

# Moral Rules, Utilitarianism and Schizophrenic Moral Education

KEVIN McDONOUGH

**ABSTRACT** *R. M. Hare has argued for and defended a 'two-level', view of moral agency. He argues that moral agents ought to rely on the rules of 'intuitive moral thinking' for their 'everyday' moral judgments. When these rules conflict or when we do not have a rule at hand, we ought to ascend to the act-utilitarian, 'critical' level of moral thinking. I argue that since the rules at the intuitive level of moral thinking necessarily conflict much more often than Hare supposes, and since we often do not have ready-made rules for our moral judgments, we must necessarily use critical moral thinking very frequently. However, act-utilitarian judgements at this level will sharply conflict with our strongly held 'intuitive' moral convictions. I show that Hare's attempt to balance these two aspects of moral judgment requires us to simultaneously adopt two conflicting sets of moral standards, and thus an attempt to inculcate such standards constitutes a 'schizophrenic' moral education. Finally, I briefly outline an alternative conception of moral education, based on Aristotelian phronesis.*

## Introduction

In this paper I discuss the grievous consequences of a view of moral education based on a strong view of moral rules in moral agency. Joel Kupperman has defended a view of moral education that would give a central role to the inculcation of strong 'inhibitions' against breaking rules such as 'Don't tell lies' (Kupperman, 1978). Kupperman argues that moral education has as one of its primary tasks the implanting of such rules and strong affective accompaniments to the rules, so that agents will not feel impelled to violate them, at least under normal circumstances. Thus, rules represent ongoing constraints on even mature moral agents. In this paper I will discuss and criticize a more sophisticated version of what I term the 'strong view' of moral rules. R. M. Hare has recently advocated (Hare 1981) and vigorously defended (1988 a-d) a view of moral agency that requires agents to rely on 'relatively simple *prima facie* principles' (or PFP's) for 'everyday' moral judgments. I argue that this view represents a naive oversimplification of moral agency. Since one important educational implication of this view would be the inculcation of PFP's, the view represents an objectionable moral miseducation. I briefly outline, at the end of the paper, a more desirable view of moral agency based on Aristotle's ethics. Aristotle considers rules to be like moral crutches. They are needed for agents to take their first halting moral steps, but are ultimately dispensable for agents who have developed mature moral judgement. Moral education, then, will seek to cultivate such judgment, or what Aristotle calls *phronesis*.

The issue of moral rules is of obvious practical concern for educators. Rules might be thought to represent the goals, the content, of moral education. For example, people who habitually lie, steal, hurt others, neglect themselves, etc., are as obvious cases of moral failure as we can find. Children who adhere to rules against lying, etc., will not be moral failures in this sense. Nevertheless, lists of general rules (e.g. 'Don't

tell lies' 'Don't steal' 'Be nice to others' etc.) inevitably neglect the complexity of moral life. Lies must sometimes be told. Justice may require that property rights be overridden. And it is not always best to be nice, even to one's friends.

These reflections point to an educational concern with children's ability to make sophisticated moral judgments in a complex, rapidly changing world. The complexity of moral judgment is seen to arise from the interplay of two aspects of judgment. First, principles represent moral standards that are not simply personal. In ethical life, they are tools by which different agents can impartially determine the adequacy of judgments. However, we do not often expect principles to determine their own application. Sensitivity to the particularity of situations, including aspects particular to the individual agents involved or to the salient human relationships involved, is needed to determine how a principle is to be applied in any particular case.

Some earlier conceptions of moral education have tended to turn a blind eye to the issue of moral judgment. For example, in the United States and Canada, the influential values clarification movement claimed to provide a procedure by which children can learn to determine their 'real' beliefs, and to act on those beliefs without internal conflict and with enthusiasm (Raths *et al.*, 1978). This method is in large part a response to alleged dogmatism and rigidity of traditional views of moral education, but in responding to such dogmatism, values clarification has reduced moral education to the mere entrenchment of arbitrary personal preferences and desires. These preferences cannot, for values clarifiers, be justified or condemned according to any interpersonal standards or criteria. Thus, a successful application of this method could generate 'morally' educated agents as diverse as Mother Teresa and Charles Manson. Lawrence Kohlberg's views have similarly been widely criticized for allowing objectionable forms of moral relativism. Hare's view offers a sophisticated, attractive, though ultimately seriously flawed, response to the weaknesses of views such as Kohlberg's and that of values clarification.

Hare's is a 'two-level' view of moral thinking. The distinction between levels, he argues, provides a rational method for establishing the moral rules on which individuals can justifiably rely for much of their everyday moral thinking. These moral rules, the application of which, for Hare, constitutes the 'intuitive' level of moral thinking, make up the moral 'content' for all individual moral agents. They exist at what Hare calls the level of intuitive moral thinking (IMT). These rules are constructed, revised and justified at the level of critical moral thinking (CMT), by a process of impartial evaluation of preferences. If Hare's arguments for his two-level theory are persuasive, he has helped to resolve important weaknesses in previous accounts of moral education. First, since IMT contains the rules by which most of us will lead a large part of our moral lives, Hare avoids the objectionable radical moral relativism often thought to accompany Kohlberg's approach and the values clarification approach. In addition, educators will have a relatively clear idea of the most appropriate moral educational aims, since children will need to learn to obey the rules when that is necessary. Furthermore, since the rules are impartially justified by CMT, Hare's view seemingly avoids charges of moral dogmatism. And, to the extent that one of the educational goals of Hare's view is to cultivate an ability to do CMT when that is appropriate, the rules are not imposed in an authoritarian way. After agents achieve competency in CMT, they will be able to select, revise, and suit moral judgments to their particular circumstances when that is necessary.

In what follows I argue that the alleged advantages of Hare's view are illusory and that his view of moral thinking itself constitutes a grievous moral miseducation. Hare has recently re-affirmed the primacy of IMT in our everyday judgments, since "moral decision-making by humans in the face of real-life risk and uncertainty... should normally be done at the intuitive level" (Hare, 1988b, p. 241). He even goes so far as to suggest that the archangel—his representative of CMT, and who is also a "wise act-

utilitarian” (Hare 1981, p. 38)—is not relevant for studying this kind of ‘real-life’ moral thinking, but only as a “device for studying critical thinking” (1988b, p. 241). Thus, CMT is to be used rarely, or never, at least by humans, for reasons to be explained below. Our moral character, then, is largely constituted by the rules we hold at the intuitive level and our allegiance in adhering to them. At this level, the rules provide a determinate method for judging right acts. However, as I will argue, the complexity of moral life makes this view of moral judgment grossly oversimplified. Our moral rules will necessarily often conflict, or will not serve for novel or unusual moral cases. For Hare there is no further method of judgment at the level of IMT. Thus, on his view, in cases where our rules conflict or where we have no rules, we are forced to resort to act-utilitarian CMT. If, in fact, CMT must be used more often than Hare usually indicates, I argue that his view of moral judgment suffers from common and crippling objections made against act-utilitarian theories. Thus, Hare is faced with a dilemma. On one horn of the dilemma moral judgment (on his view) is largely a matter of applying rules, in which case it fails to account for the complexity of human judgment. On the other horn, moral judgment represents an abhorrent act-utilitarian ideal. The apparent advantages of this view for moral education are thus illusory. I argue that a moral education based on Hare’s two-level theory would induce a paralyzing ‘moral schizophrenia’ since in complex moral cases, the standards of the intuitive level inevitably conflict, in a psychologically debilitating way, with those at the critical level. Hare offers no way for agents to resolve this conflict.

### **Intuitive Moral Thinking: Hare’s strong view of moral rules**

In the latest development of his universal-prescriptivist theory, Hare has claimed that he provides a ‘two-level’ structure and determinate method of moral thinking that is tantamount to what a ‘wise act-utilitarian’ would use (1981, p. 38). The theory is allegedly impervious to ‘vulgar objections’ often made against act-utilitarianism since our strongest common sense moral judgments cannot, except in extremely unusual circumstances, be sacrificed, as with other versions of utilitarianism, in favour of utility maximizing judgments.

Hare claims that we can rely on a set of relatively simple, general principles for most cases in which moral judgment is required. These are represented at the lower level of IMT. Furthermore, these rules must be strongly held. Violating them must induce in us a strong feeling of ‘compunction’ (Hare 1981, p. 29). Hare says that we can “[rely] in much of our moral thinking on relatively general principles” (1981, p. 38) and that we ought to have “intuitions and dispositions firmly built into our characters and motivations” (1981, p. 38). He prefers to call these “prima facie principles” (PFP’s). In order for agents best to deal with the complexities of the moral world, Hare advocates

the formation in ourselves of relatively simple reaction-patterns (whose expression in words, if they had one, would be relatively simple prescriptive principles) which prepare us to meet new contingencies resembling in their important features contingencies in which we have found ourselves in the past (1981, p. 36).

At least most of us are then, most of the time, ‘proles’ (Hare 1981, *passim*) in our moral thinking. We have a determinate ‘method’ of right moral judgement at this level, i.e. the application of rules.

At the upper, act-utilitarian level of moral thinking, agents engage in CMT. One task of CMT is to impartially determine the ‘right’ act in particular situations. This method bypasses the substantive moral intuitions, represented by our internalized ‘reaction patterns’, etc. It proceeds instead “in accordance with canons established by

philosophical logic and [is] thus based on linguistic intuitions only" (Hare, 1981, p. 40). For Hare these linguistic intuitions are represented by the linguistic properties and logical canons of prescriptivity, universalizability and overridability. This trio, taken jointly, is uniquely characteristic of words used in moral discourse. CMT also has the pre-educative role of justifying and selecting the rules which we shall be wise to internalize at the intuitive level, and apply in our everyday moral judgments. In other words, the critical moral educator or policy maker has to design a code of rules which will provide the blueprint for childrens' moral education. This is needed because, for Hare, the intuitive thinking that we normally apply in our everyday lives will be a matter of applying these rules.

It is important to see that, for Hare, these two levels of moral thinking, and the methods appropriate to each, are sharply distinct. The human critical thinker selects principles for use at the intuitive level, "on the score of their acceptance utility, i.e. on the ground that they are the set of principles whose general acceptance in the society in question will do the best, all told, for the interests of the people in the society considered impartially" (Hare, 1981, p. 156). He suggests that only an archangel could do this selecting perfectly. However, as humans using the critical method, with the aid, for example, of our ancestors' critical thinking, we can work out some approximation of the archangelic ideal. The point is that what principles we shall desirably hold is decided at the level of critical thinking, while intuitive thinking is based on our accepting those principles in a strong (uncritical) way.

The 'complete separation' of the two levels (Hare, 1981, p. 32) is made sharper by the two senses of 'right' that Hare claims characterize judgments at each level. The act-utilitarian archangelic judgments of the critical level constitute the 'definitively' right judgment in any particular case. These may occasionally override our intuitive judgments, but CMT requires an immense and daunting knowledge of facts and logic. Thus, it is usually reserved for a 'cool hour' in which one has time to collect the facts, and, in light of those facts, to consider different options carefully. However, intuitive judgments are also 'right' in one sense, i.e. for the purpose which they are intended to serve. They are 'prima facie' right insofar as the critical thinking that selects them for use has been competently done. Moral education, on Hare's view, will involve learning to ignore, in situations in which it is fitting, the judgments made at the critical level. Moral education will also seek to develop competent critical thinkers so that we can employ such thinking when *that* is fitting. We become educated, then, to perform much like well trained tennis players. Tennis players react instinctively to the exigencies of the game as it progresses, but they also analyze their shots and movements more reflectively while watching them leisurely on videotape. Both kinds of thinking have their place for the moral thinker as for the tennis player, but clearly any attempt to use one level of thinking in improper situations will yield disastrous results. Hare argues that, for various reasons, humans usually ought to eschew the methods of the critical level in favor of the rules of the intuitive level. Why? Hare gives two general reasons. First, we rarely have the time or ability to analyze particular exigencies. Attempts to do so in such cases would be foolish, and would likely lead to botched moral judgements. As in any complex practice we need simplifying guides to help us. Secondly, due to human weaknesses such as partiality to self, critical thinking is wisely reserved for those situations in which one is detached enough not to be tempted to 'cook' one's principles in one's own favor. Like good tennis players, then, we will normally rely on the best set of 'rules of good play' (1981, p. 38). This set will be the set that, if followed, gives us "the best chance of acting rightly [even though the principles making up this set] are not definitive of the 'right act' on an act-utilitarian analysis" (1981, p. 38).

Is this a desirable picture of moral judgment? I think not. Hare is right to stress the importance of the simplicity and generality of rules, or what he calls 'prima facie

principles' (or PFP's), in the moral education of young children (1981, p. 35–36). Simplicity and generality are important for the psychological reason that it would be extremely difficult for very young children to learn complex principles in such a way that they could be competently applied. Also, by remaining general, PFP's allow children to learn to pick out the morally relevant features of a "variety of situations all of which have certain salient features in common" (ibid.). However, as a representation of anything approaching mature moral judgment, this model, i.e. a model based on relatively simple, general principles, constitutes a gross oversimplification.

### Conflicts of Principle and Moral Judgment

Why is it an oversimplification? Let us consider cases of moral conflict that Hare describes as "situations in which we seem to have conflicting duties" (1981, p. 26), or "conflicting [prima facie] principles" (1981, p. 32) or, what amounts to the same thing for Hare, "conflicts between intuitions" (1981, p. 32). Suppose, for example, as Hare does, that a good, long-time friend shows up unexpectedly, just for the day, and wishes to be shown around town. I have already promised my family that I will picnic with them today. I cannot fulfil both 'duties'. How shall I decide what to do? (1981, p. 28). For Hare there are two clear options. Either there is in fact no conflict, in which case I simply apply a relevant rule. Or else there is in fact a conflict and I resolve the conflict using act-utilitarian CMT. Other options, such as making our rules more complex, or having higher order intuitive moral principles, are inadequate, as Hare himself points out. Such procedures lead to a process of lengthening our rules and making them more complex that soon gets absurdly out of hand (1981, p. 32–34). Such rules will be impossible for children—or adults—to learn; and they will be useless as guides to future conduct. Furthermore, any sort of indeterminate 'judging or weighing' process is immediately dismissed by Hare (Hare, 1981, p. 34) [1]. Indeterminacy in itself is necessarily enough for such a process to be deemed irrational and arbitrary for Hare.

If we accept Hare's description of this simple case as one of conflicting moral rules, then obviously, contrary to his claims, we are going to need to use CMT very often to resolve conflicts like this. We then need to evaluate CMT more closely, as I do below. However, the most striking thing about Hare's example is that, as J. O. Urmson suggests, it appears absurd to consider it a case of conflicting rules (1988, p. 163). What rules are supposed to conflict here? 'Don't break promises' is one perhaps. But it seems absurd to have a rule, or PFP, in this case, of "showing one's friends around", or even Hare's version which tells us "not to betray [our] friends" (1988d, p. 275). If this does not convince, remember that for Hare PFP's must be quite rigid. We need to be strongly disinclined to consider alternatives to these principles when they present themselves, or else they lack the emotional significance they are supposed to have for us. In addition, if we consider this case a conflict of principles, in Hare's sense, we must necessarily resort to CMT to resolve the conflict. However, as Urmson points out, this has the morally deleterious effect of encouraging us to ignore the fact that there are all kinds of judgments one might make in cases like the one above that avoid the extremes of, on the one hand, arbitrary choice (e.g. simply letting one principle override the other for no apparent reason) and, on the other, impartial act-utilitarian analysis (1988, p. 165). We might, and most people do, have a principle of not betraying our friends. We might even take it to apply in the kind of case described above. But such a principle obviously requires much moral judgement for its application. The kind of judgment or weighing of alternatives in such cases is not like Hare's CMT since, as Urmson claims, "there are many relevant considerations [encompassed by such a principle], none of them decisive . . . All these considerations have to be taken into account, and the general welfare may be one, but only one, of them" (1988, p. 165). Clearly, then, if Hare's view of moral rules in IMT fails to do justice to the

complexity even of rather trivial cases such as this one, it will be woefully inadequate for dealing with the complexities and uncertainties of the moral problems raised by issues such as AIDS, animal rights, world hunger and overpopulation, or control over schooling in multicultural societies.

The persuasiveness of my argument that Hare's view oversimplifies moral judgment depends on an understanding of what Hare means by a PFP (or what I call a strong rule). Are these rules as rigid as I have argued? Following Hare, Urmson describes principles as relatively simple and general. They are "rules of conduct which we set out for ourselves or accept from others to be followed (ordinarily) *blindly*. By *blindly* I mean I am unwilling to consider arguments to the contrary" (1988, p. 164). They are, within the boundaries of their application, uncritically held. For example, I have a principle against interrupting people when they are speaking. Under normal circumstances I unreflectively adhere to the rule. It has quite a clear range of application. But if someone insults me, or if they tell a blatant lie or if they are about to be run down by a car, I break the rule. Whether or not this is a comprehensive account of the word 'principle', Hare considers this characterization to be quite similar to that of his own treatment of PFP's (1988d, p. 275).

Urmson also suggests the term 'adage' for rules which are more weakly held than blind principles. I shall call the former rules of thumb. This is appropriate since they seem to be akin to the rules of thumb that Hare contrasts PFP's to in *Moral Thinking* (p. 38). These are propositions such as 'Be kind to your friends'. They may serve as rough guides for one's actions. That is, they represent potential reasons, in competition with other morally relevant considerations, for considering something morally relevant in particular situations. Rules of thumb cannot be considered a "sufficient reason [against the performance of certain actions] in all circumstances" (Urmson 1988, p. 165). Many other reasons for taking alternative actions—some of which might be expressed as competing rules of thumb—may exist. Which rule, if any, counts as most important, and for what reasons, and what counts as its fulfilment, will be determined by particulars.

### A Weak View of Moral Rules Outlined

What precisely distinguishes rules of thumb from PFP's? This is an important question, since if there is no difference, the moral complexities that rules of thumb are supposed to allow for must also be allowed for by PFP's. My objection to Hare thus fails. But there is an important distinction to be made. PFP's are habitual, mechanistic responses. Thus, a rule such as 'Keep your promises' is quite straightforward in its application. As Charles Larmore points out, if I make a promise to return your book today, the only *moral* consideration here is simply returning the book sometime today. Considerations of judgment with rules of this type involve only 'non-moral' considerations such as when today to return the book (Larmore, 1987, p. 5). [2] However, principles of courage and friendship involve a multitude of moral alternatives for their application. In what way should I honor this principle? What counts as fulfilling the principle in a particular case? With what intensity should I pursue this alternative? In cases where more than one such principle is open to consideration, the complexity of the judging process is compounded. In this sense, such principles act more like rules of thumb. That is, they provide only rough general guides for action.

Hare has recently claimed that his PFP's actually include what Urmson calls 'adages' or what I call rules of thumb. This new 'wider' sense of principle includes rules of thumb like "be nice to one's friends" under "loosely specified conditions" (1988d, pp. 275–276). However, such principles, as I have argued, require considerable moral judgment for their application. Hare's move is to suggest that experience, either one's own or someone else's, can illuminate some latent determinate method of

judgment in the application of such principles. Hare argues from analogy. He suggests that, for example, although jurors need to use judgement in evaluating the truthfulness of witnesses' testimony, experienced ones can do so within the framework of a 'method'. This kind of thinking "can be rational because there is a way of determining, perhaps later, whether a witness *was* telling the truth. So we can build up an idea of whether the intuitive thinking we were doing was methodologically sound . . . What makes moral intuitions, when unsupported by critical thinking, irrational is that there is no such method" (1988d, p. 277). Therefore, Hare seems to think that, just as critical moral thinking can establish a methodology based on a code of moral rules, and itself provides a methodology of resolving the rare disputes at the latter level, so 'factual' critical thinking can establish a methodology for factual intuitive thinking. Additionally, this kind of critical thinking can itself resolve the rare disputes at the level of intuitive factual thinking. However, this is unconvincing. For example, whatever rules might be formulated for assessing the veracity of testimony are quite obviously going to be rules of thumb. 'Ignore hyperbolic testimony' is an excellent rule, but clearly it should be completely ignored in some circumstances. I have heard of a judge who immediately dismissed a child's testimony at a sexual assault trial when the child answered the question "How often were you molested?" by saying "Millions of times" [3]. To avoid such obtuse judgements, one might say that the rule itself should be made more flexible. To do this one could suggest, as Hare might, a rule such as "Treat hyperbolic testimony with care" or "Ignore hyperbolic testimony (under 'loosely specified' conditions)". However, these clearly are nothing like Hare's strong rules. They require much judgment for their application. Similarly, in the moral case, the more the rules of the intuitive level are represented as 'loosely specified', the less they can be rules which are simply applied by some kind of 'reaction pattern'. Being nice, even to friends, is often not what is morally required of us, but obviously there is no determinate methodology available for determining when it is and when it is not.

The preceding discussion suggests that adherence to moral rules is not "the best way of characterizing" virtuous character as Hare claims (1988d, p. 279). We normally treat rules of thumb as summaries of important truths. Thus, "Too many cooks spoil the broth" and "Many hands make light work" (Hare's examples) reflect insights which we recognize from diverse contexts in our actual lives. It would be extremely unnatural to suppose that we rely on such insights in any strong way to guide our evaluations of particular situations. Artists do not paint in large groups on the grounds that 'Many hands make light work'. We may, in certain situations, use such adages as simplifying devices to represent features of particular situations but we would normally attend much more closely to the particular features of particular situations before deciding which adage was applicable. Even once we have done this, we think of this characterization of the situation as only a sort of summary of a more complex and elusive situation which underlies it. Much of moral importance is left out. Similarly we cannot superimpose upon the virtues a set of principles, declaring the two identical (Hare, 1988d, p. 280). If we do, we find that principles of courage, humility, other-regard, etc., involve considerable judgement for their application. We will find that considerations of what acts actually count as courageous and with what sort of intensity and vigor they might be pursued require judgement of considerable delicacy and skill. Such considerations cannot necessarily be characterized, as Hare claims, as simply "obedience to a principle or disobedience" (1988d, p. 168) since much more needs to be said about the adeptness or ineptitude of the application of the principles. Hare's view seems to make general principles into the tail (and a rather small one at that) wagging the substantial dog of judgement [4].

My argument so far has been that Hare, given the resources available to him in his 'two-level' theory, has two lines open to him in dealing with the complexity raised by

conflicts of moral principles. First, he could simply extirpate judgment altogether at the level of IMT. This would require some explanation of how everyday intuitive thinking would be a matter of *applying* a set of internalized, habitual rules or principles in such a way that no judgement was required. Given the nature of the relatively simple, general principles that constitute one's 'moral ammunition' at the intuitive level, this move would make his conception of intuitive thinking quite implausible. It becomes implausible because most, if not all, of our principles at this level require for their application some conception of judgement, if their application is not to be arbitrary. We need *some* way of determining what counts as a good or bad application of such principles. If these arguments are on target then CMT must play a much larger role in everyday moral judgment than Hare indicates.

### **Critical Moral Thinking: the problems of act-utilitarianism**

Hare seems to recognize this fact when he backpedals, under fire from critics, from his earlier claims regarding the primacy of IMT. For example, he has claimed that "moral development is a development toward critical thinking" (1982, p. 316). More recently, he has said that critical thinking is "the completion of the moral thought begun at the intuitive level" (1988c, p. 261). If this claim reflects Hare's actual position, then mature moral agents will, contrary to his claims elsewhere, rely less on rules (which, as we have seen, are seriously inadequate anyway) and much more on the method of CMT.

For Hare, CMT is a decision procedure involving a strongly impartial identification with all the relevant preferences obtaining in a particular case. The agent, when faced with a moral choice, must 'fully represent' to herself all of the preferences of those affected by her action. The preferences of the overall greatest intensity determine the right choice. Thus the right choice for Hare is equivalent to "situational act-utilitarian judgment" based on equal evaluation of preference intensity; but Hare's necessary reliance on CMT raises a puzzle. Such an egalitarian judgment procedure is open to well-known and much rehearsed objections. It allows too strong a role to 'evil desires' such as those of racists or Nazis; it does not allow a strong enough role for intuitive values such as that of loyalty to family and friends. I argue that Hare's CMT is a sitting duck for these kinds of objections.

First, I will clarify what CMT is. The sharp distinction Hare draws between the two levels of moral thinking is based on a theory of the meaning of concepts that are typically used in moral discourse. Understanding this theory helps us to understand the theoretical basis for CMT. CMT and IMT are levels of substantive moral thinking. However, they are established, Hare claims, by inquiry in the area of philosophical logic. The clarity of our thinking at these levels is based on our understanding of moral concepts. These meanings are established at a third level of thinking—the meta-ethical. Here, the logical canons of prescriptivity, universalizability and overridability are laid bare to reveal the meaning of evaluative words like 'ought'.

For Hare, moral judgements at both the critical and intuitive levels are universal prescriptions (1981, p. 41). The difference between the two levels is described as a difference in the specificity of the two kinds of principles. Intuitive principles, as we have seen, must be of limited specificity; they must be fairly general. This is required due to the purposes served by intuitive principles (e.g. they must be used for moral education, they must allow us to deal with situations similar in their general moral features). Critical principles, while remaining universal, can be of unlimited specificity. Indeed, Hare says, for their purpose, critical principles "have to be highly specific" (1981, p. 41). The differences between principles at the two levels do not stop here of course. They also have different purposes. Consider the case of moral conflicts. As Hare says, when we are thinking about conflicts between principles at the critical



level, “there is a requirement that we resolve the conflict” (1981, p. 26), but there is no such requirement at the intuitive level. Thus, the way we resolve a conflict between our principles (e.g. the case of the unexpected friend above) is by doing some hard CMT and tailoring a critical principle that accounts for all the relevant, particular features of the situation. This principle must necessarily lead to maximum satisfaction of preferences. This is the same judgement a “careful act-utilitarian would make” (Hare, 1981, p. 43).

There is a further feature of the logic of the moral concepts which differentiates the two levels. This is the property of overridingness. Hare notes that we often may continue to *hold* our PFP’s even if we find ourselves constrained to do the CMT leading us to a more specific principle which resolves our moral conflict. This leaves us with two prescriptions, which may conflict. But surely we are to act on the critical principle in this case (assuming we are confident about the competence of our CMT). The reason for this is that critical principles are taken as overriding. That is, we may make exceptions to intuitive principles. Hare says that this property in fact distinguishes moral evaluative judgements from other evaluative judgements like aesthetic ones (1981, p. 52).

Critical moral principles on this account are prescriptive and universal (i.e. they contain no individual constants and are governed by a universal quantifier). In addition, they are overriding. If they are sincerely held and the individual is able to act upon them and there is no external impediment to acting upon them, they must be acted upon if they are to count as the agent’s bona fide moral principles. Hare says that PFP’s are also prescriptive and universal, but they *can* be overridden. However, he says, the fact that they can be overridden “does not mean they are not prescriptive; if applied, they would require a certain action, but we just do not apply them in a certain case” (1981, p. 59). So PFP’s are universal prescriptions, but they are not overriding. We will desire to act upon them so long as they are considered morally relevant. Thus we can always be inclined to uphold them even in cases where they are overruled by competing critical prescriptions. In such cases, if the PFP conflicts with the critical principle, the former will be regarded as morally irrelevant in that case (which is not the same as simply discarding them at the intuitive level. We can still have the intuitive desire to uphold them). Nonetheless, in such cases IMT clearly must give way to CMT.

Hare distinguishes two kinds, or two uses, of CMT. Recall that IMT is necessary to take account of what seemed to be certain facts about our everyday moral thinking. Our intuitive principles must be selected or tested, or somehow based on, CMT. Thus, CMT ushers in two kinds of moral principles. First, the highly specific, universal, prescriptive and overriding critical principles. Secondly, more general universally prescriptive (but non-overriding) PFP’s. The latter are to be justified by being principles “whose general acceptance would lead to people’s actions and dispositions approximating to the greatest extent to the deliverances of a perfectly conducted critical thinking” (1981, p. 61). They have high ‘acceptance-utility’. These two procedures are the two uses or kinds of CMT.

I will call the former kind ‘direct CMT’ and the latter kind ‘indirect CMT’. I will consider direct CMT first. How is this kind of thinking act-utilitarian? William Frankena notes that direct critical thinking “consists in using the universal prescriptivist, act-utilitarian method directly to determine what to do in a particular case” (1988, p. 47).

However, what does it mean to call this method act-utilitarian? For Hare, there is a particularly strong requirement of universalizability connected to moral judgements. Briefly, if we hold a prescription, which represents a preference of ours, (since, as Hare says, to have a preference is to assent to a prescription) and if this is a *universal* prescription, we must “accept the universal application of the prescription; and this

includes its application were we in the other's position [or in the position of all others affected by my action]. So the facts we need to be cognizant of will include facts about his [their] positions as it affects him [them] with his [their] preferences . . ." (1981, p. 89). For Hare the requirement to universalize moral principles is interpreted in an ambitious way. When I put myself in another's shoes, I must take account of all others' preferences equally. Thus: "It follows from universalizability that if I now say that I ought to do [X] to some other person, I am committed to the view that the very same thing ought to be done to me, were I in exactly his situation, including having the same personal characteristics and in particular the same motivational states" (Hare 1981, p. 108).

This method of representation generates act-utilitarianism since it "involves [coming to have] equal motivations [i.e. motivations equal to those of all persons affected by our actions], with regard to possible similar situations, were we in them" (1981, p. 95). Since we come to have, ourselves, in some way, preferences equal to all of those affected by our actions, we are in effect deliberating in the same way that an impartial 'ideal observer' (1981, p. 44) would do. Through an admittedly difficult act of imagination, Hare's theory requires that we treat all the preferences of all those affected by an action, including ourselves, equally and impartially.

Holding a Harean universal principle requires that I come to 'identify' with others so strongly that I come to feel now (i.e. in my deliberations) their preferences in the same way they do, were I in their position, with their personal characteristics. Thus, if I have a strong desire to park my car in a driveway currently occupied by X's bicycle, and if X has a weak desire not to move his bike, after full representation of all preferences, I find I should move the bike and park my car. However, if X has a very strong desire not to move his bike, then I find—after full identification—that I should not move the bike. This is act-utilitarian since the preferences of all those affected are given equal weight and the strongest preferences are decisive.

It should be noted here that in cases in which the preferences of many persons are concerned, this method does not entail that the most intensely held preferences win out. For example, in a society in which a large majority prescribes slavery for a small minority the prescription may survive the universalization process. The many preferences of the slave owners (though relatively weak) may outweigh the few (though intense) preferences of the slaves. Strongest *cumulative* preferences are maximized.

### **CMT and Moral Schizophrenia**

How might Hare respond to objections raised by the example of tyrannical majorities, or by the familiar case of the 'down-and-out' kidney and heart donor who, act-utilitarians must think, ought to be killed against his will in order to save the lives of two other people (Hare, 1981, p. 132)? Predictably, his defense depends upon the distinction between the two levels of moral thinking. First of all, he says, IMT will advocate a prohibition on murder in the form of an intuitive principle. Such a principle will have a high acceptance utility. Secondly, we will be very unlikely to have good reason to overcome this prohibition in the particular case since it is very unlikely that we will be able to compile the information necessary to do competent critical thinking and it will be equally unlikely that we will be able to be justifiably confident that the information indicating that we ought to kill the 'down-and-out' is correct. Thus, all things considered, we should always or almost always stick with our intuitions. As Hare says, the 'high degree of probability' which would be needed in connection with the information we gathered (e.g. about the down-and-out's lack of connections, the discretion of all those involved in not revealing that the act has been done, etc.) "will not be forthcoming in many actual situations, if any at all" (1981, p. 133).

Hare deals with two further, and more 'genuine' (1981, p. 137) examples in a similar fashion. The first example considers the common sense moral intuition that we should "give . . . [extra] weight to the duties usually thought to exist towards particular persons, or to ties of affection or loyalty which bind us to them but not to mankind in general" (1981, p. 135). That is, we are usually thought to have special duties to our spouses and children, etc. These are stronger than our duties to total strangers. Act-utilitarianism seems to require, for example, that parents avoid being partial to the needs of their own children. In requiring us to maximize preference satisfaction, and in forcing us to treat the preferences of all persons equally, it seems to ignore our common intuitions about such relationships. To take a particular case, offered by Hare (via Bernard Williams) I am to suppose that I have been "in an air crash and the aircraft catches fire, but [I] have managed to get out; in the burning plane are, among others, my son and a distinguished surgeon who could, if rescued, save many injured passengers' lives, to say nothing of those whose lives he would save in his subsequent career. I have time to rescue only one person" (1981, p. 138). Hare's first comment is that once again, one is unlikely to have enough information, of which one can be confident that it is true, to take what is supposed to be the utilitarian's unflinching prescription—i.e. save the surgeon, let your son burn! However, even if one did have the necessary information, "You have (rightly from the critical utilitarian point of view) been brought up to attach dominant importance to these family loyalties" (1981, p. 138). Thus, you will probably rescue your son.

This characterization of the problem raises an important point. The upbringing recommends, in such cases, what is the morally *wrong* action from the critical perspective. This seems to be a serious flaw in the theory. Even seeing clearly what the right action was from the point of view of my theory, I am likely to do the wrong thing from that point of view. That is a seriously schizophrenic result.

This point strikes at the heart of Hare's theory and needs to be explained further. For Hare, the 'well-brought up' individual is, it seems, the morally educated one. This means that she is adept at thinking at both moral levels. She knows when it is appropriate to ascend to the critical level. She knows also that intuitive thinking is a mere concession to human weakness and partiality. Hare might say that it is often inappropriate to make the move to critical thinking, because of the urgency of the situation and the resulting lack of information and fellow-feeling. In the cases we are considering, Hare can allow values like family loyalty and the prohibition on murder to prevail. However, upon reflection, the agent must admit that *on the whole* the action was wrong. It is still 'right' (i.e. *prima facie* right) in one sense, as Hare claims. But we now see that this sense of right is the relatively feeble sense in which the action was the best one could do as an agent who must live by moral standards fitted to human frailty and ignorance. Thus, regardless of lacking the 'leisure' to perform critical calculations in the face of 'real life stress and urgency', Hare's agents must adopt a sort of schizophrenic outlook. They must diligently and confidently adhere to general rules at the intuitive level; but at the critical level, these same principles must be viewed as mere 'useful fictions', designed for a benighted humanity. They are fictitious because they fabricate a more or less crude substitute for genuine important moral truth. This truth, then, is commonly something which is out of the reach of humans, but also something for which they should rarely strive. However, this point raises the possibility for Hare to take advantage of another response to the objection I have raised. He might acknowledge that the ascent to critical thinking does create a psychological 'dissonance'. But since these moments of ascent are so rare, they do not create anything like a schizophrenic *identity*. Moral educators who consider a schizophrenic identity an educationally reprehensible result, might find psychological dissonance of this weaker type quite tolerable. However, if our moral principles conflict as frequently as I have suggested, then it is clear that the rules of IMT are relatively flimsy

moral props. For instance, in the case of slavery, the fact that enslavement goes against the grain (at the level of IMT) of our moral feelings must be immaterial to our moral judgment.

Hare's defense so far has consisted in his insisting that counter-utilitarian examples ignore the fact that different types of moral thinking are appropriate for different times and cases. IMT is for 'everyday situations' and CMT is reserved for a 'cool hour'. While the judgments at the two levels may exist in tension, they each have an appropriate time and place, and thus these judgments never openly or sharply conflict. However, given any fairly diverse set of principles, conflicts are bound to arise often. Also, our set of principles would need to be fantastically diverse to avoid unanticipated moral problems. Hare cannot deny that CMT must at least be attempted in such cases. In these cases our judgments at the two levels will sharply and openly conflict.

Indirect CMT does nothing to resolve this difficulty. As Frankena notes, the crux of the problem for Hare is that he is unable to avoid the objectionable implications of what is the 'fixed point' of this theory. That is: "the right act to perform in a particular situation is the act that one of archangelic powers would pronounce to be so if he addressed himself to the question and used [direct CMT] to answer it . . ." (1988, p. 48). The unrestrained use of this kind of thinking is, as Hare admits, dangerous and morally objectionable (at least for humans). Bernard Williams has also pointed out the damaging effects of such thinking to the integrity of moral agents (Williams, 1973, pp. 108–118). If Hare's PFP's are to be of any use, they must be strong enough to withstand the temptation to use direct CMT. They are uncritical—i.e. they must operate, in Williams' word, as "black box mechanisms" (1988, p. 190). One remaining line of defense for Hare would be to find some middle ground; some insulation, as it were, between the two extremes of IMT and direct CMT.

Hare has suggested a way of interpreting the two levels in such a way that all of our moral thinking is actually a combination of IMT and CMT. However, this move, rather than resolving the dilemma I have been posing, simply heightens the schizophrenic implications of Hare's view. Hare admits that "almost all acts are to some degree self-educative", although some are more so than others. He gives as an example of acts which are 'especially self-educative', acts involved in adhering to a rule of "suffering fools gladly" (1988a, p. 229). This is a good example for our purposes since obviously much judgement would be needed to apply such a principle. This would mean, as Hare himself says, that "We should . . . end up applying [direct act-utilitarianism] to pretty well all our acts to some degree . . ." (1988a, p. 229). The reason for this is one I have suggested already. We always, or almost always, have more to learn about what counts as or *constitutes* 'following' or 'obeying' a simple rule in a particular case. Some judgement is needed to tell us which rule is the best one to follow and what is the best way of following the rule.

This also makes clear another important point. The critical thinking involved in "direct application of the act-universal prescriptivist, act-utilitarian method to particular cases" and the critical thinking involved in "selecting or justifying" our intuitive principles seem to occur, in Hare's view, in the same thought process. Thus, the 'two uses' of critical thinking may often occur simultaneously. However, it also suggests that, at least very often, we are not simply applying intuitive principles in some uncritical way, i.e. as 'black box mechanisms'; rather we are using critical thinking in almost all of our acts to (a) 'cultivate dispositions' and (b) to calculate the 'optimal consequences' of the particular act dictated by the principle on a particular occasion. Hare says of the above example, ". . . if we found ourselves *not* suffering fools gladly, we should pull ourselves up *both* because it led to less than optimal consequences on that occasion, *and* because we were getting into bad habits" (1988a, p. 229).

This is a revealing example. It clearly shows that insofar as principles cannot be

simply followed by 'habit', insofar as judgement is required for their application, insofar as the acts involved in following the principle are 'self-educative', Hare needs direct critical thinking to monitor the conduct of the moral agent in a quite pervasive way. It seems that his view of the relationship between the intuitive and critical levels commits him to the view that rarely do we simply uncritically follow a given principle. Rather—at least for Hare—we usually, if not always, need to use critical thinking to determine what is morally best. Hare can certainly say: "... we should pull ourselves up both because it led to less than optimal consequences on that occasion, and because we were getting into bad habits" (1988, p. 229). But as soon as 'habit'—conceived as a rule—fails us in a particular case, direct critical thinking, and the act-preference utilitarian standard, is needed to judge that particular case. Thus, the form of schizophrenia inherent in Hare's theory is particularly severe. *We do not adopt critical thinking on rare and rather bizarre occasions. Indeed, we are not just required to alternate moral personae, one instant rational archangelic strategist, the next mindless prole. We are required somehow to adopt both personae at once. Thus, the moral educator who educates for competence at both of these levels must necessarily encourage the internalization of such schizophrenia in her students.*

### Aristotle on Moral Rules and Virtue

Obviously, the failure of Hare's theory to account for the complexity of the moral life does not compel us to throw up our hands in despair. I hope to argue more fully in a future paper that Aristotle provides us with a much more promising picture of moral judgment, and of a moral education based on it, but I can give the rough contours of such a view here.

Aristotle's understanding of moral development exhibits a harmonious relationship between moral principles and reflective moral judgment that Hare's view is unable to account for. This view avoids the schizophrenic results of Hare's view. Hare's theory entails that our habitual, trained moral responses can be sufficiently characterized in terms of simple general rules. These rules provide determinate standards of moral judgment at the intuitive level. When these rules conflict, the rules must be altogether abandoned in favour of a new kind of reflective thinking—CMT—and its determinate moral standard of maximal preference satisfaction.

In spite of the fact that Hare claims that his 'two-level' view owes some inspiration to Aristotle (Hare, 1981, p. 25), Aristotle differs sharply from Hare. A rough similarity undoubtedly exists between Hare's level of IMT—which represents our habitual and trained moral responses—and what Aristotle calls the excellences of character. Similarly, there is a rough correspondence between Hare's CMT—which represents more reflective moral judgement—and Aristotle's notion of the excellences of intelligence, but the similarities are superficial. Aristotle does not see the excellences of character as codifiable. They do not function as mere 'noble lies' for the benighted masses, as in Plato and Hare. The habitual, ingrained moral responses, products of our moral training, are in fact constitutive of our more reflective understanding of the good. Unlike Hare, Aristotle recognizes that we can rarely shed our moral 'habits' or principles, as a snake sheds its skin, when moral conflicts require us to do some reflective thinking.

For Aristotle, our habitual concerns and principles represent for us the vague outlines of the moral landscape. Practical wisdom, or *phronesis*, is needed to fill in more detailed features of the landscape. For Hare, the moral habits we express as rules may also be said to paint a hazy and general outline of the moral landscape. But when these elements 'conflict', one must, as it were, erase the outline and begin again using a completely different method (CMT). Alternatively, the Aristotelian view conceives of

moral habits as essential tools for 'composing the scene' within which our reflective judgments will be made. This is not to say that we cannot, and do not often need to, refer to our habitual concerns, and to modify our 'moral outline' an indefinite number of times in doing so. The main point for Aristotle is that our final picture is constituted by a reflective consideration of the complex and varied concerns we are brought up to pay attention to. For Hare, our final judgments must ignore these concerns of character and virtue. Furthermore, just because, following Aristotle, we need not adhere to the independent standard of maximizing utility, this does not mean that we can adhere to no standards at all. Standards of friendship, regard for family as well as for the most disadvantaged, etc., must be weighed against, not subordinated to, standards of maximizing overall intensity of preferences. And the sort of ability required for making such judgments cannot be characterized in terms of rules. Obviously, for educators this conception of virtue has much importance, including the importance of teaching with complex, detailed and rich examples, both from life and from literature, rather than teaching by precept. I will save this discussion for another occasion.

### Acknowledgements

For their comments on earlier versions of the ideas presented here, I would like to thank Allen Pearson and especially Eamonn Callan.

*Correspondence:* Kevin McDonough, Department of Educational Policy Studies, College of Education, University of Illinois at Urbana-Champaign, 1310 S. Sixth Street, Champaign, IL 61820, USA.

### NOTES

- [1] Hare does allow that in some very simple cases of intuitive thinking "we may 'feel sure' that some principle or some feature of a situation is in *that* situation more important than others . . . We shall then be able to sort the matter out intuitively, letting one principle override the other in this case, without recourse to critical thinking" (1981, p. 50), but any intuitive choice between principles makes sense only insofar as it converges with 'critical' judgment.
- [2] Even here actual cases are rarely this simple. I could fulfil the promise by ringing your doorbell at 11:59 p.m., when I know you are asleep and must wake up early the next morning for an important job interview. Such alternatives will obviously be ruled out by virtuous agents. So even straightforward rules are apt to engage judgment in this minimal way.
- [3] I owe this example to Eamonn Callan.
- [4] I owe this devastating one-liner to G. J. Warnock (1979), who used it to make a slightly different point against Hare.

### REFERENCES

- FRANKENA, W. (1988) Hare on the levels of moral thinking, in: DOUGLAS SEANOR & N. FOTON (Eds) *Hare and Critics: Essays on Moral Thinking* (Oxford: Oxford University Press).
- HARE, R.M. (1981) *Moral Thinking: Its Levels, Method and Point* (Oxford: Oxford University Press).
- HARE, R.M. (1988a) Comments on Frankena, in: *Hare and Critics* (op. cit.).
- HARE, R.M. (1988b) Comments on Harsanyi, in: *Hare and Critics* (op. cit.).
- HARE, R.M. (1988c) Comments on Scanlon, in: *Hare and Critics* (op. cit.).
- HARE, R.M. (1988d) Comments on Urmson, in: *Hare and Critics* (op. cit.).
- KUPPERMAN, J.J. (1978) Inhibition, *Oxford Review of Education*, 4.3, pp. 277-287.
- LARMORE, C.E. (1987) *Patterns of Moral Complexity* (Cambridge: Cambridge University Press).
- RATHS, L. *et al.* (1978) *Values and Teaching* 2nd edn (Columbus, Ohio, Merrill).
- URMSON, J.O. (1988) Hare on intuitive thinking, in: *Hare and Critics* (op. cit.).

- WARNOCK, G.J. (1979) Morality and Language: a reply to R.M. Hare, in: D. B. COCHRANE, C. M. HAMM & A. C. KAZEPIDES (Eds) *The Domain of Moral Education* (New York: Paulist Press).
- WILLIAMS, B. (1973) A critique of utilitarianism, in: J. J. C. SMART & B. WILLIAMS (Eds) *Utilitarianism: for and against* (Cambridge, Cambridge University Press).
- WILLIAMS, B. (1988) The structure of Hare's theory, in: *Hare and Critics* (op. cit.).