

HARE'S DEFENSE OF UTILITARIANISM

(Received 18 July, 1985)

R. M. Hare's *Moral Thinking* is surely one of the most compelling defenses of utilitarianism to appear in many years. Hare defends utilitarianism at some length against the objection that it has consequences that are inconsistent with our common-sense or intuitive moral judgments. Hare also offers a positive argument for utilitarianism. In this paper I shall only concern myself with the latter argument. In the first part of the paper, I shall set out Hare's argument in some detail. In the second part of the paper, I shall suggest criticisms of Hare's argument. I shall argue that two of the assumptions upon which his argument rests are false and that a third assumption is open to serious criticisms. The final section of the paper is a criticism of Hare's proposed solution to the problem of interpersonal comparisons of the strengths of desires (a solution which enters importantly into his argument for utilitarianism).

I

Hare's positive argument for utilitarianism rests on his defense of a version of the ideal observer theory. After identifying the term "archangel" with "ideal observer", Hare writes the following:

Provided that we do not give it a 'subjectivist' or 'relativist' interpretation (12.1), there is no harm in saying that the right or best way for us to live or act either in general or on a particular occasion is what the archangel would pronounce to be so if he addressed himself to the question.¹

Hare argues that facts and logic will so constrain the moral judgments of ideal observers that they will be compelled, in practice, to make the same moral judgments about all cases, concerning which they make any moral judgments at all (*MT*, p. 6). The moral judgments that facts and logic compel us to make are all ones that would follow from a preference version of act-utilitarianism. According to Hare, rational constraints alone cannot compel us to make any

moral judgments — amorality is a possibility. However, if one makes any moral judgments at all, then facts and logic compel one to make judgments that are consistent with those that follow from a preference version of act-utilitarianism.

Hare's positive argument for utilitarianism begins with a restatement of his argument for a prescriptivist analysis of the meaning of moral judgments. According to Hare, moral judgments are prescriptions that are 'universalizable' and 'overriding' (*F & R*, p. 55ff.). Moral judgments prescribe actions in just the same way that commands and requests do. They differ from mere commands or requests, e.g., "close the door", in that they are universalizable and overriding. To say that moral judgments are universalizable means roughly that, if one makes a moral judgment about a particular case, then one is committed to making exactly the same moral judgment about any similar case, unless there is a relevant difference between the two cases. To say that moral judgments are overriding means that one takes the prescriptions expressed in the moral judgments that one accepts to be prescriptions that ought to be followed in cases of conflict with other sorts of prescriptions, e.g., conflicts between morality and considerations of prudence, etiquette, and the law. According to Hare, one of the rational constraints that the universalizability of moral judgments imposes on moral appraisers is that, if one makes a moral judgment about a particular case, then it must be the case that one is willing to make the same judgment about the otherwise identical cases in which one occupies the position or place of one of the other parties who is affected by the action in question. So, for example, if one says that slavery is morally permissible, then it must be the case that one is willing to say that, under certain conditions, it would be morally permissible for someone else to hold one as a slave (*MT*, pp. 89 and 109 and *F & R*, Chapters 6 and 11).

An important and highly controversial feature of Hare's argument, which I shall criticize later in the paper, is his denial of the existence of, even the possibility of, moral weakness. Hare claims that a necessary condition of one's accepting a moral judgment is that one act in accordance with it whenever it is possible for one to do so (*F & R*, p. 79; and *MT*, pp. 58–59). Hare holds that a necessary condition of S's accepting a moral principle to the effect that S ought to do *x* is that he do *x* if it is in his power to do it and he is aware that the principle in question commits him to doing *x*. If this is taken together with the requirement of universalizability, we get the following

result: if S accepts a moral principle that implies that he (or someone else) ought to do x at t and if he knows that this principle commits him (or the other person) to doing x at t , then it must be the case that S now prefers or "wills" that x be performed at t , even if he were in the position of one of the other parties affected by his doing x . So, if I say that killing Jews is morally obligatory, then it must be the case that I now prefer that I be killed myself if I am a Jew. The vast majority of us strongly prefer that we not be killed under almost any circumstances. Therefore, most people are rationally constrained to reject the judgment that killing Jews is morally obligatory or morally permissible. This is the kind of argument developed at length by Hare in *Freedom and Reason*.

However, such arguments cannot always succeed in compelling people to abandon views that we would regard as morally perverse. For, as Hare himself notes in *Freedom and Reason*, one might be so fanatically devoted to a perverted ideal that one would consent to policies that advance the ideal, even if one were to find oneself in the position of someone who might be grievously harmed by those who act in accordance with that ideal. Hare gives the example of the fanatical Nazi who discovers that he is Jewish and consents to being murdered (*F & R*, Chapter 11). The machinery of *Freedom and Reason* does not seem to offer any satisfactory way of dealing with such cases. Hare and many others regard this as a serious shortcoming of his earlier view. One of the things that allows the case of the "fanatic" to be a problem for the position defended in *Freedom and Reason* is that when the "fanatic" imagines himself in the position of the person who might be his victim he retains his present fanatical preferences and desires. He asks himself the question "what would I want to happen if I were in the victim's position with my present desires?" The actual preferences of the victim to the effect that he not be harmed are now[†] allowed to come into play; all that is allowed to come into play are the fanatic's *present* desires concerning those hypothetical cases in which he is in the position of one of his victims. In *Moral Thinking* Hare, in effect, argues that when we place ourselves in the position of another person and ask whether we can still accept the prescriptions that we make we must have taken on his desires and preferences, at least conditionally; we must have a desire that the other person's desires be satisfied, if we find ourselves in his situation with his desires and preferences, etc. This requirement is justified on the grounds that one can't understand what the other person's situation is like *for him* (with his preferences and desires), unless

one now has a preference that if one were in the position of the other person (with his preferences and desires) those preferences that one would have should be satisfied.

I am to imagine myself in his situation with *his preferences*. Unless I have an equal aversion to myself suffering, forthwith, what he is suffering or going to suffer, I cannot really be knowing, or even believing, that being in his situation would be like *that* (MT, pp. 94–95).

I cannot know the extent and quality of others' sufferings and, in general, motivations and preferences without having equal motivations with regard to what should happen to me, were I in their places, with their motivations and preferences (MT, p. 99).

It is important to be clear exactly what this requirement amounts to. Hare is not saying that in order to understand another person's situation I must have or acquire exactly the same desires that the other person has.² Nor is Hare's requirement equivalent to saying that if I were in S's position and had his wants and desires for *x*, *y*, and *z*, then I would desire that *x*, *y*, and *z*. This is something that is trivially true. If I were in S's position and had his desires, then I would have his desires. Hare's view is that in order for me to understand what it would be like to be in S's position with his wants and desires for *x*, *y*, and *z*, I must now have the following desire: "Should I ever be in S's position and have his desires for *x*, *y*, and *z*, then let it be the case that *x*, *y*, and *z*." Suppose that, on balance, S prefers that *x*. We might take Hare's view to imply that in order for one to understand S's situation, one must now have the *on balance* preference that *x*, should one find oneself in S's circumstances with his wants and desires, etc. However, in a subsequent reply to other critics, Hare explicitly rejects this interpretation of his view.³ He claims that in order to represent S's situation adequately it is only necessary that one now have a *preference* (equal in strength to S's preference for *x*) that *x*, if one finds oneself in S's circumstances with his wants and desires, etc. This preference or desire for *x* if _____ may or may not be counter-balanced by other desires that one has concerning what should happen if _____.

In order to make a moral judgment about some issue, one must understand the position of *all* of the affected parties. In order to make a moral assessment of an act that affects more than one person, one must have conditional desires for the satisfaction of all of the preferences of all of the individuals in question. Suppose that act *a* affects persons $p_1 \dots p_n$. In order to represent everyone's position, I must have all of the following conditional desires: (1) Should I be in the position of p_1 with his wants and desires (d_1),

then let things happen in accordance with $d_1 \dots (n)$ Should I find myself in the position of p_n with his wants and desires (d_n), then let things happen in accordance with d_n . Whether or not an act is right depends on the strengths of these conditional desires and thus ultimately upon the strengths of the preferences of the affected parties concerning that act. Hare says that we must treat each person's preferences impartially and that, therefore, we are led to a preference version of act-utilitarianism (*MT*, p. 129). He attempts to avoid the notoriously difficult problem of making comparisons of the strengths of the desires of different people. Since the ideal observer has all of the preferences of the affected parties (at least conditionally), he needn't sum up the different preferences, he only needs to find out what his own overall preferences are. Referring to these features of his view, Hare writes the following:

they enable us to reduce comparisons between other people's preferences to comparisons between our own. ... in so far as I fully represent to myself the strengths of other people's preferences, I have preferences, myself now, regarding what should happen to me were I in their positions with their preferences. If I do this with two other people or more, I can then compare the strengths of their preferences by comparing the strengths of my own corresponding ones (*MT*, p. 128).

For Hare's purposes it is not enough that the ideal observer has the conditional preferences discussed above and that he determines what he most prefers, all things considered. For this allows for the possibility that the ideal observer knows that he is in the position some particular individual, in which case, he might give absolute priority to the interests of that individual. That this is a possibility can be seen by considering by the following example. In order to understand the positions of both person A, who desires x , y , and z and person B who desires that x , y , and z not be the case, one must now have desires of the following conditional form:

- (1) Should I ever be in A's position with his wants and desires, then let it be the case that x , y , and z .
- (2) Should I ever be in B's position, with his wants and desires, then let it be the case that not- x , not- y , and not- z .

An ideal observer must then form an *overall* preference on the basis of such desires. This overall preference will be a "sum" of the preferences of the individuals in question. Being a utilitarian, Hare, of course, thinks that each individual's preferences should be given equal weight. However, since desires (1) and (2) are conditional in form they cannot influence the overall prefer-

ence formed by the ideal observer, unless he thinks that there is at least some chance that the antecedent condition is satisfied, i.e., unless he thinks that there is at least some chance that he might occupy the position of the particular individual in question. If an ideal observer knows that he will never be in the position of person A, then his desire that "should he find himself in A's position with A's wants and desires then those things that he would want should be the case" cannot possibly influence the overall preference that he (the ideal observer) forms.

Clearly, some restriction must be made on the beliefs of the ideal observer concerning whose position he occupies. He cannot know or believe that he will fail to occupy the position of any particular person, since the interests of the person in question may not be taken into account in that case. It might be suggested that we conceive of the ideal observer as occupying the positions of all of the affected parties simultaneously. Hare says that this is "too much" to expect of an ideal observer (indeed it hardly seems possible at all) (*MT*, p. 129). He suggests the following instead: an ideal observer assumes that either: (1) he will live the lives of all of the affected parties in some random sequence, being in one person's shoes at one time and another person's shoes at a later time, or (2) he will live the life of just one of the affected parties and he has an equal chance of being any one of them (*MT*, p. 129). (Presumably, the preferences of the ideal observer will be the same in either case.)

The preferences of others that one adopts in the conditional form explained earlier may conflict with one's present preferences. In some cases, an individual may now prefer, on balance, that his future, on balance, preference for *x* be thwarted rather than satisfied, should he find himself in the position of someone who desires *x*, all things considered. Hare claims that an archangel should resolve such conflicts between his now-for-then and then-for-then preferences in accordance with the following principles:

This conflict will be resolved if we add to our present preferences what we may call the *requirement of prudence*. This is that we should always have a dominant or overriding preference now that the satisfaction of our now-for-now and then-for-then preferences should be maximized. ... I propose the following simplifying assumption. We are to assume, when we come to universalize our prescriptions, as morality demands, that we have to consider only those preferences of others that they would retain if they were always prudent in the sense just defined (*MT*, pp. 105-106).

I shall now argue at some length that the requirement of prudence is a necessary assumption for the purposes of Hare's argument. The requirement of prudence is essential for his claim that ideal observers would all agree in

whatever moral judgments they might make. In turn, the claim that "facts and logic" constrain all ideal observers to make the same moral judgments, is essential for Hare's claim that we are compelled (on pain of error) to make moral judgments that are consistent with a preference version of act utilitarianism, provided that we make any moral judgments at all.

Suppose that X and Y are both involved in "open marriages" to other individuals and that each desires to have an affair with the other. Suppose also that whether or not they have an affair will not affect the happiness or welfare of anyone else. X desires, on balance (with strength s), to have an affair with Y. In order to represent X's circumstances, an archangel must now have a desire (of strength s) that Y sleep with him if he is in the place of X with X's wants and desires, etc. But, as Hare is at pains to stress, this is not necessarily an all things considered desire that he (the archangel) sleep with Y, if _____. This desire might be counter-balanced by other desires, including (presumably) a commitment to certain non-utilitarian ideals. An archangel might be appalled at the thought of his being the sort of person who has an open marriage, and might now prefer, on balance, that his desire to engage in extramarital affairs be thwarted if he becomes the sort of person who desires to have such affairs. This undermines Hare's claim that all ideal observers would agree in whatever moral judgments they make. In this sort of case, it would be possible for different ideal observers to make incompatible moral judgments. (For different ideal observers might have different moral ideals that would incline them to make conflicting moral judgments in certain cases.) That this is so can be more clearly seen when one reflects that there are certain to be borderline cases in which an ideal observer's balance of preferences concerning the choice between x and not- x is nearly equal. In such cases, an ideal observer's commitment to a particular set of ideals could affect his on balance preference. Given the requirement of prudence, cases of this sort could not arise, since this requirement makes it impossible that an ideal observer or archangel could be the type of person who allows his now-for-then preferences to override his then-for-then preferences in cases in which they conflict.

In the foregoing argument, I have assumed that, unless justified by *appeal to* other essential features of an ideal observer, e.g., impartiality or complete information, any requirement to the effect that all ideal observers must either have or fail to have certain ideals would be implausible. The ideal observer theory is intended to be a standard for determining the truth or correctness

of any given moral judgment or moral ideal. The theory cannot serve this function if accepting one particular moral ideal, as opposed to another, is taken to be an explicit condition of one's being an ideal observer. For the claim that accepting a particular moral ideal is an essential feature of an ideal observer could only be justified if one had grounds (independent of the ideal observer theory) for thinking that the ideal in question is true or correct. Hare and other proponents of the ideal observer theory deny that there are any standards of truth for moral judgments other than the attitudes or judgments that would be held by ideal observers (see, *MT*, p. 46).

Here, it might be objected that, not only is it a mistake to assume that an ideal observer *must* accept *some particular* moral ideal or moral standard, it is also a mistake to *allow* that an ideal observer could be predisposed to accept *any* moral ideal or moral standard. One could stipulate that the only ideals or moral judgments that an ideal observer can hold are those that he holds as a consequence of his other essential characteristics, e.g., rationality and knowledge of all of the relevant facts, etc. However, my argument to show that Hare needs the "requirement of prudence" can be revised so as to be consistent with the proposal in question. For, differences in mere preferences that one does not take to be demanded by any moral norms could still create disagreements between ideal observers. One might now have a strong preference that one never be a partner in an open marriage, even if one should come to desire to be a part of such a marriage at some time in the future. Having such a preference does not require that one believe that there is anything morally wrong with having an open marriage. Given that some ideal observers could have such preferences and others not have them, ideal observers would disagree in their judgments about at least some cases. Hare needs to presuppose the requirement of prudence in order to rule out this possibility.

Hare's argument can be summarized as follows:

- (1) Moral judgments must be universalizable; if one makes a moral judgment about a particular case, then one must make the same judgment about all cases that are the same in all morally relevant respects, including those hypothetical cases in which one occupies the positions of the different individuals who are affected by the action or policy in question. (If a hypothetical case (*c'*) is just like another actual or hypothetical case (*c*) except that a particular person occupies a position in *c'* that he does not occupy in *c*, then *c* and *c'* are identical in all morally relevant respects. The fact that a particular individual occupies a certain position in one case and a different position in another case is not a morally relevant difference.

For instance, it is not a relevant difference if I am in the position of the master in one case and the position of the slave in the other.)

- (2) Moral judgments must be impartial: if one makes a moral judgment about a particular case, then it must be the case that one would be willing to make the same judgment if one believed that one had an equal chance of being in the place of any of the people involved in the case in question.

The conjunction of (1) and (2) entails the following:

- (3) If one makes a moral judgment to the effect that an act *a* is obligatory, and if *a* affects individuals *x*, *y*, and *z*, then it must be the case that either (or both?): (i) one would judge that *a* is obligatory if one believed that one were going to live the lives of *x*, *y*, and *z* in some random sequence, (ii) one would judge that *a* is obligatory, if one believed that one had an equal chance of being *x*, *y*, or *z*.

Hare adds the following assumptions:

- (4) Moral judgments require that one have knowledge of all relevant facts – including facts about the feelings and experiences of others. Among other things, one must understand what it would be like to be in the positions of other people and have their wants and desires.
- (5) In order for one to understand what it would be like to be in S's position (with his wants and desires for *x*, *y*, and *z*), it must *now* be the case, that one has a *desire* (equal in strength to S's desire for *x*, *y*, and *z*) for *x*, *y*, and *z*, if one should find oneself in S's position with his wants and desires, etc.
- (6) If S accepts a moral judgment or moral principle that implies that he ought to do *a* at *t*, then it must be the case that S does *a* at *t*, provided that: (i) he is aware that the judgment or principle in question commits him to doing *a* at *t* and (ii) he is able to do *a* at *t*.

Hare not only wants to say that any moral judgments that one makes about *one's own actions* must be consistent with the principle of utility, he wants to claim that any moral judgments which one makes about *the actions of others* must be consistent with the principle of utility. Therefore, the following assumption is necessary for the purposes of Hare's argument:

- (7) If one judges that Z ought to do *a* (where Z may or may not be identical with oneself) then it must be the case that one would prefer that Z do *a* and choose Z do *a* if it were within one's power to decide.⁴

From the conjunction of (3), (5), and (7) it follows that:

- (8) If one makes a judgment to the effect that act *a* is morally obligatory, and if *a* affects *x*, *y*, and *z*, then it must be the case that one would choose or prefer that *a* be performed if both of the following were the case: (i) for each person affected by *a*, one has a conditional desire that his preferences be satisfied, if one is in his position with his desires and preferences, etc., and (ii) either one believes that (a) one will live the entire life of

each of the affected individuals in some random sequence, or (b) one believes that one will live the life of one of the affected parties and that one has an equal chance of living the life of any one of them.

The following assumption can then be conjoined with (8):

- (9) A rational moral appraiser must have a dominant or overriding preference that the satisfaction of his now-for-now and then-for-then preferences be maximized and when universalizing his prescriptions as morality demands, he has to consider only those preferences of others that they would retain if they always satisfied the above condition.

An ideal observer believes that he has an equal chance of being in the position of anyone. He desires that all of his future desires be satisfied — no matter what they are; he takes all future desires of equal strength to be equally important. Thus, an ideal observer will, on balance, prefer the performance of all and only those acts that maximize the satisfaction of the preferences of individuals. Thus we see that the conjunction of (8) and (9) entails the following:

- (10) If one makes a moral judgment to the effect that an act *a* is morally obligatory, then that judgment must be one that is consistent with the judgments that would follow from a preference version of act-utilitarianism, i.e., if one makes moral judgments about the rightness or wrongness of actions, then one must make those judgments in accordance with the following standard: an act is morally right if and only if there is no other course of action open to the agent which would result in a greater net satisfaction of the desires of individuals.

Note that Hare does not claim that everyone is rationally constrained to be a utilitarian. In accordance with his earlier view, Hare holds that one can refrain from making any moral judgments without pain of error. His claim is that *if* one makes a moral judgment about a particular case, then “facts and logic” require that one’s judgment be the same judgment that would be entailed by a preference version of act utilitarianism.

II

Here, I shall state some reasons for thinking that, for all of its merits, we should not accept Hare’s argument. I shall argue that premises (5), (6), and (7) are false and that premises (9) is open to serious question.

(a) Hare’s view about what is required in order to understand what it would be like to be in the position of another person (premise (5)) is much too strong, as can be seen by the following example. Suppose that a corrupt

and ambitious politician feels great frustration and humiliation on account of having lost an election. Suppose also that the politician is willing to do almost anything, including bring about the deaths of many innocent people, in order to reverse this outcome. According to Hare, in order for me to represent the politician's feelings and situation adequately, it must be the case that I now have a *desire* that all of the politician's preferences be satisfied if I should find myself in his position with his wants and preferences, etc. Suppose that the politician is in a situation in which he needs to kill in order to gain office. Hare's view implies that, in order for me to represent the politician's situation adequately, I must have a *desire* (equal in strength to the politician's desire that the necessary people be killed) that the necessary people be killed, if I find myself in the position of the politician with his wants and desires, etc. I don't have any such desires. It is not just the case that I have a *desire* that people be killed in order that I gain office, if I am in the position of the politician and that this desire is counter-balanced by other stronger desires. (Hare thinks that this possibility is compatible with my representing the politician's situation adequately.) Rather, I now have no particular desire that the people be killed, should I find myself in the politician's situation. The politician, by contrast, has an intense desire that innocent people be killed, should this further his political ambitions. But, Hare to the contrary, this is in no sense a limitation on my ability to represent the politician's experiences. All that is necessary for an adequate representation or understanding of the politician's position is that I understand the preferences that he has (how he is disposed to act in various actual and hypothetical situations) and that I understand the experiences that he is undergoing and facts about how he is likely to feel in different situations. These conditions are all satisfied in the present case. I know what the politician's preferences are. I know, for example, that he is willing to kill in order to gain power. Second, I know how he feels; I know what it is like to feel frustration and humiliation myself. That I would not feel humiliated if I lost an election myself and that I lack any kind of desire that things should go as he wants them to go, if I should find myself in his situation, in no way detracts from my understanding of his situation. Surely it is possible to *understand* what it would be like to be in another person's situation without it being the case that one shares all of the same desires as the person in question. It is equally clear that one can understand what it would be like to be in another person's position without having conditional desires (of the sort that Hare requires) that the other person's desires be satisfied.

(b) Premises (6) and (7) are tantamount to a denial of the existence of moral weakness. I suggest the following as a rough analysis of the concept of moral weakness:

S exhibits moral weakness in doing *a* at *t* iff: (1) S believes that his doing *a* at *t* is wrong, (2) S is able to refrain from doing *a* at *t*, and (3) S does *a* at *t* deliberately.

According to Hare, it is *logically impossible* for there to be cases of moral weakness – so defined. Hare's view also implies that it is logically impossible for one to believe that S ought to do *x* at *t*, unless one would choose or prefer that S do *x* at *t*, if it were within one's power to decide. Hare has not given us sufficient reason for denying what seems to be one of the most obvious features of the moral life, that we often act in bad conscience and do things that we know or believe that we ought not to do. Suppose that I say that I believe that I ought to give a large portion of my income to famine relief but fail to do so, even though it is clearly within my power to do this. Hare would take this to be evidence for the insincerity of my judgment (or at least, evidence for the claim that I am mistaken in thinking that I accept this judgment). If I really thought that I ought to give away a lot of money, then I would. But this is implausible. For one thing, Hare's position does not seem to be able to explain the guilt that is often felt by those who fail to do what they profess to believe that they ought to do. The most plausible explanation of the fact that people often feel guilty or remorseful on account of their conduct is that they often do things that they believe to be morally wrong.

Hare attempts to reply to this objection in a number of places in his writings. He suggests that those cases of guilt and remorse that constitute ostensible cases of moral weakness are not cases of feeling guilty for failing to do what one believes one "really" ought to do, but rather cases of one of the following two sorts: (1) acts that are contrary to conventional moral standards, but, which one does not consider to be "really" wrong (clearly, one might be disposed to feel guilty for acting contrary to conventional moral principles in which one has been indoctrinated, even if one no longer regards those principles as correct or reasonable), or (2) acting against an intuitive principle that one takes to be only *prima facie* correct, e.g., telling a lie (*MT*, pp. 31 and 58–59; and *LM*, pp. 166–170). According to Hare, a person who is "properly brought up" will be inclined to feel guilt or remorse

for telling lies, even if he thinks that, all things considered, it is permissible for him to lie in the case in question. No doubt, many ostensible cases of moral weakness fall into one of these two categories. However, Hare needs to say that *all* such cases can be so classified. He needs to say that in all cases in which people feel guilt and remorse for failing to act in accordance with moral principles that they claim to accept, the principles in question are merely conventional principles or *prima facie* principles. However, some ostensible cases of moral weakness cannot be explained away in this manner. Consider the following case. Suppose that a member of *this society* believes that he ought to eat meat and that he ought to give away at least one half of his income to famine relief. Suppose also that this person acts contrary to these beliefs and that he feels guilty as a result. Surely this is a genuine case of moral weakness. The individual's principles are not mere *prima facie* principles; he believes that, *all things considered*, he should give away half of his income and refrain from eating meat. Nor can these principles be said to be part of a conventional code in which he was once indoctrinated, but which he now rejects. (He was not brought up to believe that it is wrong to eat meat or to fail to give away half of one's income.)

In *Freedom and Reason*, Hare suggests yet another way to explain away ostensible cases of moral weakness. Individuals who appear to be morally weak may be "psychologically incapable" of doing what they think that they ought to do (p. 79). To my knowledge, Hare never gives an explanation of what he means by being "psychologically incapable" of doing something. In the absence of an explication of this notion and reasons for applying it to the types of cases in question, I see no reason to suppose that all cases in which a person fails to act as he believes he ought to (all things considered) when he is physically capable of doing so involve some kind of psychological inability. As Hare suggests, in the case of certain kinds of mental illnesses it seems *prima facie* plausible to say that someone is psychologically incapable of doing something that he would be able to do otherwise. For example, we might say that a kleptomaniac is incapable of refraining from stealing. However, we still need to be given an account of what it is to be psychologically incapable of doing something in order to compare such cases with ostensible cases of moral weakness. On the face of it, such cases do not seem comparable. For example, it does not seem at all plausible to say that I am *incapable* of giving more money to charity, but only that I am *unwilling* to do so.

(c) I shall now argue that there are strong reasons for questioning the

truth of Hare's "requirement of prudence". I will not attempt to try to show that the requirement of prudence is false (indeed I am not sure that it is false). Instead, I will simply try to show that the falsity of the requirement of prudence follows from the truth of certain plausible and traditional theories of the good, theories which Hare has not given us any reasons for rejecting.

The requirement of prudence states that, in order to be rational, one must have "a dominant or overriding preference now that the satisfaction of our now-for-now and then-for-then preferences be maximized" (*MT*, p. 105). Most importantly for our purposes here the requirement of prudence implies that it is in some sense irrational for one to (now) desire the frustration of one's future desires. If the requirement of prudence is taken to imply that it could *never* be rational for one to desire the frustration of one's future desires, then it is most implausible. For, surely it would be rational for a morally good or morally decent person to desire the frustration of immoral desires that he might acquire in the future. It is perfectly rational for me to (now) desire the frustration of any desire to decapitate strangers that I might acquire in the future. (Here, we might say that I don't desire the frustration of my future desires as an end in itself, but only as a means to insuring the welfare of others.) The requirement of prudence is plausible, only if it is restricted to "self-regarding" or "purely self-regarding" desires, i.e., only if it is understood in something like the following way:

In order to be rational, one must have a dominant or overriding preference for the satisfaction of one's *purely self-regarding* now-for-now and then-for-then preferences over the satisfaction of other kinds of *purely self-regarding* preferences, in cases of conflict.⁵

Given certain conceptions about the nature of the good or the good life, it makes perfect sense for a person to now desire the frustration of his future self-regarding desires. Suppose that we accept a theory of the good according to which having a good life requires (or consists in) developing one's "higher" capacities, e.g., one's artistic and intellectual capacities. Such a theory of the good has been defended by many philosophers, including Aristotle, Marx, and Nietzsche. Suppose that I now accept such a theory and that a reliable fortune-teller informs me that the following future awaits me. I will become a rock-and-roll star, move to California, and come to live a life devoted to the pursuit of physical pleasures (derived from sex and drugs) and the adulation

of my numerous fans. I will at this time cease to desire self-fulfillment in the Aristotelian sense and will strongly prefer the continued pursuit of sensual pleasures to development of my intellectual and artistic powers. Suppose that I am also informed that my future preferences could be thwarted if I lose my popularity as a singer and am thus compelled to earn my living as a philosopher again. Given the truth of the Aristotelian conception of the good life, it is perfectly rational for me to *now* desire the frustration of these future preferences. For, on these assumptions, the frustration of my future preferences would give me a better life than I would otherwise have.

The foregoing argument, in effect, claims that if an Aristotelian conception of the good is true or correct, then the principle of prudence is false. A more forceful argument can be stated as follows. Any plausible version of the ideal observer theory ought to be able to address the question of the truth or plausibility of an Aristotelian conception of the good. Further, this ought to be, in some significant sense, an open question. It is not a question the answer to which should be determined by some formal constraint on the concept of rationality. However, Hare's insistence on the requirement of prudence as a formal constraint on rationality automatically implies that the preferences of ideal observers will not be compatible with an Aristotelian conception of the good. This is unacceptable. Surely it is a mistake to think that a theory such as Aristotle's can be refuted by mere appeal to formal features of rationality.

I confess to being very perplexed by the phenomenon of changing desires and I have no definite idea of what it is rational for one to want for the future, if one has reason to think that one's future desires will conflict with some of one's present desires.⁶ *Anything* that one says about this topic is problematic and that Hare *needs* to take a stand on this issue is a serious problem for him.

III

Hare claims to have given a solution to the long standing problem of comparing the strengths of the desires of different individuals. Such comparisons, he claims, can be reduced to comparisons of the strengths of one's own desires. I am not persuaded that Hare has succeeded in "reducing comparisons between other people's preferences to comparisons between our own". His theory *presupposes* a solution to the difficult problem of making interpersonal

comparisons of the strengths of desires. Presumably, there is no problem in making (ordinal) comparisons of the strengths of the desires of a single individual at a given time. To do this we only need to ask about how the individual in question acts (or is disposed to act) in various actual and hypothetical situations. For example, given certain verifiable assumptions such as that I was not motivated by any considerations other than those having to do with my health, the fact that I quit smoking allows us to conclude that my past desire for good health was stronger than my past desire to smoke. The problem of making interpersonal comparisons is not just one of being in a position to claim to *know* that one individual's desire or preference is stronger than another's, the problem is one of making clear sense of what it would *mean* for the one desire to be stronger than the other.

Consider a case in which two people have conflicting desires concerning x . On balance, A desires that x and, on balance, B desires that not- x . Suppose that A and B are the only people affected by whether or not x occurs and the only people who care whether or not it occurs. An archangel or ideal observer must have conditional preferences for both x and not- x . He must now have a preference that x , if he is in the position of A and he must have a preference that not- x , if he finds himself in the position of B. The archangel will also have some kind of overall preference between x and not- x . In our present example, one of the following must be the case: (i) A's desire for x is stronger than B's desire for not- x , (ii) A's desire for x is weaker than B's desire for not- x , or (iii) the two desires are equally strong. As a utilitarian, Hare believes that the relative strength of two preferences is morally significant; other things being equal, we ought to be more concerned about satisfying the stronger preference. We should take the archangel's overall preferences to be a criterion for the truth or falsity of moral judgments, because the ideal observer's preferences accurately reflect the strengths of the preferences of A and B. Hare presupposes that a person's preferences or desires for something have a strength that can be specified independently of the strengths of his other desires. However, to my knowledge, Hare never explains how this might be done.

It is *always* in order to ask whether the archangel's conditional desires and the overall desire or preference that emerges from them accurately represent the strengths of the desires of the other individuals in question. To suppose that a particular overall preference is the on balance preference that accurately reflects the strengths of the desires of different individuals is to presuppose

that there is an *independent* basis or ground on which it can be said that one desire is stronger than another. Without some such independent standard to judge the strengths of the desires of different individuals, it makes no sense to say that the overall preference of the archangel or ideal observer does or does not accurately reflect the strengths of the preferences of the different individuals. This shows that Hare cannot analyze or make sense of the notion of an inter-personal comparison of the strengths of desires in terms of the notion of an intra-personal comparison of the strengths of desires. The comparisons of the latter sort which he makes use of cannot make sense, unless there are independent grounds (at least in principle) in terms of which we can make sense of comparisons of the former sort.

On pp. 128–129 of *MT* Hare considers the question of whether he should give a bun to Smith or to Jones. Both want the bun and, since he has only one bun, he can't give it to both of them. Hare uses this case to illustrate how his theory can reduce the comparison of the strengths of Smith and Jones' preferences to a comparison of the strengths of his own preferences. However, in Hare's description of how he would represent their preferences to himself he presupposes that Jones has a stronger preference for the bun than Smith:

I have to give a bun to either Jones or to Smith, and Jones wants it more than Smith (*MT*, p. 128).

Assuming that we can rank their two preferences in order of strength, then a certain set of conditional preferences that one has may be said to accurately reflect the strengths of their respective preferences. In this case, in order for Hare's conditional preferences to accurately reflect the strengths of the preferences of Smith and Jones, he must have an on balance preference that Jones get the bun.

Clearly, given the *assumed* (my emphasis) strengths of Smith's and Jones' preferences, my own preference will be for the first of these alternatives, and I shall, in accordance with the method of Ch. 6, say that that is the one I ought to choose (p. 129 *MT*).

But all of this presupposes that interpersonal comparisons of the strengths of the preferences of different individuals can be made independently of the sorts of intrapersonal comparisons that an archangel makes.

It would be well to conclude by noting that this is hardly a problem for Hare alone. Any plausible moral theory must contain a requirement of benevolence. Therefore, any plausible moral theory will require that one be

able to compare the strengths of the desires and interests of different individuals (see *MT*, p. 118).

ACKNOWLEDGEMENT

Portions of this paper are derived from the final appendix to my book *The Status of Morality*, (Reidel, *Philosophical Studies Series*, 1984). I am indebted to Professor Hare for helpful criticisms and comments on an earlier version of this paper. I also wish to thank Judy Carson, Harry Gensler, Mark Overvold, and my colleagues at Loyola University to whom an earlier version of this paper was presented.

NOTES

¹ *Moral Thinking* (Oxford, 1981), p. 46. Hereafter, I shall refer to this and other of Hare's works with the following abbreviations: *MT* = *Moral Thinking*. *F & R* = *Freedom and Reason*, (Oxford, 1963); and *LM* = *The Language of Morals*, (Oxford, 1952).

² This is a bad misreading of Hare. I am guilty of this misreading myself in the final appendix to my book *The Status of Morality*. Among other things, the requirement under consideration has the consequence that it would be impossible for an archangel (or anyone else) to represent adequately the experiences of two or more people who have *any* incompatible preferences. For it is impossible for one to have two or more incompatible sets of preferences simultaneously. For all practical purposes, the present requirement would mean that an ideal observer could never represent adequately the experiences of more than one person at a time. For it is unlikely that there are any two people who have no incompatible desires.

³ 'Some reasoning about preferences: A note on papers by I. Persson, F. Feldman, and Scheuler', *Ethics*, October 1984, pp. 82–83. (Also see, *MT*, pp. 223–225.)

⁴ To my knowledge, Hare never explicitly realizes the need for anything like premise (7). It might be thought that (7) follows from the conjunction of (1) and (6). However, this is not the case. Suppose that I make a moral judgment to the effect that S ought to do *x* at *t*. According to (1), I must accept the judgment that I ought to do *x* at *t* if I were in S's position. Thus, the conjunction of (1) and (3) implies that if I accept the judgment that S ought to do *x* at *t*, then I must prefer on balance that *if I am in S's position* that S do *x* at *t*. However, it does not follow that I must now prefer that S do *x* at *t*, if I am not in S's position. Thus, (7) cannot be inferred from (1) and (6) alone. I am inclined to think that (7) is no more problematic an assumption than (6). If this is so, then the introduction of (7) as a needed assumption does not create any further problems for Hare that are not created by (6) alone.

⁵ I suggest the following as an analysis of the concept of a "self-regarding desire":

S's desire or preference that P is self-regarding if, and only if, the object of the desire (P – the desired state of affairs – the state of affairs whose occurrence would "satisfy" the desire) is such that, necessarily, if it occurs at *t*, then S exists at *t*.

So, for example, my desire that I be happy, that I be famous, or that I save your life are

"self-regarding desires". My desire that you be happy and that your life be saved are not self-regarding desires. Alternatively, we could say the following:

S's desire that P is a self-regarding desire if, and only if, necessarily if P obtains at *t*, then S exists *at some time or other*.

Unlike the first analysis, the second analysis has the consequence that one's desire for posthumous fame counts as a self-regarding desire. I owe this notion of a self-regarding desire to Mark Overvold. See his papers 'Self-Interest and the Concept of Self-Sacrifice', *Canadian Journal of Philosophy* 10, March 1980, 'Self-Interest and Getting What You Want', in *The Limits of Utilitarianism*, Harlan B. Miller and William H. Williams eds., (Minnesota, 1982), and 'Morality, Self-Interest, and Reasons for Being Moral', *Philosophy and Phenomenological Research*, June 1984. Using Overvold's notion of a "self-regarding desire", we can define the concept of a "purely self-regarding desire" as follows:

S's desire that P may be said to be purely self-regarding if, and only if:
(i) S's desire that P is self-regarding, and (ii) it is possible that P obtains, even if S is the only person in the universe.

My desire that I be happy and my desire that I acquire certain traits of character are purely self-regarding desires. Such desires as my desire that I be loved by others or that I be the cause of your welfare are not purely self-regarding.

⁶ For a good illustration of the many difficulties created by the phenomenon of changing desires see Brandt's recent paper, 'Two Concepts of Utility', in *The Limits of Utilitarianism*, *op. cit.*, Brandt says that his concerns with this problem led him to reject a desore satisfaction analysis of personal welfare in favor of a hedonistic theory. (Also see *A Theory of the Good and the Right*, pp. 247, ff.)

*Department of Philosophy,
Loyola University of Chicago,
6525 North Sheridan Rd.,
Chicago, IL 60626,
U.S.A.*