Hare's *Moral Thinking*

A. PHILLIPS GRIFFITHS

This book[1] is the fruit of well over thirty years' reflection on moral philosophy. A complete appreciation of it requires reference to *The Language of Morals* (Oxford University Press, 1952) (*LM*), *Freedom and Reason* (Oxford University Press, 1962) (*FR*) and many of Hare's extensive shorter writings. To some, it will appear to represent a radical, if gradual, reversal of his early views. His early position was thought by some to be one similar, in certain respects, to that of certain existentialist thinkers: that the most fundamental moral attitudes must be the outcome of sheer choice, or commitment. Whereas in the latest book Hare argues that to think fully rationally about moral questions requires us to be Utilitarians (unless we are *amoralists*, as we shall see below).

This view of his early work might most easily arise from a reading of Chapter 4, 'Decisions of Principle', in *LM*. Moral judgments 'can only be verified by reference to a standard or set of principles which we have by our own decision accepted and made our own' (*LM*, p. 78). At the most fundamental level, he said, we can only ask a man to make up his mind about how he ought to live, '. . . for in the end everything rests on . . . a decision of principle' (*LM*, p. 67). This impression is reinforced when, modifying Kant's dictum slightly, he says 'to ask whether I ought to do A in these circumstances is to ask whether or not *I will* that doing A in such circumstances should become a universal law' (my italics). This makes it look as if, at the fundamental level, the only injunction can be 'legislate what thou wilt, for that is the whole of the law of laws'.

Nevertheless, what Hare said in *LM* can be read, without the preconceptions introduced by the comparison with existentialism, in a way which leaves it quite compatible with his present views. Kant's dictum, unmodified, is about what a man *can* rationally will (again, my italics). For Kant, there are rational constraints on what a man can will shall be a universal law. One such is suggested on p. 69 of *LM*: a man must *try* to live that life he accepts as the one he ought to live. Since one obviously cannot rationally accept a way of life one cannot live, this is a constraint on what one can accept. Hare's later thinking may be represented as a progressive investigation of what such constraints there are: progressive, because step by step they become more severe. In *FR* the utilitarian, who treats all

[1] R. M. Hare, *Moral Thinking: Its Levels, Method, and Point* (Oxford: The Clarendon Press, 1981), viii and 242 pp., £11.00, £3.95 paper.

preferences as having an authority proportionate only to their strength, cannot 'rationally constrain' someone into abandoning a position in which some preference is treated as overriding with respect to all others whatever their number or strength (and whether preferences of the agent or someone else). The latter kind of person Hare calls a 'fanatic' and in the present work argues that even this leeway must be closed. 'If we assumed a perfect command of logic and the facts, they would constrain so severely the moral evaluations we make that in practice we would be bound all to agree to the same ones'.

The constraints, then, are of two kinds: logic, and the facts. By logic Hare means not purely formal logic, but 'philosophical logic': in this case, about the 'formal, logical properties of moral words'. This yields 'a system of moral reasoning whose conclusions have a content identical with a certain sort of utilitarianism' (p. 4). I shall not enter far here into the question of whether philosophical logic is relevant; it seems to me that Hare is surely right that if we are going to ask how moral questions are to be answered, we must examine what the questions are, that is, what is meant by the sentences we are using to ask them. It could be objected that a prior question is, how do we determine the class of moral questions? Hare's answer is that while he does not abandon the claim that he has a sound idea of what determines the class of moral questions, he does not rely on it. The discourse he takes to be moral discourse is in a language in which we can reason about practical questions, and will (if his subsequent arguments are correct) produce rationally acceptable answers for those who use it seriously.

While no facts can be ruled out as relevant absolutely *a priori*, it is the task of moral thinking to determine which among possible preferences are relevant. We can do this by a consideration of the formal, logical properties of the moral words alone, which is why such considerations can yield a utilitarian system of moral thinking. To use the utilitarian system of moral thinking to reach moral conclusions in the form of substantial moral judgments will require the empirical task of determining which among the class of possible relevant facts are actual.

The class of possible relevant facts turns out to be the class of facts concerning people's possible preferences. The reason that it is *all people's* preferences (or, more strictly, all conscious beings' preferences) which are relevant is that moral judgments are universalizable: that is to say, it is a contradiction to assert two differing moral judgments about cases supposed to be identical in their universal descriptive properties. Hare's logical thesis concerning universalizability has been exhaustively discussed in his earlier works and by others, and in what follows I shall discuss Hare's views on the assumption that it is correct. Moreover, I think it is correct: in so far as we bear in mind the caveat that even if we are wrong about the import of the word 'moral' or the logic of the term 'ought' we are setting

out a possible way of thinking. We will not be subject to a criticism such as Kolakowski's (*Religion*, Oxford University Press, 1982, 190) of Kant, that 'I am not inconsistent at all if I prefer other people to follow rules which I do not want to abide by', for the claim is not being made either that I *must* use the term 'moral' in such a way that this must be inconsistent, or that if I am using the term 'ought' with the suggested logic there *must* be practical consideration of what I ought to do over and above considering what I want. (See Hare on the rationality of the amoralist, 10.7–8.)

The reason that people's *preferences* are relevant is that moral judgments are prescriptive: that is to say, they must always entail at least one imperative. What this involves, for the purposes of the argument, is that being subject to the requirement of universalizability any moral judgment must entail a first-person imperative; e.g. 'You ought to pay your debt' addressed to another requires that I assent to 'when in exactly his circumstances I ought to pay my debt', and this entails 'Let me pay my debt'. I do not, however, accept 'Let me do x' unless I do x, or at least am motivated to do x. To be motivated to do x is to have a preference for doing x. Hence, sincerely accepting a prescriptive moral judgment is to have a preference, and all prescriptions express preferences.

Assuming a perfect command of logic and the facts, Hare says, we would all agree. Relying on nothing but the logic of the moral words and the facts, we are engaged in what Hare calls critical thinking. It is contrasted with intuitive thinking, when we rely on more or less general principles about what we ought to do.

We have to rely on intuitive thinking most of the time: because in urgent practical situations we rarely have the time to engage in the difficult process of critical thinking, whereas intuitive principles can be regarded as the outcome of critical thinking engaged in before the event and with precisely these difficulties in mind; because we cannot rely on our individual knowledge of the facts, especially facts about people's preferences, whereas intuitive principles are ones about which there tends to be a (local) consensus and which are therefore more likely to reflect common preferences; and because the difficulty of honest and careful thinking lays it open to the abuse that we misuse it to put aside those intuitive principles which conflict with our own individual preferences. Thus critical thinking itself demands that for most of the time we form our moral conclusions through intuitive thinking.

On the other hand, we are forced to engage in critical thinking, for two reasons which reflect facts about our intuitive principles. First, they notoriously can conflict (principles of benevolence versus justice, welfare versus liberty, etc.). This may lead to the modification and increasing specification of intuitive principles; but this process cannot lead to useful principles with very little generality and a point must come where the case is considered on its merits from a critical point of view. Second, different

people may hold, or one person may hold, incompatible intuitive principles, and such a disagreement where one or the other intuitive principle has to be abandoned can only be settled at the critical level. These points are reminiscent of Mill's arguments for the need for a supreme principle at the beginning of *Utilitarianism*; and indeed Hare points out that the distinction has frequently been adumbrated in the history of moral philosophy; and he regards it as of the first importance. The first section of the book is mainly devoted to distinguishing the two levels, and the two succeeding sections are on 'method'—how rational critical thinking about moral issues must proceed; and on 'point'—why we should engage in critical thinking. Hare uses the distinction to show how objections to utilitarianism very often arise from a failure to appreciate it. For example, one might see that while a calculation of utilities would justify a policeman's torturing a prisoner in some cases, that would conflict with our intuitive principle that no policeman should ever resort to torture. Thus utilitarianism is incompatible with our intuitive morality, and thus stands condemned by it. Hare's discussion of such examples is lively and instructive, but it is enough here to give his general answer. If utilitarianism (the product of critical thinking) gives us a different answer from the accepted code then it is the intuitive not the critical conclusion which stands condemned, because it is the function of critical thinking to judge the acceptability of intuitions. However, it does usually turn out that the cases which are supposed to involve counter-intuitive utilitarian conclusions are merely logically possible and never actual. Hare turns the critic's ammunition against himself. In the light of our intuitions utilitarianism, it is claimed, seems at least silly. Hare ripostes that it is only by producing silly, unreal examples that it can be made to look as if critical thinking conflicts with general principles intended only to cover cases which are likely to happen in reality. We need not enter into a discussion of Hare's particular examples here, except to make Hare's quite firm position quite clear. Bentham (Twining and Twining, 'Bentham on Torture', *Northern Ireland Legal Quarterly* **24**, No. 3. 1973) thought there very well could be actual circumstances in which it would be right and indeed a duty for a policeman to inflict torture, and that we ought to bring about the conditions in which such circumstances could occur. These would be where a properly constituted court finds with due process and the strictest judicial safeguards that a prisoner is a part of a criminal conspiracy and has knowledge of accomplices who are still at large and pursuing their secret and nefarious plans. It could then instruct a policeman to put the criminal to the question, using a degree of severity laid down by the court, and certainly no more than would be efficacious. Bentham points out that such instructions by a court are less undesirable than those given to a hangman or a jailer, since the disutility of the order can be minimized or avoided by the prisoner's willingness to give the information. Hare's response here would be either to deny

Bentham's factual claims or to accept his conclusion. As he says on p. 167, if the facts were other than be believes they are, then the method of argument he advocates 'might make me advocate slavery and tyranny'.

Hare argues that the demands of universalizability and prescriptivity rule out the possibility of a self-regarding or egoistic or selfish morality on the critical level. He would not, I think, deny that if the facts were very different indeed, a purely egoistic morality would be required at the intuitive level. But this would not be to employ an egoistic principle at the critical level: on the contrary, it would result from the critical demand that a policy of every man for himself be accepted only because it was to the benefit of all rather than of one particular person. Hare's view is, however, that the world is such that the contrary is really true: far from its being that rational principles require practical policies which serve self-interest, the rational pursuit of self-interest requires practical policies of concern for the interests of others; that is, it is prudent to be benevolent in general (11.5–6).

It may not seem immediately obvious why the critical level of moral thinking could not rest on a universal prescriptive principle which constitutes a universal egoistic morality. In trying to explain why this is so I shall not be expounding any particular argument of Hare's, but relying on what I have (perhaps mistakenly) gleaned from the present book and *FR*. The parenthesis is important because I shall try to suggest some arguments against Hare's major conclusion which rely on the way I put things here.

The principle of egoism (*E*) is a prescriptive one which could be expressed: *If I have a preference for some state of affairs S, and my doing x is a necessary condition of S, then,* ceteris paribus (cp), *let me do x*. The content of the *cp* clause, for an egoist, could concern only his other preferences, whatever they might be, that is only as being his preferences *as such*; which means that only their strength (and possibly their alterability) could enter into consideration. This is a rational principle, and one which rationality demands should be satisfied if one accepts it as subject as a whole to a further *ceteris paribus* clause, the content of which we can for the present leave open. This is because if prescriptions are the expressions of preferences, then one cannot have a preference without accepting the prescription. (It would seem that the primary expressions of preferences or desires are optatives, but these can be regarded as having a logical relation to imperatives, illustrated as follows. If someone being unemployed says 'Oh, if only I had a job, any job' and is then offered a job and refuses it, then his not accepting the imperative 'Let me take the job' throws doubt on the sincerity of his optative utterance. That is why I expressed the consequent of the egoist's hypothetical principle in terms of *doing x* rather than in terms of *S*. There will, of course, be some optatives which because of step-motherly fortune can never in practice give rise to imperatives; a lot of people wish they were in employment but can't do anything about it.)

The consequent of *E* is a particular imperative and remains so when we eliminate the particularity of the antecedent to produce a principle of universal egoism (*UE*): *If some agent A has a preference for some state of affairs S, and doing x is a necessary condition of S then, cp, let A do x.* Now while this seems both prescriptive and universal, it seems to involve a kind of pragmatic prescriptive contradiction: that whoever prescribes the universal principle (*UE*) prescribes that he shall make only particular prescriptions, given that the *cp* clause ranges only over the agent A's preferences. Or, another way of putting it, it requires every thinker to accept only particular imperatives, so since it is universal no thinker who accepts it can accept it. So it cannot be a principle for a critical thinker. If it were regarded as a possible principle at the intuitive level, then it would be *prima facie*, and hence governed by an external *cp* clause, and the content of *that* would have to be determined by a critical principle which could not be *UE*.

E is none the less a rational principle, and there is no rational necessity for anyone to adopt another. One who does not, however, will not prescribe in accordance with any universal principles; that is, not seriously engage in moral discourse: he will be what Hare calls an amoralist. It is important for Hare that this is so: reason cannot force us to accept universal prescriptive principles; rather, in doing so we 'choose in freedom' (Sartre): the quasi-existentialist 'decision' thus vanishing to a not quite extensionless point. The great difference is that the existentialist's anguished plumpings are all at the ground level of specific individual choices (with the possible exception of choosing *millefeuilles*) so principles are useless: whereas Hare's choice is at the highest level. Hare's view of the amoralist is like Schopenhauer's contempt of the solipsist: he lives in an impregnable tower but can be simply ignored (by the theorist, if not the policeman) since he cannot sally forth to attack the moralist; and, Hare would add, the fact is that he will find the interior of the tower so uncomfortable that he will be forced to come out and surrender (11.2–4). Leaving the amoralist in his tower, then, we ask how the erstwhile egoist can accept the external *cp* clause, how he can test his particular imperatives by requiring that they are compatible with universal imperatives *that he accepts* (one might even say, legislates: so that he remains a law unto himself).

It is part of rationality for everyone that he should prescribe that, *cp*, if he has a preference for S and his doing x is a condition of S, then *cp* he should do x. The second *cp* clause exclusively concerns the agent's other preferences, whereas the content of the first (external) *cp* clause is left open. The latter thus leaves it open that the agent can consider other matters than his own preferences, and this of course leaves room for him to consider the preferences of others. The restricted principle governed by the external *cp* clause does not refer to the *content* of the agent's preferences, but only to their *being preferences*, which means that only charac-

terizations of preferences *as such* can be relevant, i.e. of their strength or alterability. If the restricted principle is to be seen as following from a universal principle, then at least the limitation constituted by the reference to the agent must be removed. This minimal requirement would be satisfied by abstracting the possibility of considering others' preferences from the external *cp* clause, and building it into the restricted principle which would then be universal, thus (*L*): Ceteris paribus, *if anyone has a preference for S, and doing x is a condition of S, then,* cp, *x ought to be done.* The internal *cp* clause of *L* will concern only the strength and alterability of preferences of anyone. This new unrestricted and universal principle *L* is got by a consideration of no more than prescriptivity (which gives the restricted principle) and universalizability. If only the strength and alterability of preferences are relevant in the application of *K*, then the first *cp* clause becomes redundant, being empty. In that case *L* is prescriptive and universal, but also *overriding*, since if the external *cp* clause is empty there could never be any counter-considerations with regard to what follows it (that is, there is no principle with which it can conflict). If, however, the *cp* clause can range over other matters—such as the content of preferences in determining what to prescribe (i.e. the nature of what the preference is *for*, its intentional object)—then it ceases to be overriding and becomes *prima facie*. In Hare's system that means it is not a principle of the critical level, because it could conflict with other principles, whereas the principle of the critical level must be sovereign and overriding in order that such conflicts can be resolved. Those who disagree with Hare must either claim that there is some other over-arching principle, other than the utilitarian one, at the critical level, or accept that as moral thinkers we simply have to put up with fundamental *prima facie* principles. The second position is clearly uncomfortable, but only cripplingly so if it could be shown that such fundamental principles would not deserve to be called moral at all. I do not think that, as far as the criteria of universalizability and prescriptivity are concerned, this need be the case. If overridingness is a condition of being a moral principle, then this may seem to make irresolubly *prima facie* principles, which can always be overridden by another, non-moral. But surely we have the alternative of treating the class of moral judgments as overriding in relation to those of other classes (aesthetic, prudential, medical, strategical, etc.), and *prima facie* moral principles as those necessary to moral judgments.

The issue is, then, whether it is ever possible to introduce considerations which do not concern the strength or alterability of preferences, at any remove, in moral thinking. First, I wish to state an objection on a subsidiary point which, I think, does not affect the central issue.

At one point in his account of what must be relevant considerations (5.2, p. 90), Hare seems to narrow these not just to the strength of preferences, but to the strength of preferences for experiences. Now, even taking

the term 'experience' to have the widest possible extension (to include knowledge or awareness of any state of affairs) I can see no justification for such a restriction. As Hare uses the term, I cannot have a preference without being conscious of it or being able to be conscious of it. That does not imply that its content or intentional object need involve reference to anything of which anyone is or might be conscious, and as Hare uses the term an experience is something of which its subject necessarily is conscious. To insist on this restriction would, it seems to me, introduce a factor concerning the content of preferences (that they should be for experiences of some sort) where only their strength and alterability can be given weight. It is the latter requirement, and the powerful moves Hare can make in support of it, that I want to question.

The requirement comes out most clearly I think in Hare's discussion of the 'pure' fanatic. For a particular reason I want to consider this in terms of what will be from a common sense point of view a totally unreal and fantastic case. Hare is willing to do this at the critical level, for unreality comes in as a criticism only when we are considering intuitive principles tailored for the world that common sense inhabits. Imagine that someone, called Moore, holds that any world, whether devoid of observers or not, which is more beautiful than another, is thereby more preferable. Let us suppose he also holds, what his namesake G. E. Moore did not, that there being such a world is better than there being one which while ugly satisfies more preferences of presently living beings. He and all other people, who are few enough to join in a debate, firmly believe that in a year's time all possible observers are to be destroyed by the arrival of a massive asteroid, and with them all their artefacts unless they are buried at the bottom of the deepest mine in existence. All agree that the Elgin Marbles are the most beautiful artefacts in existence and that they could be preserved indefinitely if everyone's efforts for the remaining year were bent to that aim. Everyone else advocates a year of eating, drinking and merrymaking (things for which Moore also has a taste) but Moore urges them to forgo this for the sake of the Elgin Marbles, which, he claims, is what they *ought* to do. Now common sense suggests that Moore's proposal would be rejected out of hand: Moore's preferences are absolutely dotty, for what is the point of preserving works of art nobody can look at? But this is not a move Hare would make at the critical level; that Moore's preferences are dotty, or immoral, would be a judgment at the intuitive level. For, says Hare, no preference has any greater authority or dignity than any other at the critical level (10.6, p. 179). Taking Hare's advice, however, the majority could argue with Moore as follows. You agree, accepting *L*, that everyone's preferences ought to be taken into account. To this you oppose another consideration, the preservation of the Elgin Marbles. If that is a genuine prescription arising from that consideration, then it is one of your preferences. It may, extraordinarily, indeed, be your strongest preference,

outweighing all of your others. But reflect that, if you were any one of us, grasping that your preferences would then be the same as ours *you* would prefer that *those* preferences were satisfied. So surely your preferences are that were you to be each of us simultaneously the preferences to be satisfied should be all of ours rather than those you happen to have where you are not considering yourself as all of us simultaneously (or as successively each of us in random order).

If Moore does so reflect, but then responds 'Ah, but I am not you, so I can stick to my prescriptions, the hypothetical condition of my being you not being satisfied', he could be answered 'But this is only to say you are one distinct individual amongst others, and what difference does that make in the application of any universal principle? Or that *your* preference ought to be satisfied just because it is *yours*? But for everyone, his own preferences are his own: so again, what difference does that make? It seems you are just refusing to accept universal principles; you are not what Hare called in *FR* a pure fanatic, you are simply an amoralist. Leave the debate.' But Moore might say that he does claim to be a pure fanatic, because having reflected on the matter as asked, he finds that whether he considers the hypothetical condition or not, his preferences come out the same, because his preference is not only stronger than all his other preferences put together, but also stronger than everyone else's put together, so he must prescribe the preservation of the Elgin Marbles. If the others believe Moore about the incredible strength of his preferences, then their only recourse is to agree that they ought to preserve the Elgin Marbles; unless they agree to being amoralists and leave the debate, remarking that Moore's claim to fanaticism was mistaken and that he is the only true utilitarian among them (hence Hare's conclusion that there is no logical space for the 'pure' fanatic between the impure fanatic and the utilitarian.)

Before giving up, however, the majority might try a further ploy suggested by Hare. Moore's preferences are alterable, they might claim, whereas theirs are not. I can't see that this is a strong ploy, however. I take it that this move seems legitimate for two reasons. The consideration concerning alterability can be introduced since like the factor of strength one it does not involve anything over and above what pertains to preferences as such. Furthermore, if at the critical level one is prescribing that each person's preferences, whatever they are, be satisfied, one must prescribe that any conditions under which this is possible be satisfied. It cannot be satisfied where people's preferences are incompatible, so one should prescribe that they be made compatible. This will require altering at least one preference, so where of two preferences one is alterable and the other not, the alterable preference should be altered. Thus it might be pointed out to Moore that after a bit of aversion therapy, of the kind applied to potential soldiers in *Brave New World*, he would lose his preference and become

like everyone else. While Moore might not deny this, he could point out that he has had to school himself to endure pain, fear, hunger and cold, loneliness and derision in his fight to preserve the Elgin Marbles, and if they respect his overriding preference, so much stronger than the totality of others, they can and should school themselves to the hardships of the straight and narrow path. One cannot in this matter appeal to the fact that it is undoubtedly true that the vast majority of people suffering pain cannot *but* prefer that they were not: the question is whether they can alter the *strength* of that preference against another. Hare does say that there are autonomous, unalterable preferences, but not what they are or how they may be determined. Reading of the dedicated behaviour of Japanese soldiers in the war in the Pacific, I wonder what they might be. One might talk about needs in this context, but that would be to move away from conscious preferences. I do not know whether Hare could accommodate the view that the conscious preferences of Japanese soldiers or Californian voters are a consequence of false consciousness, of alienation from their true humanity.

Is there any way that Moore could escape being driven, by a majority using Hare's method, to the conclusion that he must admit either to being an amoralist or a critical thinker who happens to have fantastically strong and unalterable preferences? The vital move seems to be Hare's dictum that at the critical level *no preference has any greater authority or dignity than another*. Now this is a principle which has been of great importance in the history of ethics. Kant and Bentham hold it, Mill tries not to. In the case of the first two they can do so because they accept a doctrine which can be called hedonism; and in so far as Mill's attempt to think otherwise fails, it is because he tries to remain a hedonist. What I am calling hedonism is a view held most starkly and crassly by Kant, and it is his example I shall discuss. Kant (*Critique of Practical Reason*, Bk I, § III, Remark I) argued that while our inclinations may be characterized as 'higher' or 'lower' according to the status of their objects (as poetry might be on a higher intellectual level than pushpin, as it presents itself to the 'higher' *cognitive* faculties) they are not so characterizable from the point of view of motivation (as they present themselves to the *appetitive* faculty), i.e. *qua* preferences, since all desires or preferences are for one and the same thing, namely that pleasure which results from attaining their objects, whatever the latter might be. If this were so then it would seem that the only way of rationally assessing preferences as such is in terms of their strength, their strength being relative to the degree to which each achieves more of one and the same thing, satisfaction (or, as some hedonists say, pleasure; others, utility). The crassness of the mistake has been exposed often enough. It arises because the pleonasm that if someone desires some state of affairs under the description D then he desires the satisfaction of the description D, has been mistaken for the falsehood that if someone desires something

then he also desires some consequent state of himself, some experience other than anything referred to in the description of the object of the desire. As one might expect: when one misidentifies two different things one ends up believing in a chimaera.

I do not mean to accuse Hare of this mistake (far from it: see p. 143) but rather to point out that since he does not make it his dictum that in assessing a preference at the critical level we can advert to nothing other than its mere strength needs some other justification. But what else than strength *could* be appealed to? Moore might answer 'the nature of its object' or 'the nature of the state of affairs preferred'. That the consideration of the object of a possible preference leads to the existence of an actual preference does not show that the existence of that preference is all that is germane. When the majority point out to Moore that in recommending the preservation of the Elgin Marbles he is only recommending his own preference, and he has no right to think that his preferences just by being his are more important than anyone else's, he can reply that of course his recommendation expresses his preference, and would not be recommending if it it did not, but that is not *why* he is recommending it but merely a logically necessary condition of his sincerely doing so. The majority's move seems to be as infuriating as that of people who say in answer to a contentious claim 'That's your opinion' (implying that it can have no greater authority or dignity than anyone else's); to which one's response may be to bang the saloon bar and say 'Yes, of course, it's my opinion; do you suggest I might be a hypocrite? But it isn't its being my opinion that makes it true; its being true is what makes it my opinion.' If I recommend or prescribe that some state of affairs should obtain, then if I am sincere it is logically required that I prefer that that state of affairs should obtain. It does not follow that my preferring it is my reason for prescribing it. If I believe that the Pope is John Paul II, then if I am rational I must accept that either the Pope is John Paul II or I am not riding my bicycle, but not that my reason for believing the Pope is John Paul II is that I am riding my bicycle.

The distinction between a recommendation or prescription *depending* on the nature of the object of recommendation rather than on a preference seems to be inherent in our everyday discriminations of preferences. I prefer coffee to tea, but not through any consideration of what it is, and I am quite indifferent to what I find my preference to be: I do not care if tomorow it changes. But my preference for people being free, autonomous, self-directed (if discontented) individuals as opposed to happy slaves must come from a consideration of what it is to be each, and (however much it may contribute to my discontent that there are not enough of the former and too many of the latter) I do care that my preference should not change tomorrow, though of course if it did I should presumably not then care that it had. One feels that the latter sort of preferences seem much more like preferring a man's views to preferring his face. My

preferences for tea over coffee are obviously different from my aesthetic preferences. In the former case *de gustibus*, etc., in the latter; not. In disputing another's aesthetic preferences I recommend that the other adopts mine (and that I should stand by my own); not because they are my preferences, but because of what sorts of things they are for. Hare finds an explanation for the disputability of aesthetic judgments in their universalizability; I would add that their universalizability rests on the fact that the preference is supposed to be based on objective consideration of the nature of the object, in which matters concerning the identity of the subject of the consequent preference drop out. Thus, in cases of egregious failure of aesthetic preference (say for the music of the Tweets over that of the Berlin Philharmonic) one tends to attribute the failing to the judge: not because one's preferences are different or stronger (they might not be the latter) but because of what they are *for*. Such preferences are demanded, though not expected: as if the object measures the preference, not the preference the object.

Hare characterizes someone (5.6, p. 105) as an 'autofanatic' with respect to prudence who will sacrifice the satisfaction of more and stronger future preferences for the sake of fewer and weaker present preferences. Imagine a man who has a preference for people's being free, autonomous, self-directed individuals rather than happy slaves, being told that since he is one of the most dangerous opponents of the Happy Slave party in power, consistently with their policy of preferring slaves to be happy, in a week's time he will be brainwashed into becoming as dedicated a happy slaver at the rest. If he is prudent and certain of the threatened outcome, he should do all he can to frustrate the policies of the Free Man opposition, because the more successful they are the less pleased he will be after being brainwashed. If he is imprudent, he might do the opposite: continue to try to frustrate the policies of the Happy Slave government, and indeed any actions he might undertake after the end of the week. I accept the first course as prudent and the second as imprudent, using the term in Hare's sense in which it is prudent to allow one's future preferences the same weight as present ones. But why should one be prudent in this sense? Hare's answer is that the agent is about the business of how optimally to satisfy *his* preferences; and the mere time at which such preferences occur cannot rationally be relevant. The only way therefore that he could discount any preference, such as some future one, would be to claim that it will not be *his* preference. But couldn't he say he will imprudently but rationally discount certain future preferences, not just because they are future ones, but because of what they are *for*? He could say we mistake his business: it is not optimally satisfying his preferences whatever they are and whenever they occur, but of satisfying those preferences which are for a certain sort of object. His present preferences are for certain objects, S, and he prescribes his pursuit of them, but not because

they are his, or present, but because of what they are for, S. However, Hare thinks this is a move he cannot make if he is using the 'prescriptive' sense of 'I'.

The 'prescriptive' sense of 'I' is explained as follows (5.3, pp. 96–97):

> The suggestion is that 'I' is not wholly a descriptive word but in part prescriptive. In identifying myself with some person either actually or hypothetically, I identify with his prescriptions. In plainer terms, to think of the person who is about to go to the dentist as myself is to have now the preference that he should not suffer as I believe he is going to suffer. In so far as I think it will be myself, I now have in anticipation the same aversion as I think he will have.

The words 'now have' and 'have now' in this passage are very important. Hare is not simply appealing to the principle that if a is identical with b, a has all the characteristics of b (including b's preferences), which would not give him what he wants; which is, that if I fully appreciate some other person's preferences, then, if I think of myself *as* he, knowing at the same time I am not he, I now share that preference, that is, *now* prescribe as he would. The word 'I' is attached to a variety of identifying properties, but what we do when we regard the term as having a prescriptive sense in addition to these properties is recognize that 'by calling some person "I", I express at least a considerably greater concern for the satisfaction of his preferences than for those of people I do not so designate' (p. 98). Now this cannot depend on a mere psychological generalization; indeed it may be a very doubtful psychological claim with respect to some people who are considerably less self-concerned than most of us. The point must be a linguistic one. But if there is such an implication of the use of the word 'I', it is surely an eminently withdrawable one. It may be that by using the word 'I' in this way, one is recommending this attitude towards whoever is designated by 'I'; but someone who does not wish to recommend this attitude could without great difficulty use the word otherwise. One could use the word 'I' primarily in a token-reflective sense, so that when a sentence is uttered containing the word 'I' (leaving aside quotations etc.), 'I' refers to the person uttering that sentence. The person will be identified and re-identified by whatever are the normal criteria of identity of persons but the prescriptive addition is dropped, so that in using the term in this way one is not forced by logic to be prudent in Hare's sense. Of course it may well be that persons who use the term 'I' without the prescriptive requirement will in fact show more concern for their own preferences, perhaps because they feel better able to know what they are, as Mill claimed: perhaps because for a very great range of preferences it is vastly easier for the person who has them to satisfy them than for someone else to attempt the task—such as clearing one's throat, jogging, or acquiring a foreign language, as those who have brought horses to water lugubriously

admit; because, perhaps, there is some built-in tendency for each of us to do so because of its value to our genes: or, because perhaps the present stage dictated by our relations to the means of production requires us to be acquisitive individualists. Again, someone might adopt this attitude towards some individuals and not others: a man might generally give his own preferences priority over others', but show greater concern for the preferences of his wife. But either way it need not be because one is or is not attaching a prescriptive feature to the term 'I'.

If these reasons throw any doubt on the requirement of prudence, then we can also see that the imaginary Moore need not give in to the majority's claim that if he properly appreciates their preferences, knowing what they would prescribe in their position, he must *now* prescribe it himself.

It would seem then that there is still logical space for the fanatic between the amoralist and the utilitarian, as there is for the rational autofanatic. Hare originally introduced the notion of the fanatic in *FR* in terms of conflicts between interests or preferences, and ideals. Those who follow the utilitarian doctrine may have ideals, but they will treat them as some among other preferences in making interpersonal decisions; whereas the fanatic will allow his ideals to override (his own and others') other preferences. The discussion of the fanatic's position was in terms of Nazi ideals, and Hare anticipated objections against the use of such 'loaded' examples. I agree that these were not telling objections, because Hare needed to make his case in terms of the requirements of philosophical logic, and it is the form and not the content of the examples that matters. That is why I chose to make logical room for so egregious a fanatic as the imaginary Moore. Hare also gives reasons of a very general empirical sort for holding that the pure fanatic must be very rare. But it is difficult to decide whether this is so. Hare characterizes the non-fanatical utilitarian position as the liberal one; but among those who are called liberals there may be in addition to those who take into account only the strength of preferences, others for whom the concern for the satisfaction of preferences and the concern for autonomy are distinct and, indeed, able to conflict. Some who defend the virtues of the free market may do so because they regard it as a wonderful device for the optimum satisfaction and distribution of satisfaction of preferences, wherein self-seeking individuals pursuing their own preferences are held in the hidden hand of a utilitarian god; though, in consistency, they should agree to advocate slavery and tyranny if it in fact turns out that people prosper better as slaves. Others who advocate the free market may continue to do so because they have ideals of freedom which they may put above ideals of prosperity. Some of the latter may even be subject to an error opposite to one Hare suggests may be made. He points out that some who hold respect for preferences and for autonomy as *prima facie* intuitive principles which are not only distinct but fundamental would on reflection see that their ideal of autonomy is justifiable only by the

application of critical utilitarian principles. But another may think his respect for autonomy is dependent on critical utilitarian principles when, as might be shown by his use of it, it is not. Mill begins *On Liberty* with a justification of liberty in terms of the best way of generally satisfying preferences, but the way the discussion subsequently goes often suggests that for Mill it is an end in itself. Certainly, at the end of the *Principles of Political Economy* he suggests that the (then) Chinese contented acceptance of customary ways of life should not be tolerated by more enlightened idealists, who should force them to be free.

University of Warwick

https://doi.org/10.1017/S0031819100066651 Published online by Cambridge University Press