantic vision of the moral life according to which the direct emotional commitments of the individual man or woman stand more or less opposed to the structure of formal rules and conventions inculcated by, and representing the interests of, 'society'. This proto-Romantic outlook in turn represents a shifting of the polarities of an earlier and still sharper Augustan distinction between morality as a socially imposed structure of rules on the one hand and individual self-interest on the other. Hume, along with Hutcheson, Butler and other writers of the mid-eighteenth century, softened this doctrine to allow for the possibility that some natural impulses may be morally virtuous. But in Hume in particular the hard lines of the older doctrine show through: moral rules represent the interest of society and are imposed by society on the individual by training or sanctions, while the moral impulses 'natural' to individuals have no intrinsic connection with moral rules or conventions at all, but represent simply the workings of an associative mechanism whereby 'all the affections readily pass from one person to another, and beget correspondent movements in every human creature'.[10]

If the foregoing arguments are on the right track, however, any such separation of moral rule from moral impulse involves a false dichotomy. Moral emotion and moral commitment on the one hand, moral rules and moral concepts on the other, are simply different aspects, inextricably because conceptually connected, of moral relationship. The interest we take in morality is not a matter of conditioned habits of response imposed upon each individual through social and parental training by 'society' conceived in Hobbesian style as a supra-personal individual possessing 'interests' which the commandments of morality articulate. Morality is social, but not on that account societal. It is what connects individual to individual, not what connects individuals taken one by one to a notional individual representing the abstract collectivity of the common life which is founded upon and made possible by moral relationship: the distinction is one which social theorists and social engineers neglect at their, and our, peril. Nor, on the other hand, is the interest we take in morality a matter of each bosom resonating in harmony with the winds of feeling which agitate its fellows. If that were the case it would be hard to see why morality should generate a conceptual scheme at all, or lend itself to the irritatingly fine distinctions, for example between heroism and heroics, friendship and flattery, sympathy and a sense of desert, which obstruct the free passage

[10] Op. cit. note 2, III, ii, 1.

of the more grandiose and schematic kinds of theorizing about morality. Morality is rooted more deeply in us than either social conditioning or nervous sensibility could root it, because its imperatives spring from the formal conditions for the existence of types of relationships into which individual human beings must enter with one another, because such relationships provide an essential framework around which the personalities and goals of individuals organize themselves.

To set in place this final keystone of the argument we must return briefly to the question of B's motivation. Let us assume B to be on a footing of casual but genuine sporting friendship and trust with A. If we have argued correctly so far, part of what it means to say that A and B are on that footing with one another is that each would feel some degree of moral outrage at the thought of anyone cheating the other out of his turn, and some degree of moral repugnance at the thought of doing so himself. A and B have become subject to such feelings as part of the process of acquiring habits of good sportsmanship, and it is because each knows that the other is subject to such feelings that they do trust and are at ease with one another in situations of sporting competition. B, in short, is 'inside' a system of moral relationships, and to his insider's view the repugnance he feels towards the act of usurping A's turn appears as a feature of the act;[11] it seems to him 'mean', 'a dirty trick' and so on. So of course the fact that it is A's turn is motivating for B, and is offered to C as a reason for B's reluctance to take practice shots on the assumption that C is also 'inside' the same system of moral relationships.

But perhaps the problem of B's motivation can be reconstituted at a higher level. Granted that B, by entering into relationships of mutual trust such as the one he enjoys with A, has fostered the growth in himself of patterns of emotional commitment and response which obstruct him in the pursuit of his private interests, why should he be prepared to allow himself to be ensnared by moral relationship in this way? Why should he not observe the outward forms of moral relationship, so far as these serve his private ends (which may include imposing and politically complex projects for the reform of society in general), while remaining inwardly free from the corresponding forms of emotional commitment?

The answer, I think, is that to be effective such a policy must be systematic, but that if it is pursued systematically it has the effect, not of liberating the agent to pursue his private goals, but of grossly impoverishing the range of things he can intelligibly posit as goals: the effect, in short, of draining most kinds of *meaning* from his life.

It seems clear that the conditions for moral relationship can make no reference to particular, named persons. Particular persons may become intrinsic sources of meaning and goals of activity for me, as Gilberte Swann

[11] Cf. David Wiggins, op. cit. note 5, 347.

does for the young Marcel, without that necessarily placing me on a footing of moral relationship with those individuals. Indeed, as we have already seen, infatuation with a particular individual may actually inhibit friendship with that individual: if B's repugnance at the thought of usurping A's turn arises solely from the fact that the turn belongs to A, whom he hero-worships, and if A is well aware of this, then A will not trust B, or be on a footing of friendship with him, though he may in an amused way take it for granted that B will not double-cross him so long as his bout of hero-worship lasts. If B and A are to stand to one another on a footing of casual sporting friendship, in other words, B's commitment to fair play cannot be *ad hominem*: it must be commitment to a general *schema* of conduct; to kinds of moral reluctance which have nothing to do with any warmth of affection or admiration he might also, as it happens, feel towards A personally. It is because B's propensity to feel those kinds of reluctance fits him for sporting friendship with *anybody at all* that, paradoxically enough, he is able to be on a footing of friendship, and not just hero-worship or infatuation, with particular individuals, A included.

Hence, if B wishes to deaden the feelings of emotional repugnance which prevent him from dealing self-interestedly with A, he cannot do so *selectively*. It is no good his silencing his repugnance towards one particular act of betrayal of trust if he leaves alive in himself the kinds of repugnance towards cheating and betrayal of trust in general which fit him to participate in the *type* of relationship in which he stands at present to A; for then he leaves open the possibility that these generalized repugnancies will appear in forms—that of subsequent remorse, for instance—as paralysing to his pursuit of his private interests as the specific feelings of repugnance towards cheating A that he has succeeded in repressing. What he needs if he is to emerge into the radical freedom proposed by the philosophical egoist is not just the power to repress his moral feelings for a time or in particular cases, but a radical cautery of such feelings. Whether such a radical cautery is psychologically possible may be doubted, but it will only be logically possible on condition that B withdraws not just from one or two relationships, but from all relationships whose specification proceeds partly in terms of full-blooded trust.

If B takes this radical step, however, then at the very moment when he recovers radical freedom to pursue his private ends, the range of things which he can intelligibly posit as ends suffers a correlative and equally radical diminution. It will now be of no intrinsic concern to B whether another person's justified convention-backed claims are ignored or overridden. Thus he no longer satisfies a condition which he is logically required to satisfy if he is to stand to others in any of the wide range of moral relationships which include, for example, those of friend, co-worker, fellow-sportsman, fellow-citizen, parent, spouse, and, less formally, the casual relationships of implicit trust which complete strangers strike up with one

another in passing many times a day. For non-egoists such relationships are, directly or indirectly, the source of most of the goals and interests which attach people to life: they underlie the desire to bear children and to see them flourish, to organize works outings to Southend, to see teams win, to enjoy casual companionships in pubs, to free nations from servitude to foreign powers, to write novels that will be read, to undertake heroic feats of competitive angling, and so on. All this fabric of commonplace involvement in life is closed to the radical egoist because he has placed himself outside it: he is no longer *in* any of the kinds of relationship which give sense and meaning to such interests. So far as he keeps up an appearance of involvement in such relationships, that practice must have as its object some set of goals which lie beyond the charade of seeming trust and seeming concern: it cannot be *the marriage* and its shared goals, *the friendship* and its companionable pleasures which interest him; it has to be *something else*. But what can this *something else* be? The problem here is essentially that of ascribing intelligible motives to Iago.[12] If friendship, family, personal affection, group loyalty, sporting companionship, the general solidarity of human beings in distress, are all painted shams, they can no more engage interest or give meaning to life than paintings of shop-windows, once a *trompe l'oeil* illusion has dissolved, can draw the window-shopper. B in his radical egoist phase must define his goals entirely outside the circle of intrinsically motivating concerns created by moral relationship: but what is left outside that circle? Certainly the basic appetites: food, drink, exercise, warmth, and sex provided it is not allowed to dethrone calculation. Beyond that there remain mockery, revenge, the enjoyments of arbitrary power and, curiously enough, certain kinds of meaning-creating relationships to others which escape the circle of moral relationship, and which include power-worship, the devotion of a willing slave to his master and the kind of infatuation which dogs the footsteps of Proust's characters; though these last the egoist, who desires freedom to pursue 'private interests', must presumably avoid like the plague. Beyond that there is curiosity and the pursuit of knowledge so far as these can be detached from the pleasures of discussion and shared activity with co-workers; and that concludes the list.

It may well be psychologically possible to construct a kind of human life out of these meagre materials, and in any case we can afford to remain agnostic on that issue. The serious difficulty which the philosophical egoist must confront is the conceptual one that there is, logically speaking, no halfway house between the extreme motivational poverty of such a life and some degree of commitment, of a more than merely formal and outward kind, to moral relationship. Radical egoism 'liberates' us to pursue

---

[12] Cf. Gabriel Josipovici, *Writing and the Body* (Princeton: Princeton University Press, 1982), 47–48.

our interests only at the cost of demolishing the fabric of interests which alone could make such 'liberation' intelligible.

In short there is a real price to be paid, in freedom to pursue private interests, for entry to the moral life. Some philosophers, like Butler, have tried to argue that this entrance-fee is simply an investment attracting dividends payable in the same currency: that the moral man will find it possible to *satisfy* more of his interests than the man who rejects morality; but many obvious counter-instances undermine this claim. So far as there is any 'reason for being moral' which might weigh with someone contemplating adopting the stance of radical egoism, it is the altogether bleaker and more negative one which we have just uncovered; that we have no choice, if we wish merely to *have* the interests we have, never mind satisfying them. The dividends payable on our investment in the moral life appear, not in the form of increased, or deferred, *interest-satisfaction*, but in the possibility of life taking on certain kinds of significance which it could not otherwise possess. It is perhaps this, or something like it, that Wittgenstein had in mind when he spoke of good or bad acts of will as altering 'the boundaries of the world', so that 'The world must, so to speak, wax or wane as a whole. As if by accession or loss of meaning.'[13]

V

Perhaps, it might be suggested, we have a (second-order) interest in having the (first-order) interests that we have. But this way of putting things will not allow us to say that being moral is a matter of pursuing my private interests after all, for allowing myself to be drawn into moral relationship with others logically involves the acquisition of general habits of moral response capable in principle of conflicting with and overriding any specific interest of mine. Nor would it be any less misleading to characterize moral scruples as desires. Very often we do not at all desire to say the words or do the things that justice, or common decency, or the duties of doctor or teacher extort from us. It seems, in fact, that desire and interest, despite Philippa Foot's persuasive arguments to the contrary,[14] do not exhaust between them the list of ultimate grounds for action. It is evidently not in B's interests to refrain from taking practice shots; but equally what leads him to neglect his interests by refraining from taking any is not a desire (compare the case in which B neglects to cheat only because he is occupied in placing a bet on the 4.30 at Kempton Park) but a moral

[13] Ludwig Wittgenstein, *Notebooks 1914–1916*, G. H. von Wright and G. E. M. Anscombe (eds) (Oxford: Basil Blackwell, 1969), 73e.

[14] Philippa Foot, 'Reasons for Actions and Desires', *Aristotelian Society Suppl. Vol.* 46 (1972).

repugnance, which is something different from either an interest or a desire because it issues directly from the foundation of interwoven convention and commitment of will and feeling upon which the structures of interest and desire which characterize human personality are in part erected.

It is not surprising, therefore, that moral concern or repugnance can conflict with, and at times override, considerations of interest and desire. We rightly feel that moral concern does not represent just one more interest or desire that we happen to have, and hence that it is not barterable against interests or desires in the way that one interest or desire is barterable, if we wish, against another. We feel, in short, that moral imperatives are categorical. And there is a sense in which we are right about this, too. Moral relationships clearly differ in emotional intensity, from the more profound kinds of personal or group commitment, to the casual intercourse of ordinary social morality. But it is a logical requirement for the existence of any moral relationship, even the most casual, that the participants should not respect the demands of the relationship merely because doing so happens to coincide for the moment with their interests or desires. Of course, that does not mean that they have to grant the demands of the relationship *unquestioned* or *automatic* priority over interest and desire. Some moral demands are intrinsically less weighty than some prudential ones; and Philippa Foot is clearly right to argue[15] that moral reasons are not always regarded, even by very 'moral' people, as automatically overriding prudential ones. I think it could be argued, however, that in distinguishing between categorical and hypothetical imperatives Kant is, as is usual with him, less concerned about how people (even very moral people) behave, than about what morality is. His concern in drawing the distinction is to isolate the manner in which moral claims, whether we heed them or not, bear on us; and the thesis enshrined in the distinction is simply that they bear on us in a manner which makes no reference to any of the specific interests or desires which we may happen to have as individuals.

So we are left, perhaps, with a kind of Kantianism, but one which is not incompatible, for instance, with Bernard Williams's[16] putatively anti-Kantian thesis to the effect that what we take to be the emotional basis of an action is not irrelevant to the question of what moral judgment we are to pass both on the act itself and the actor.

On the other hand, the position I have been outlining here is not, or at least is not supposed to be, a rationalist or 'objectivist' one; though it is supposed to be in some sense a Realist one.

To begin with, it evades the fact-value distinction, and all its non-

---

[15] Philippa Foot, 'Are Moral Considerations Overriding?', *Virtues and Vices* (Oxford: Basil Blackwell, 1978), 181–188.

[16] Bernard Williams, 'Morality and the Emotions', *Problems of the Self* (Cambridge: Cambridge University Press, 1973), 207–229.

cognitivist consequences, by the obvious strategy, remarked on by my namesake Jonathan Harrison in his *Hume's Moral Epistemology*,[17] of appealing to a 'desire to do what is right' (although it would be less misleading to speak of a repugnance towards doing wrong) which accounts for our sometimes being moved to action merely by the knowledge of what is, morally speaking, the case, without the need to ascribe 'intrinsic action-guiding force' to moral judgments themselves. It gives a new twist to this hoary tactic of moral rationalism by the manner in which it explicates the relationship between moral judgment and moral action. The account of the relationship propounded by Hume and by modern non-cognitivists goes something like this: moral judgments are intrinsically action-guiding because such judgments serve the logical function, not of stating facts but of giving verbal expression to some mental state or act (a sentiment, a choice, or an attitude, for instance) whose corresponding practical expression is an action. What I have done is to move the point of union between the practical and the theoretical several steps further back. On the one hand we are subject to moral impulses and emotions of various kinds—repugnance, remorse, righteous indignation, moral concern, and so forth—because we build the systems of interest and desire which define us as individuals around modes of relationship to one another which (a) depend on trust in the full-blooded sense, and which (b) because of the dependence of the notion of trust on a notion of *wronging someone* which has to be defined in terms of interest-equalizing conventions, ultimately go to ground in systems of convention of that kind. On the other hand we make factual assertions, about whose turn it is, who owes whom what, what X ought to do in given circumstances, who trusts whom, and so on, couched in a vocabulary of conceptual distinctions which ultimately go to ground in the very same conventions. The statements in question thus interest us practically because the conceptual vocabulary in which they are couched, while entirely 'descriptive', is not just any old conceptual vocabulary, but the conceptual vocabulary *of the moral life*. The relation between the theoretical and the practical is thus not one of *inference*, and so raises no problems about the logical relationship between descriptive assertions and 'ought'-statements. A very large class of moral 'ought'-statements *are* descriptive assertions of a kind: they describe how someone who intends to participate straightforwardly in a given interest-equalizing convention must order his conduct if he is to carry out that intention. As Julius Kovesi has argued on other, though related grounds, 'moral notions do not evaluate the world of description; we evaluate that world by the help of descriptive notions'.[18]

---

[17] Jonathan Harrison, *Hume's Moral Epistemology* (Oxford: Clarendon Press, 1976), 13–15.
[18] Op. cit. note 8, 161.

We are thus left with a picture of the conceptual scheme of morals as a structure containing several quite distinct layers of conceptual organization; most notably a level containing practices such as turn-taking or promising which lend themselves to essentially game-theoretic modes of description, and a level containing a range of concepts concerned with moral relationship (*friendship, citizenship, common humanity*), which do not; the linkage between the two levels being provided by the (non-game-theoretic) concept of *trust* and the (game-theoretic) concept of *wronging someone*. This differentiation of conceptual levels has some interesting consequences.

First of all it offers us some of the advantages of the kind of moral realism advocated lately by Mark Platts[19] and others. Because, on the present view, the truth and falsity of moral judgments is established by appeal to the terms of conventions rather than by appeal to effective procedures for determining truth and falsity for particular judgments, we can both offer a generally realistic account of truth and falsity for moral judgment and explain how a speaker's grasp on a moral concept can transcend his recognitional capacities. Thus, for example, it is (or is not) true that it is A's turn *given the terms of the turn-taking convention*, whether or not we happen to have lost hold (because, say, we have forgotten the sequence of turns, or who began) on the possibility of effectively deciding that.[20] Again, the concept of *wronging someone*—or of *someone's having been wronged*—is not tied to any particular set of procedures for establishing that wrong has or has not been done, because it is not tied to any given set of interest-equalizing conventions: new ones can arise.

On the other hand we can avoid some of the least plausible features of recent moral realism. We are not committed, for instance, to speaking of 'moral reality' (in a way rightly castigated by Mackie[21] as viciously 'objectivist') as if the morally important features of a situation simply presented themselves to our recognition in the way that the colour of an emerald and the sound of a waterfall do. We can allow with modern Humeans like Mackie or Gilbert Harman[22] for the role of convention in partly constituting moral reality, and that in turn permits us to do justice to some of

---

[19] Mark Platts, 'Moral Reality and the End of Desire', *Reference, Truth and Reality*, Mark Platts (ed.) (London: Routledge and Kegan Paul, 1980), 69–82.

[20] For a similarly Realist account, along corresponding lines, of truth and falsity for sentences involving proper names, see my 'Description and Identification', *Mind* **XCI**, No. 363 (July 1982), 321–338.

[21] J. L. Mackie, *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin Books, 1977), 15–49.

[22] Gilbert Harman, *The Nature of Morality* (New York: Oxford University Press, 1977), 91–114.

the more legitimate claims of moral relativism. We can envisage the possibility that different societies might give rise to different kinds of interest-equalizing convention, and different kinds of resulting moral relationship, from ours; and that our kind of moral intuition might not be capable of judging, at any rate in a hasty and *prima facie* way, the resulting practices. But at the same time we are not thereby committed to the view that different systems of moral practices are hermetically sealed off from one another so far as moral judgment or the transmission of values are concerned. Values leak across the boundaries of social groupings because moral conventions, and the modes of moral relationship and moral feeling which go with them, are socially contagious in the way that all human practices—the making of pottery or the use of the plough, for instance—are socially contagious. 'Social conditioning', the inculcation of unquestioning respect for socially preferred standards of conduct, may hinder the resulting leakages of values, but in the end can neither explain nor prevent them.

Finally, although on the present view moral reality is constituted in part through the adoption by human communities of conventions whose object is to equalize advantages, it does not follow that morality is susceptible, above a certain level, to any form of analysis which ascribes an ideal egoism to participants, or which reconstructs moral relationship wholly in terms of the quantitative bartering of interests against one another. Utilitarianism, and other inherently quantitative modes of moral theorizing, are often opposed by an entirely *ad hoc* appeal to 'moral intuition'. It seems to me possible that what we call 'moral intuition' involves a confused perception of the fact that not all our moral concepts and judgments are on the same logical level: that the conceptual scheme of morals has at least the kinds of depth and complexity which I have tried to exhibit in the foregoing pages. The trouble with quantitative models for the analysis of moral decision-making may not be, in fact, that they are irrelevant to 'morality' *tout court*, as revealed *ad hoc* to 'moral intuition', but merely that they are applicable to some levels of our moral thinking and not to others.[23]

University of Sussex

[23] In its progress through various versions this paper has profited from criticism by Professor D. W. Hamlyn, Ruby Meager, Professor Kurt Baier and Richard Rosenbluth among other contributors-to-discussion, and by Professor J. J. C. Smart, who read and commented on an earlier draft of the present version.