J. Value Inquiry 18:69–80 (1984).

Review article

HARE ON MORAL THINKING

D.W. HASLETT
University of Delaware

R.M. Hare, *Moral Thinking: Its Levels, Method and Point* (Oxford: Clarendon Press, 1981. pp. viii + 228, bibliography and index. $19.95 (cloth); $6.95 (paper).)

In the preface to his new book, *Moral Thinking: Its Levels, Method and Point,*[1] R.M. Hare says that, "Nobody can hope to write the last word on a philosophical subject; the most he can do is to advance the discussion..." (p. v). That is no doubt true, but on the philosophical subject of ethics Hare has, in this new book, come closer to saying the last word than he, and perhaps anyone else, ever has. This is a book about how to do our moral thinking more rationally. In it we find much that is of interest and, I think, much that is true about many topics – topics ranging from moral intuitionism to moral fanaticism – but, most importantly, in this book we find a potent moral theory, a moral theory the cornerstones of which are, in my opinion, threefold. The first of these cornerstones is a two level analysis of moral thinking, an analysis which Hare puts to use in answering traditional objections against utilitarianism, and which, incidentally, provides him with the means by which he attempts to reconcile act and rule utilitarianism. The second cornerstone is the generation of his version of the principle of utility from "the logic of moral concepts," an enterprise begun in Hare's earlier works, especially in *Freedom and Reason,*[2] and developed further in this book. The third cornerstone, an enterprise not begun in his earlier works, is a justification of the logic of moral concepts – i.e., of the moral point of view – a justification in terms of prudence. In what follows, I shall concentrate upon each of these three aspects of his theory.[3]

I

The first of Hare's two levels of moral thinking he calls the "intuitive" level. Moral thinking at the intuitive level is done by appeal to prima facie moral principles, ones which are relatively simple, and resemble the sort of moral principles most of us are brought up to observe – principles forbidding such things as promise-breaking, theft, murder, and falsehoods. These are, Hare says, the *only* sort of principles that the moral intuitionist recognizes, and therefore when faced with a conflict between them, or when confronted by someone with intuitions different from his, the intuitionist will be at a loss.

What the intuitionist fails to realize, Hare claims, is that there is another level of moral thinking, one which Hare calls the "critical" level. Moral thinking at the critical level is done by appeal to what is, in effect, the principle of utility, a principle which prescribes that we choose that alternative which maximizes overall preference satisfaction, where equal preferences are to be given equal weight, regardless of whose preferences they might, in reality, happen to be.[4] Hare himself does not speak of thinking at the critical level as proceeding by appeal to the "principle of utility," or any other principle. Rather than referring to what he is proposing for use at the critical level as a "principle," he refers to it as a "method." But he does tell us that the judgments this "method" yields are "the same as a careful act-utilitarian would make" (p. 43), and, of course, we always speak of act

utilitarians as proceeding by appeal to some version of the principle of utility. Therefore, I do not think it misrepresents Hare's view if we, for convenience, speak of what Hare is proposing as a version of the principle of utility also, provided we keep in mind that we are not speaking of the classical version of this principle, but one which prescribes that preference satisfaction be maximized. Furthermore, Hare limits what preferences are to be taken as the relevant ones to preferences for *experiences* (pp. 90–1).[5] And, what is equally significant, only ideal preferences are to count; that is, only preferences as they would be had they been *fully exposed to facts and logic* (pp. 104–5). It is at this critical level of thinking – i.e., the level at which we are to appeal to the principle of utility – where those conflicts which arise between the prima facie principles we use at the intuitive level are to be resolved, and where these prima facie principles are to be chosen in the first place – chosen on the basis of their "acceptance-utility" (p. 50).

What exactly it means to "accept" a moral principle is, unfortunately, left unclarified by Hare, but I assume that the "acceptance" of some principle *P* includes a certain willingness to try to comply with *P*, as well as a certain willingness to help bring about or maintain social pressure for the compliance of others with *P*. It should be noticed that the acts prescribed by prima facie principles with maximum acceptance-utility would not correspond exactly to those that would maximize utility; their prescribing acts corresponding exactly to those that would maximize utility is incompatible with these principles being as *simple and easy to comply with* as they would have to be in order to maximize acceptance-utility.

But if these prima facie principles do not always prescribe the act that maximizes utility, this means that the act prescribed at the intuitive level of thinking does not always correspond to the act that would be prescribed, after full exposure to facts and logic, at the critical level of thinking. This then raises the question of which act, the one prescribed at the intuitive level, or the one prescribed, after full exposure to facts and logic, at the critical level, is definitive of (i.e., in the final analysis, delineates) the morally right act. Hare answers as follows: "[The prima facie principles] are not definitive of 'the right act'; but if we wish to act rightly we shall do well, all the same, to follow them" (p. 30). So if the prima facie principles are not definitive of the right act, we may assume that, for Hare, the definitively right act is the one prescribed, after full exposure to facts and logic, at the critical level; in other words, the one prescribed by the principle of utility.

But if it is the principle of utility, not the prima facie principles, that is definitive of the morally right act, then why, it may be asked, shall we do well, all the same, to follow the prima facie principles? The answer is that, ordinarily, we do not have the time or the facts necessary for determining reliably what the definitively right act is; the prima facie moral principles then tell us what our "best bet" is, given this lack of time and facts – i.e., given human fallibility; they tell us, in other words, what we most likely would have chosen at the critical level if we *had* had sufficient time and facts (pp. 35–39). And since we lack sufficient time and facts for a reliable choice most of the time, the intuitive level of thinking, carried out in terms of these prima facie principles, is the level of moral thinking that is appropriate for most everyday moral decisions. But these prima facie principles are not, Hare insists, mere "rules of thumb." Owing to our upbringing, they are associated with "very firm and deep dispositions and feelings;" thus, unlike mere rules of thumb, a breach of them normally excites in us a feeling of compunction (pp. 38–39).

Having seen how Hare distinguishes between two levels of moral thinking, we can now see why he claims his position reconciles act and rule utilitarianism. According to Hare, both the act utilitarian and the rule ufilitarian are right about one

level of moral thinking: the act utilitarian is right about the critical level; the rule utilitarian is right about the intuitive level (p. 43). We can now also see how, in a delightful piece of philosophical writing that must be read to be fully appreciated (ch. 8), Hare counters the intuitionist's standard argument against utilitarianism. The intuitionist's standard argument is simple: he produces some example where what maximizes utility appears to contravene one of our deeply held moral principles; and since (the intuitionist assumes) contravening this deeply held moral principle is morally wrong, he concludes that utilitarianism must thus be false. In reply Hare, first of all, points out that if the example which the intuitionist produces is a realistic one, one that might well occur under ordinary circumstances, then he (Hare) has no significant disagreement with the intuitionist about what moral principles should be used to decide the case: intuitive level, prima facie moral principles should be used, principles that, for all practical purposes, correspond to the very ones the intuitionist is claiming should be used. Intuitive level principles should be used – and used even though what they prescribe might not, in the case at hand, maximize utility – because, in ordinary cases of the sort in question, they are, given our normal lack of time and facts, our "best bet." In order to provide an example where, on Hare's view, we must resort to critical thinking, and thus to the principle of utility itself rather than these prima facie moral principles, the circumstances in which the example is couched would have to be so extraordinary, the case so fantastic, that even the intuitionist would have to doubt the appropriateness of using the prima facie principles to decide it, principles designed only for deciding ordinary, everyday cases. But then if we have doubts about the appropriateness of using the prima facie principles, the fact that the act which maximizes utility contravenes these principles is no longer a very compelling reason for rejecting utilitarianism.

II

Hare's two level analysis of moral thinking amounts to what is probably the strongest defense yet of a utilitarian position, but I shall now argue that this defense does not quite succeed. Yet it comes so close to succeeding that, if I am right, all that would be necessary for success is a minor change in Hare's position – or what would appear to be a minor change but, as we shall see, perhaps is not. Hare's defense of utilitarianism in terms of a two level analysis of moral thinking does not succeed because, in the final analysis, the intuitive level of moral thinking cannot be kept distinct from the critical level; it "collapses" into the critical level leaving Hare not with a reconciliation between act and rule utilitarianism, but with just act utilitarianism. The problem is not that the prima facie principles to be used at the intuitive level end up prescribing exactly the same acts as the principle of utility to be used at the critical level; as we saw, they would not. Rather, the problem lies in determining, for whatever case is at hand, which level of moral thinking, the intuitive or the critical, we should be using – that is, in determining whether we are pressed enough for time, or whether the case before us is "ordinary" enough, for us to invoke the prima facie principles, or whether we should go directly to the principle of utility instead. There are only four alternative means of determining this: (1) by means of some third level of moral thinking more basic than either the critical or intuitive levels; (2) by means of the critical level; (3) by means of the intuitive level; and (4) by means of no standard at all – i.e., arbitrarily. Hare does not provide us with any third level of moral thinking, and surely such decisions should not be made arbitrarily, which leaves only alternatives (2) and (3). Hare does say that, with any deep experience in moral thinking, a person will

have acquired some *methodological* prima facie principles which will tell him when to use critical thinking and when to use intuitive thinking (p. 52); so perhaps the answer is that these decisions should be made at the intuitive level, by means of methodological prima facie principles. Methodological prima facie principles may help – depending upon how simple these methodological principles themselves are – but they cannot be the whole answer. For one thing, we are still left with the question "How are we to determine what these methodological prima facie principles should be, and when they should be invoked?" Hare's answer is that we are to determine these matters by means of critical thinking (p. 50). It is then critical thinking – i.e., the principle of utility – which is the *final* arbiter for when we are to use critical thinking and when not; that is to say, it is the final arbiter as far as *morality itself* is concerned. Therefore, in every case when we do use intuitive thinking, we are, as far as morality itself is concerned, to do so only because this is how the *principle of utility* tells us to proceed. It is in this sense that the intuitive level collapses into the critical level, leaving us, I suggest, not with a reconciliation of act and rule utilitarianis, but with just act utilitarianism – a very sophisticated act utilitarianism to be sure, but act utilitarianism nevertheless. So it must be concluded that, in the final analysis, the code of morality that Hare supports is an act utilitarian code of morality after all, a code of morality in which the principle of utility is fundamental. It is true that Hare himself never says that he supports an act utilitarian *code of morality*, as opposed to an act utilitarian *method of moral thinking*, but if my arguments so far are sound then what Hare does say commits him, I claim, to supporting such a code nevertheless (if, in fact, the distinction between an act utilitarian "code of morality" and an act utilitarian "method of moral thinking" is a meaningful one to begin with).

Hare might reply that, although perhaps his view does, in this sense, collapse into act utilitarianism, he is not particularly bothered by this; he is not particularly bothered since his defense of utilitarianism against the intuitionist's argument remains sound. That is to say, a perfect practitioner of his method of moral thinking or code of morality (whichever you may want to call it) would still always choose, at the critical level, to think at the intuitive level, unless confronted by a genuinely fantastic case, and so forth.

This may well be so. The only problem is that none of us is, or ever could be, perfect practitioners of this method, or even come close to being perfect practitioners. To be perfect practitioners of this method we would, among other things, have to be able to do perfect critical thinking about what our prima facie principles should be, and when they should be invoked. But, as Hare himself admits (p. 45), perfect critical thinking – and I see no reason why this should not include critical thinking about what our prima facie principles should be, and when they should be invoked – requires "superhuman powers of thought, superhuman knowledge, and no human weaknesses" (p. 44). Critical thinking is, in other words, very difficult indeed. This raises an objection to Hare's utilitarian position that is, I think, even more serious than the one raised by the intuitionist. Since critical thinking is so very difficult, people's *interpretations* of what is prescribed at the critical level are not likely to be very uniform. In particular, people's interpretations of what prima facie principles are prescribed, and of when to invoke them, are not likely to be very uniform. After all, decisions at the critical level on these matters, just as on other matters, must sometimes be made quickly, must often be made with a scarcity of facts, and must always be made by us very fallible human beings. But if these interpretations were not very uniform, then people's moral behavior, stemming as it would from these interpretations, would not be very uniform either. And finally, without much unformity, and thus predictability, in people's moral behavior, there would not be much security or mutual trust; there would, in

short, be an intolerable degree of chaos.[6]

Hare might reply by pointing out that so far I am overlooking one very important point: the prima facie principles, as we saw, are not exactly mere rules of thumb, but are backed by social pressure, social pressure which creates in people deep feelings of attachment to these principles, feelings which generally give rise to remorse or guilt on the part of those violating them. Hare might then conclude that, because of this social pressure, and thus these deep feelings of attachment, the "chaos" referred to above would never materialize; the social pressure would be sufficient to maintain adequate uniformity in people's moral behavior in spite of differing interpretations of what is prescribed at the critical level, just as, in most societies, social pressure manages to maintain a tolerable degree of uniformity in the face of differing moral opinions.

This reply might well, once again, be good enough, except for one remaining problem: although social pressure is indeed sufficient for maintaing tolerable uniformity in most societies, in the moral code of most societies the principle of utility is not fundamental while, as we have seen, in the moral code Hare supports, it is; and this is crucial. It is crucial because, with a moral code where the principle of utility is fundamental (i.e., where conformity to it, and it alone, is definitive of moral rightness) the requisite social pressure for compliance with subsidiary prima facie principles could never be generated. That is, it could never be generated, assuming only that the code recognizes, as it should, excuses for innocent mistakes of fact. Let me explain. In every code of morality of which I am aware, people are excused from failing to adhere to the moral principles within the code if their failure is attributable not to improper motivation, but to an innocent mistake of fact, one that anyone might have made. Say, for example, you have no reason whatever to know that someone has by mistake put arsenic into your sugar bowl and that you, therefore, serve some unfortunate person poisoned tea; your innocent mistake of fact will excuse you from any blame for having poisoned him. The recognition of such an excuse is, I submit, crucial, if for no other reason than for people's peace of mind. Were it not for the recognition of such an excuse people would remain in constant jeopardy of severe moral condemnation for that over which they have virtually no control – namely, innocent mistakes of fact – a situation hardly conducive to the very peace of mind for which we have a morality in the first place. But it is precisely because of this excuse for innocent mistakes of fact that adequate social pressure for compliance with simple, prima facie moral principles could never materialize with a moral code in which the principle of utility is fundamental. People with false opinions about *which* prima facie principles the principle of utility requires compliance with, and *when* exactly this compliance is required, could never legitimately be held morally blameworthy for behaving in accordance with these false opinions. They, as well as anyone else, would be doing their best to comply with the code's overriding moral duty to maximize expected utility; thus, if, in adhering to the wrong prima facie principles, or in adhering to the right ones but at the wrong times, they did not succeed in maximizing expected utility, their failure would have to be attributed not to improper motivation, but to a mistake of fact as to what the principle of utility requires. And, given the extreme difficulty there normally is in determining what the principle of utility requires, if any mistake of fact should excuse one from blameworthiness, then surely a mistake about this matter should. Yet it is (roughly speaking) only through holding people morally blameworthy for certain acts and not for others that adequate social pressure for compliance with moral principles can be generated. So if, because of their mistakes of fact, people could not normally be held morally blameworthy for failing to comply with the principle of utility, then adequate social pressure for compliance with it and, *a fortiori*, with

any prima facie principles adherence to which is said to be required by it, could never be generated. And, finally, if adequate social pressure for compliance with any such prima facie principles could never be generated then, with a code in which the principle of utility is fundamental, adequate uniformity of behavior would never materialize.

But am I not, it might now be asked, creating a lot of fuss about what is only a minor aspect of Hare's position, an aspect that could easily by changed? Hare could, after all, avoid my objections with a simple change in the status that the principle of utility has in his moral philosophy, a change in its status from that of a fundamental *moral* principle to be used, among other things, for choosing moral principles to that of a *non*-moral principle[7] to be used for choosing moral principles. This change in status would not affect *what* moral principles were chosen; those chosen would no doubt be the same simple, prima facie moral principles that Hare thinks would be chosen. We might even find that among the moral principles within the code of morality chosen by appeal to the principle of utility would be the principle of utility itself; but if so, then, for the reasons already given, its role *within* the code would not be that of a moral principle that was fundamental. It would have a much more modest role, such as that of being a principle of last resort, a principle to be used *only* in conflict and borderline cases after appeal to all the other moral principles within the code had been exhausted. If the principle of utility did have this role within the moral code, then compliance with it would be definitive of the morally right act only in those (relatively rare) cases that could not be resolved by appeal to any of the other moral principles of the code. In all other cases it would be compliance with the other moral principles – i.e., the simple, prima facie moral principles – that would be definitive of a right act.

How this relatively simple change in the principle of utility's status – a change from that of a fundamental moral principle to that of a non-moral principle having much the same function – would make my objections inapplicable should be apparent. With the principle of utility not being morally fundamental, an innocent mistake of fact about what it calls for could not be used as an excuse for failing to comply with prima facie principles. And without this standing excuse for failing to comply with prima facie principles, adequate social pressure for compliance with them could then be generated. Finally, if adequate social pressure for compliance with them were generated, then (in view of how easy it is to *succeed* in complying with such principles) there would be adequate uniformity in people's moral behavior.

Of course if Hare were to change the status of the principle of utility from that of a fundamental moral principle to that of a non-moral principle for choosing moral principles, he would thereby become a *rule* utilitarian, and have to abandon the claim to be reconciling act and rule utilitarianism. But if this were *all* he would have to abandon, perhaps it would be a small price to pay, especially since, if I am right about his position as it now stands collapsing into act utilitarianism, he must abandon this claim anyway. But I doubt if this claim *is* all that he would have to abandon; he would probably also have to abandon the attempt to generate[8] his version of the principle of utility from the logic of moral concepts – that which I referred to earlier as the second cornerstone of Hare's position. To see why he would probably have to abandon this attempt, let us now look briefly at how his argument proceeds.

Hare's first step in generating this principle from the logic of moral concepts is to show that, if any one person, say Jones, is fully informed about the fully informed preferences of any other person, say Smith, then Jones will acquire preferences corresponding to Smith's regarding what should happen to Jones *if Jones*

were in Smith's situation. It is in the next step in the argument that Hare makes use of what he sees as the logical properties of moral concepts: prescriptivity and universalizability. Moral principles are, Hare says, prescriptions, and these prescriptions must be ones that are universalizable, universalizable in the sense that if, say, Jones prescribes that a certain thing be done to Smith, Jones is committed to prescribing that the very same thing be done to him, were he in exactly Smith's situation. In other words, moral principles must be formulated only in terms of *universal properties*, not in terms of *proper names*; as Hare says, "morality admits no relevant difference between 'I' and 'he'..." (p. 223). Hare then goes on to conclude that, therefore, Jones' (and anyone else's) fundamental moral principle must be indifferent between the satisfaction of his preferences and Smith's preferences, which means that it must prescribe only that preference satisfaction be maximized, *no matter whose preferences these are* – in which case we have thus succeeded in generating Hare's version of the principle of utility from the logic of moral concepts (sec. 6.2). And then, of course, from the use of the principle of utility at the critical level we go on, according to Hare, to derive more specific, prima facie moral principles to be used at the intuitive level.

We are now in a position to see why, even if necessary for avoiding my objections, Hare might be reluctant to change the status of the principle of utility from that of a moral principle to be used for deriving moral principles to that of a non-moral principle to be used for deriving moral principles. To do so would threaten his generation of the principle of utility from the logic of moral concepts since, after this change in status, the principle of utility, no longer a *moral* principle, would no longer qualify as one to which the logic of *moral* concepts is applicable. (What I mean, of course, is that it no longer would qualify when playing the role of a principle for deriving moral principles, as opposed to the role of principle of last resort *within* the moral code, as explained earlier.) So it seems that only by insisting upon the principle of utility being a fundamental principle of *morality* can Hare protect his argument. One question, however, remains: is Hare's argument successful, and thus entitled to this protection?

III

Before trying to answer this question, a somewhat embarrassing problem which Hare's argument creates for him should be noted. No one has done more to discredit descriptivism as a theory of moral language than Hare. According to descriptivism, merely from the meanings or logic of moral concepts, we can derive substantive moral principles and judgments. But by generating his version of the principle of utility from the logic of moral concepts, it appears as if Hare is himself subscribing to a version of the very theory he has been so successful in discrediting.

Hare fully recognizes the dilemma he seems to have constructed for himself, and has left himself several ways of escaping from it. What is, I think, the most important of these ways is the following: although Hare does say that, once one adopts the logic of moral concepts, then one is logically compelled by these concepts to adopt the principle of utility too, he also says that one need not adopt the logic of moral concepts; in other words. one need not play the morality "language game" in the first place. One can, instead, be a consistent "amoralist," one who either refuses to make any moral judgments at all, or views all such judgments with indifference. So if, on Hare's view, one does choose to play the morality "language game," – that is, to adopt the "moral point of view" – one's choice is made independently of any logical or linguistic constraints whatever which, Hare concludes, is enough to preserve his status as a bona fide non-descriptivist after all (sec. 10.8).

But this absence of any logical or linguistic constraints upon our choice of whether or not to adopt the moral point of view raises still another problem for Hare: why then, it might be asked, *should* we adopt the moral point of view anyway? Hare has already granted that we cannot appeal to reasons of logic for adopting the moral point of view, and obviously we cannot appel to *moral* reasons, for that would be blatantly circular. Hare's solution is that we can appeal to prudential reasons, or reasons of self-interest, for doing so – which brings us, finally, to what I referred to earlier as the third cornerstone of Hare's position. Hare admits that it is impossible to show that acting morally is *always* in one's best interests. But, he says, if we were educating a child we should, for the child's own sake, nevertheless advise him to adopt the moral point of view *altogether*, rather than the policy of trying to pick and choose when to be moral; that is, rather than the policy of being moral only when one thinks that, in the case at hand, it is in one's best interests to be moral. Why? Here, again, we must read Hare himself to appreciate his answer (ch. 11), but, very briefly, it is as follows. In trying to pick and choose when to be moral, we fallible human beings, subject as we are to misinformation and rationalizations, are likely to miscalculate the benefits for us of immorality often enough so that, in the long run, we would have been better off following a policy of consistent morality instead. As Hare points out:

> ...successful crime is for nearly everyone an impossibly difficult game and not worth the candle. If it is alleged against this that in the past people have amassed large fortunes in business careers which were far from unspotted, I reply that the money did not on the whole bring them happiness, and that with their talents they could have done better for themselves by making less money in a more socially beneficial career. If there are exceptions, they are rare enough to be unpredictable by an educator (p. 196).

Therefore, Hare concludes, it is in a child's own best interests to educate him so that he has a firm disposition to act morally, so that, in other words, he adopts the moral point of view altogether, rather than the foolish policy of trying, in these matters, to calculate prudential costs and benefits on a case by case basis. And, finally, if it is in the child's best interests to adopt the moral point of view (i.e., the morality language game), then it should be, for the same reasons, in the best interests of the rest of us as well. Thus, for Hare, the logic of moral concepts carries us a long way toward the justification of his position, and where logic leaves off, as it inevitably must, prudence enters to complete the justification. So the justification of Hare's position consists, in effect, of two components a logical component, and a prudential component.

IV

How adequate is this two component approach to justification? To answer this, let us return now to the question posed earlier about the logical component – about, that is, the generation of the principle of utility from the logic of moral concepts: is this argument successful? As will be recalled, the first step in the argument is to show that if we are fully informed about a fully informed preference of another – say, Smith's preference for experiencing A rather than B – then we shall (for the hypothetical case in which we are in the other's position) acquire a preference corresponding to his. I have no reservations whatever about accepting this claim. But surely it is perfectly rational (i.e., it violates no constraints of logic) to allow preferences for experiences to influence our behavior or choice of prin-

ciples only in so far as we believe our behavior or principles will have some bearing upon whether we shall actually *have* the experiences in question. For example, the atheist will no doubt prefer the eternal bliss of heaven over the eternal torment of hell every bit as strongly as does the religious fundamentalist. But since the atheist does not believe that any principle he chooses will have any bearing upon whether he will actually *have* these experiences, surely it it perfectly rational for him not to allow this preference, as strong as it may be, to influence his choice. Likewise, just because we prefer experiencing A to B just as strongly as does Smith, who (as opposed to us) is actually going to experience either A or B, it does not follow that we must allow this preference to influence our choice of what principles to accept. And, just because we prefer experiencing those experiences resulting from adherence to the principle of utility to those resulting from adherence to any alternative, we need not (logically speaking) allow this preference to influence our choice either – especially since most of those experiences resulting from adherence to the principle of utility will be ones experienced not by ourselves, but by others. Competing with the principle of utility for our acceptance will be other principles which, if adhered to, might be more likely than the principle of utility to satisfy our preferences for experiences we might actually have ourselves. If, for example, we are white rather than, say, black, such a principle might be one which prescribes that we always give *special* consideration to the interests of whites. So with this first step in his argument, Hare has not given us a reason, logical or otherwise, why we must accept the principles of utility rather than this discriminatory one. He has, so far, given us no reason because, once again, our choice need not be determined by preferences for experiences which, no matter what, we are unlikely ever to experience ourselves.

But, if I understand Hare adequately (and on this matter, I must admit, I am not sure I do), it is here where the logic of moral concepts is supposed to enter the picture – in the form of universalizability – and carry the argument to a successful completion. Universalizability, it will be recalled, is a logical constraint upon the formulation of moral principles, this constraint being that they must be formulated only in terms of universal properties, not proper names. But this constraint does not, as Hare himself admits (p. 64), prevent our "rigging" these principles in our favor; that is, formulating them in terms of just those universal properties that happen to be such that the resulting principles, if successfully followed, will satisfy only those of our preferences with respect to experiences that we are likely to have ourselves. This then still would allow us, logically speaking, to accept the discriminatory principle rather than the principle of utility. In other words, if Hare has not succeeded in generating the principle of utility *before* universalizability enters the picture, I do not see how universalizability can, by itself, possibly "turn the tide." Hare does point out, correctly, that universalizability implies that principles must be just as applicable to hypothetical cases as to real ones (pp. 112 ff.), hypothetical cases where we ourselves are in the positions of others; but this does not meet my objection. I am not claiming that (logically speaking) we need not view the principles we accept as applicable to hypothetical cases; I am claiming instead that (logically speaking) we need not take hypothetical cases as seriously as real ones in deciding whether to accept the principles in the first place. Thus I am forced to conclude that Hare's argument does not succeed.

Although my rather brief criticism of Hare's attempt to generate the principle of utility from the logic of moral concepts may be far from conclusive, I wonder, for other reasons, whether his continuing this attempt is really in the best interests of his overall position. For consider the advantages of his abandoning the attempt. First, his abandoning the attempt has the advantage of establishing, once and for all, his credentials as a bona fide non-descriptivist; and, incidentally, leaves his

classic analysis of moral concepts, set out in *The Language of Morals*,[9] fully intact. Secondly, and more importantly, by abandoning the attempt Hare would no longer need to insist upon the principle of utility being a fundamental *moral* principle; he could, as explained, view it instead as a *non*-moral principle with much the same function. This would have the advantage of avoiding the objections I raised earlier against a code of morality in which the principle of utility is fundamental. And, by avoiding these objections, Hare's ingenious defense of utilitarianism against intuitionist objections would then be on firmer ground.

But if Hare were to abandon his attempt to generate the principle of utility from the logic of moral concepts then what, it might now be asked, would become of his justification for the principle of utility – a justification which we have seen consists, in effect, of two components: a logical component, and a prudential component? If Hare were to abandon this attempt, he would, of course, be left with only the prudential component. But would being left with only this component really matter? Perhaps not; let me conclude by briefly suggesting why. First, suppose that, in the final analysis, the logical component of Hare's approach to justification does prove successful; what will he really have accomplished by it? He will have shown that the *logic* of the morality language game is such that either we must play it his way (i.e., by means of the principle of utility), or else we must decide what we shall do without reference to any of the concepts of this language game – concepts such as "should," "ought," and so forth. Will the opponents of the principle of utility find this a convincing reason for abandoning their opposition to it and becoming utilitarians? I doubt it. Surely the opponents of the principle of utility will instead say something like this:

> So much the worse for the morality language game. Let us now replace it with, say, the "smorality" language game, a language game which is played with the concepts of, say, "smould," "smought," and so forth – concepts which, logically speaking, *do* allow us to accept that principle which, we think, is more justified than the principle of utility; more justified, that is, for *substantive*, not merely *formal*, reasons.

In other words, if, in the final analysis, the logical component does prove successful, the only thing Hare will have accomplished by it will be to have forced us to pose the question of justification somewhat differently than we do now; he will have forced us to ask not how we are to justify our *principles*, but how we are to justify our *language games*. So we will, in reality, be no closer to answering the question of justification than we ever were; we simply will be forced to pose it in a less familiar context. Next, consider what reasons might plausibly be given in justification of our accepting the morality language game. Obviously these could not be reasons of morality itself, for this would be to argue in a circle. In short, it is hard to see what these reasons might possibly be other than reasons of prudence, which are, as we have seen, exactly the sort of reasons Hare does give. So Hare's two component approach to justification comes down to the following: he argues that we have good *prudential* reasons for accepting the morality language game, and that if we accept the morality language game, then we have good logical reasons for accepting the principle of utility. Thus, for Hare, to have good prudential reasons for accepting the morality language game is, *in effect*, to have good prudential reasons for accepting the principle of utility. But if, in effect, we have good prudential reasons for accepting the principle of utility then why, it might be asked, do we need the logical reasons as well? Could we not just go directly from the prudential reasons to the principle of utility, and skip the logical reasons? I suggest we could. I suggest, in other words, that if the prudential component of

Hare's approach to justification is sound, then the logical component is not even necessary. So if, in the end, Hare must abandon the logical component of his approach to justification, his overall position will not, I suggest, be weakened; on the contrary, for the reasons already mentioned, it may be strengthened. But if, on the other hand, he must abandon the *prudential* component, then the loss will be serious indeed.

Is then the prudential component of his approach to justification successful? Since his arguments on this matter do depend upon empirical premises about which Hare, by his own admission (p. 192), cannot be certain, his arguments are far from conclusive. I happen to believe, however, that, with this component of his approach, he is on the right track, and that, here as elsewhere, although he has not said the last word, he may, as I suggested before, have come closer than anyone else.[10]

NOTES

1. R.M. Hare, *Moral Thinking* (Oxford: Clarendon Press, 1981). All references throughout that are in parentheses are to *Moral Thinking*.
2. R.M. Hare, *Freedom and Reason* (Oxford: Clarendon Press, 1964).
3. A fourth cornerstone of Hare's theory might be what he says in Chapter 5 (perhaps the most difficult chapter in the book) about why, if we are fully informed about the fully informed preferences of another, we shall, for the hypothetical case in which we are in his place, have the same preferences as he does. But what Hare says in Chapter 5 is, as we shall see, one step in his generation of the principle of utility from the logic of moral concepts, this generation clearly being one of the cornerstones of his theory. Thus I am inclined to view what Hare says in Chapter 5 not as itself one of the cornerstones, but as a step in the establishment of one of them.
4. Hare also mentions that critical thinking can be given not only a utilitarian interpretation, but a Kantian interpretation as well, an interpretation where critical thinking is viewed as an appeal to Kant's categorical imperative (p. 50). Thus Hare views his position as a reconciliation of utilitarianism and Kantianism. The utilitarian interpretation, however, seems to be less problematic; therefore it is the utilitarian interpretation that I shall emphasize, as does Hare himself.
5. Focusing upon preferences for experiences only is, for Hare, merely a simplifying device; he believes that a more full account, which he says he has so far been unable to work out, would treat *all* preferences as, to some degree, relevant (pp. 103–104). I happen to believe that, by focusing upon preferences for experiences only, Hare has got it right already, and that a more full account would only be a step backwards, but I shall not be able to pursue this matter here.
6. For a more full account of the need for uniformity, and why a moral code in which the principle of utility is fundamental would not generate it, see G.J. Warnock, *The Object of Morality* (London: Methuen, 1971), pp. 31–34.
7. If not a moral principle then, it might be asked, what kind of a principle is it to be? There are a number of possibilities. If, for example, it could be shown that our using the principle of utility for choosing moral principles is, generally, in our best interests, then we could call it (when it is being used for choosing moral principles) a principle of "prudence." Indeed, a principle of prudence is exactly what it would be if my brief suggestion below about how the prudential component of Hare's justification for the principle of utility can stand alone is sound. But the question of what kind of principle, if not a moral principle, it is to be, although important, need not be answered here.
8. Hare refers to what he is doing as "generating," not "deriving," the principle of utility from the logic of moral concepts (see, e.g., p. 111), because, as we shall see, he does not

take his argument to be a "derivation" in the strict sense – i.e., in the sense this term for logicians.

9. R.M. Hare, *The Language of Morals* (Oxford: Clarendon Press, 1952).
10. I cannot do more here than simply state my *belief* that, with a prudential approac justification, one is on the right track. I do, however, attempt to provide in detai *arguments* for this belief in a book, soon to be completed, entitled *Equal Consideratio*