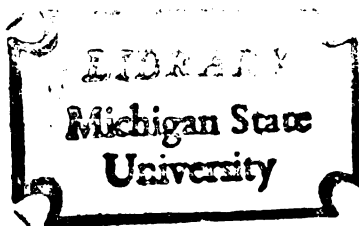




THESIS



This is to certify that the  
thesis entitled

The Use of Principles in Moral Reasoning

presented by

Thomas Tomlinson

has been accepted towards fulfillment  
of the requirements for

Ph.D degree in Philosophy

*George C. Keruer*  
Major professor

Date May 13, 1980



OVERDUE FINES:

25¢ per day per item

RETURNING LIBRARY MATERIALS:

Place in book return to remove  
charge from circulation records

*123*  
APR 26 85

SEP 15 1999

SEP 26 1999

JUL 21 2001

MAR 22 2005  
MAR 22 2005

030706

THE USE OF PRINCIPLES IN MORAL REASONING

By

Thomas S. Tomlinson

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Philosophy

Spring 1980



## ABSTRACT

### THE USE OF PRINCIPLES IN MORAL REASONING

By

Thomas S. Tomlinson

It has traditionally been thought that the question whether moral judgments can be justified is one that can be answered in the affirmative if and only if there can be discovered some one principle or consistent set of principles from which one may deduce the proper moral conclusion, after a determination of all the relevant facts. It is only at the level of principle that an objective moral truth will be found, and it is by deduction from these principles that we must determine the truth on more particular ethical questions. It is my thesis that this picture of moral reasoning is false.

After first showing that the universalizable nature of particular moral judgments does not imply the necessity of their being derived from a general principle that relates the non-moral facts to the moral conclusion, I go on to an examination of one of the crucial problems for any deductive moral system, which is how to account for the variety of moral judgments within a consistent set of

principles, for there seem to be at least two major moral considerations which might at times pull us in opposite directions--the injunction to respect the integrity and worth of individuals, and the obligation to do good. If we make "respect" the supreme requirement, as in Alan Donagan's fundamental principle to respect all human beings as rational creatures, the principle becomes too vague to support any clear lines of deductive argument to particular conclusions, and we are thrown back upon reliance on particular considered judgments to supply an interpretation of the principle. The other alternative, utilitarianism, is forced to face much the same problem, though in this instance the vague expression is the "happiness of mankind," which can also be understood only in terms of considered judgments.

This suggests that as principles become more general, they require the use of vague terms of moral assessment, making their application to specific moral issues problematic unless we bring in our judgments about particular cases to supply an interpretation, which is not a procedure involving appeal to principle at all, much less the principle being interpreted. This is a possibility which is elaborated and further supported by showing how it materially affects two attempts, by Alan Gewirth and John Rawls, to provide theoretical arguments for certain moral principles or particular moral judgments.

It is then shown more generally how the application of moral concepts require the use of moral judgments, and that this is a process not clarified by the appeal to principle, but instead by appeal to the indefinite body of particular moral convictions which underlie our understanding of principles. The application of principle, then, though necessarily part of moral argument, is not itself a matter to be handled by deduction, using principles. Whether a principle applies to the case before us is instead resolvable only by appealing to analogies with settled cases, which provide relatively fixed points which anchor the understanding of the moral principle. Though not formally explicated, this process of analogical reasoning is illustrated through an examination of several arguments concerning specific moral questions, and a critical response to a contention that such issues can be settled only by the use of moral theory.

To Debbie, who generously indulged this  
indulgence of my love of  
philosophy.

## ACKNOWLEDGMENTS

I would like to thank Dr. George Kerner for the practical guidance necessary for bringing this project to completion, and Dr. Martin Benjamin for his honest and penetrating criticism, one of the marks of "respect for persons."

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION . . . . .	1
II. CONSTITUTIVE PRINCIPLES AS A LOGICAL NECESSITY. . . . .	13
III. THE SPELL OF THE DEDUCTIVE SYSTEM. . . . .	21
IV. THE PRINCIPLE OF UTILITY . . . . .	38
V. THE FAILURE OF <u>A PRIORI</u> PROOFS . . . . .	50
VI. THE INSEPARABILITY OF PRINCIPLES AND CONSIDERED JUDGMENTS . . . . .	56
VII. VAGUE PRINCIPLES AND DEDUCTIVE REASONING . . .	75
VIII. ANALOGICAL REASONING: SOME ILLUSTRATIONS. . .	89
Familial Obligations. . . . .	90
Limits of the Right to Life . . . . .	99
IX. SUMMARY AND CONCLUSIONS. . . . .	113
NOTES TO THE TEXT . . . . .	121

## CHAPTER I

### INTRODUCTION

At least since Hume it has been a received opinion among philosophers that reasoning necessarily proceeds through either of two forms, deduction or induction. Take any proposition one will; if its truth is susceptible of proof at all, it must either (1) be because it itself represents a "relation of ideas" that simply connects the meanings of the words contained in it (as in a definition) or is connected through a chain of such meaning relations to some other proposition whose truth is accepted as given (as in a logical or mathematical demonstration), or (2) be because it is part of or an explanation for observed "associations" between events or properties, as in simple induction by enumeration or in the testing of a scientific hypothesis against the experimental evidence.<sup>1</sup>

This view that deduction and induction together exhaust the possible kinds of argument has had its influence in moral philosophy. With the exception perhaps of Dewey, no philosophers have thought that moral reasoning can be profitably thought straightforwardly

inductive in character, for induction takes as given the "data of observation" on which it may generalize. In the case of morality, such data could only be the sets of moral beliefs that individuals actually have. But the primary function of moral reasoning would presumably be to provide independent grounds on which to criticize and amend at least some of the moral beliefs actually held by people. And so the results of such a process of induction would have to be used to undercut the evidence on which those results were based!

Similar problems attend recent attempts to work out analogies between scientific theories and moral principles;<sup>2</sup> each sort of claim, it is said, is justified insofar as it systematizes and explains the relevant data, whether that be the result of empirical observation or the collection of our settled moral convictions. The analogy fails at the crucial point, for explanations take as given that which is to be explained, whereas moral principles may ostensibly be used to critique at least some of our settled moral convictions.

Regardless of what the ultimate merits of such views may be, the historical fact remains that the search for reason in ethics has assumed that moral reasoning must be deductive in character, once the possibility of inductive moral argument has been thought ruled out. Thus, with few exceptions (notably the short-lived period



of emotivism), the history of moral philosophy may be seen largely as the story of attempts to specify a fundamental principle or set of principles such that once the relevant non-moral considerations have been established for any particular case, the proper moral conclusion telling us what we should do follows by a logical necessity from an application of the appropriate principle.

Certainly this is true of the two opposed traditional normative systems, utilitarianism and the Kantian categorical imperative. But even those who have no commitments in either of these directions usually suppose that moral issues are matters of principle, which are settled by identifying and arguing for a suitably refined general principle which is decisive in the case before us. How do we determine whether we should lie to a patient who has a terminal illness, so that he may live his last days in his accustomed manner, free from anxiety? We must decide upon a principle specifying the permissible use of beneficent lies, which will then tell us whether we may lie under the circumstances before us (together perhaps with some ancillary principles which will tell us whether "living his last days in his accustomed manner, free from anxieties" counts as a "benefit").

The consequence of taking the Humean view of reason then, is to suppose that if morality is rational,

all particular moral issues can be adequately addressed by moral principles alone, however those principles themselves might be argued for, so that the paradigmatic form of moral argument must be

1. All X are right (wrong, good, bad).
2. This is an X.
- ∴ 3. This is right (wrong, good, bad).

where "X" specifies some moral or non-moral conditions (perhaps quite complex) which can themselves be established as present in the particular case before us using either deductive or inductive forms of proof.

There are several directions from which one might attack this view of the use of principles. One might first claim that there are no non-vacuous and exceptionless moral principles that can occupy the place of the major premise in such arguments.<sup>3</sup> From this perspective, the best one can do is utter inanities like "Do the right thing," for any substantive principle that might be offered will have exceptions, and hence will be false. Though the argument the principle would be used in might be deductively valid, the argument would not establish its conclusion, since it would have a false premise.

As an objection against the deductive model of moral reasoning, this carries a heavy burden of proof, for it must show of each principle as it is advanced that it has exceptions, and that refinements to take account

of this first set of exceptions result in a new principle with a different set of exceptions, and so on. Thus, the positive argument for claiming that there are no exceptionless moral principles is necessarily an anecdotal one, pointing out the failure of past theories to provide exceptionless principles. It is never relieved, however, of the necessity of taking up each new principle or refinement as it is offered, and so its truth remains uncertain, perpetually open to challenge.

The more serious problem with this objection is that it's just not true that there are no non-vacuous and exceptionless moral principles. There is, for instance, the principle, "One may break a trivial promise to prevent a great harm." I'll just assert this has no exceptions, since it's impossible to demonstrate this, and leave the burden of proving the opposite to those who think there are no exceptionless principle. And though this principle may be obvious, it's not vacuous, tautological, or analytically true. It would be analytically true if "trivial promise" meant "promise which may be broken to prevent great harm." But "trivial promise" does not mean this, for certainly there are some non-trivial promises which may be broken to prevent a great harm (e.g., under serious enough circumstances, it would be permissible to fail to repay a loan on time). The principle would also be vacuous if one was always

permitted to do whatever was necessary to prevent a great harm. But this, obviously enough, could not be the basis of an argument that there are no exceptionless principles.

There is actually a whole class of principles which may be viewed as exceptionless. These are those principles called by Ross<sup>4</sup> and others prima facie principles. They identify morally relevant features of actions and situations, and tell us what we should do if there are no vountervailing considerations. "Always keep your promises" is an exceptionless prima facie principle, for it tells us that if there are no contrary indications (if no other prima facie principles apply which indicate otherwise), we should always keep a promise. But even if one concedes that these are properly called exceptionless principles, they can't be the sort of principles had in mind by those who think moral argument is deductive. If moral principles were only prima facie, then moral argument could not be purely deductive, for where there is a conflict between prima facie principles, there is ex hypothesi no principle to settle the conflict, and so the deduction of the actual duty in such a case is impossible. In the case described above whether to lie to the person with a terminal illness, we may readily admit the prima facie principles "Lying is wrong," and "One should be beneficent," but if such principles are the best we can

do, principles are clearly of little help in deciding this or any other significant moral question.

It might be thought that though prima facie principles would be of no use in a deductive form or moral argument, they might be the elements in an inductive form of moral argument. Each application of a prima facie principle, one might suppose, identifies a piece of evidence, some perhaps more significant or "weighty" than others. In deciding whether a particular act is morally right, we assemble the evidence on both sides of the issue by identifying all the prima facie principles that apply. For example, in deciding whether to lie to a terminal patient who has expressed a desire to know the truth, we have on the one side the principles "Lying is wrong," and "One should respect the self-regarding desires of rational persons," and on the other side the principle "One should be beneficent," if we think that telling him the truth will harm him. Whichever side of the question is favored by the preponderance of evidence is what we should do. This is no different than what jurors do in making a factual determination of whether someone committed a crime, when they have to weigh the conflicting testimony of defense and prosecution witnesses. If what the jurors are doing in arriving at their factual conclusion is properly called "inductive" argument, then so

is the use of prima facie principles in deciding moral issues.

This argument is certainly correct that deciding moral questions often requires weighing conflicting moral considerations; and prima facie principles identify these considerations. It's also true that many times factual issues require weighing or comparing conflicting evidence. Only the latter process, however, could be called the making of an inductive argument.

One way in which inductive arguments differ from deductive is that only in the former can evidence accumulate. When one has provided a valid deductive argument with true premises, one has done as much as possible to support the conclusion. Any additional facts are superfluous to that purpose, since the "weight of evidence" of a sound deductive argument is absolute. With inductive reasoning, on the other hand, it is always possible to strengthen the case for the conclusion by bringing in additional relevant facts. If a person is seen fleeing the scene of a fatal shooting, that is a reason to suspect him as the killer. If he is discovered to be carrying a gun, the case against him is strengthened. And if ballistics tests show that gun to be the murder weapon, we arraign him on a charge of murder.

If prima facie moral principles, then, are parts of inductive moral arguments, they too must be the sort

of evidence that is cumulative. But they are not that sort of evidence. If I have made a promise to do X, then the prima facie principle "Always keep your promises" tells me I should do X. If doing X will benefit someone, then the prima facie principle "You ought to benefit others" also tells me I should do X. But though the second principle indicates another reason for doing X, it does not make it more likely that doing X is my actual duty, as each additional piece of evidence in the murder case makes it more likely that a particular suspect committed the crime. If there are no contrary indications, then I need only know that one of these prima facie principles applies to know with a certainty what my actual duty is. If I know that, other things being equal, I should always keep my promises, then knowing that keeping this promise will be beneficial does not make me any more certain that I should keep it.

It might be objected that though the process of identifying the prima facie principles on each side of a question does not itself make for inductive arguments, the decision between conflicting prima facie principles must require inductive reasoning, since there is no necessary or analytical relation between those principles and the conclusion we reach in determining what our actual duty is. But this is an unacceptable mode of argument, since it takes "inductive" simply to mean "not deductive,"

blurring the real distinctions between non-deductive forms of argument, and arbitrarily stipulating the truth of the Humean doctrine that induction and deduction together exhaust the forms of legitimate argument.

If the claim that moral argument must be purely deductive remains untouched by the two sorts of objections mentioned so far, what remains to be said against it? There is still another obstacle, I believe, in the way of the view that moral reasoning proceeds deductively through the application of principles, an obstacle I would like to systematically develop.

Broadly speaking, it is that any principle which is plausibly thought to be exceptionless will be vague in its criteria of application, so that the question of whether such a true moral principle applies to the case at hand will be problematic, and incapable of resolution by appeal to moral principles. If this is so, then reasoning about moral problems must involve more than the appeal to principles, and some form of reasoning in addition to deduction.

This does not mean that principles are dispensable, or that it is impossible to construct valid deductive arguments that use them. The schema

1. All X are right
2. a is an X
- ∴ 3. a is right



might still adequately characterize some part of moral argument (though a relatively insignificant part, it turns out). My thesis is that the terms that comprise "X" in any significant principle in the form of (1) are vague, with no clear criteria of application. Thus, whether (2) is true and hence also whether an argument like the above can settle the issue before us, turns entirely on how the major premise can be interpreted and shown to apply to a, the case at hand.

The nature of the vagueness of moral principles will develop quite readily from an examination of both utilitarian and non-utilitarian attempts at providing systematic moral principles. Each of these will be seen to be infected with this problem in fundamental ways, and the discussions will provide illustrations of how pervasive the problem of vagueness is in moral argument.

Examination of these systems will also offer some lessons on the source of the difficulty and its implications for moral argument, which I will develop in the course of a general theoretical discussion of the problem. In that chapter I will suggest that the use of moral principles must be supplemented by a sort of argument by analogy, in which certain kinds of paradigmatic moral judgments, or cases, are compared with the case which is at issue, and to which a moral principle is to be applied.

I am not now prepared to give a detailed account of the nature of this analogical reasoning, or the rules to which it must conform. That must be the subject of further study. What I can do at this point is show that such reasoning is used in discussion of moral problems, and that it is both unavoidable and persuasive. To this end, I will present and elaborate arguments concerning two particular moral problems as examples of analogical reasoning at work.

## CHAPTER II

### CONSTITUTIVE PRINCIPLES AS A LOGICAL NECESSITY

As a first step in the investigation of the role of principles in moral reasoning, it is necessary to dispose straightaway of the notion that moral language itself implies that principles are the single foundation upon which moral knowledge is based, and that therefore the suggestion that there must be more to moral reasoning than appeals to principle is not worth pursuing.

I am thinking of several arguments by R.M. Hare, at least one of which has become a stock-in-trade of moral philosophers, which purport to establish first of all that we in some sense "need" principles and, more fundamentally, that there must as a matter of necessity be some principle(s) at the bottom of our moral judgments and practices.<sup>1</sup> Hare does not take himself in these arguments to be establishing any particular principle as the foundation of moral judgment (although he seems to think it's the principle of utility--see Chapter VIII), but only the abstract proposition that such a principle exists.

Hare offers three arguments concerning the need for principles, but only one of them bears directly on the claim the moral principles are epistemically necessary, so I will just touch on the other two briefly.

This first is that principles are a necessary ingredient in teaching morality, a process that could not proceed efficiently or to good effect unless what was taught were rules applicable to more than one sort of case.

The second is that since moral judgments depend at least in part on evaluations of the consequences of different courses of action, and since we cannot ever predict the future with certainty, we are better advised not to attempt to make independent case-by-case evaluations, but to guide our action in accord with well-established principles. A variation of this same argument appeals not to the uncertainty of predictions, but to the weakness of human will: it is better that people act on general moral principles rather than on the assumption that moral situations are complex and ambiguous, since the former is more likely to produce right action than the latter, which allows too much leeway for the powerful forces of narrow self-interest, given the degree of interpretation, analysis, and individual judgment that it would require.

As Bernard Rosen has ably pointed out, neither of these arguments establishes the necessity of constitutive

principles<sup>2</sup> (that is, principles which are to serve as the logical basis of moral evaluation). The good purposes which the use of principles serve in teaching or in promoting the best behavior from the average man can be just as well achieved by the use of summary rules, which are taken as more or less reliable guides to behavior, not as the logical foundation of moral evaluation. Neither argument has anything to say about how the principles which are to be used are themselves justified--it may be that they are summaries of past judgments about particular cases. Indeed, both of them allow the possibility that the principles used are not precisely true and unexceptionable.

To the extent that Hare thinks these arguments establish the necessity of constitutive principles, it can only be because of a confusion between two sorts of claims: (1) that principles are useful, or even that it would be wrong not to use them; and (2) that principles are themselves the ultimate ground for judgments of right and wrong. The first does not entail the second. We might of course promote and teach principles because doing so has the most morally desirable results, but this reasoning presupposes only that we have some beliefs about what constitutes "morally desirable" behavior, and nothing about whether principles, or principles alone, were or must be used to arrive at those beliefs. Hare's first

two arguments bear only on the first of these claims, and so have no implication for the second of them.

Hare's third argument is much more slippery than the first two, and a great deal more persuasive. As a simple requirement of consistency, Hare tells us, it follows from any particular moral judgment that some more general moral judgment or principle is true. If, for instance, I say of an action that it is "right," but refuse to call "right" a second action which I admit is exactly like the first, then I am inconsistent. "This is right," then, implies "Everything exactly like this in every respect is right." Now since no two things are exactly alike, we may safely say that

- (a) "This is right" implies "Everything like this in relevant respects is right."

This has not yet given us a "principle" properly so-called, Hare goes on, since it contains the individuating term "this." But we may readily eliminate this term by specifying what the "relevant respects" are, which yields the universal principle

- (b) Everything with features a...n is right.

Bernard Rosen addresses this argument also, but it seems to me not as successfully as the others. His objection is simply that we know of no unexceptionable principles. In the past, any proposed principles have been refuted with counter-examples, and so as a matter of

experience we should conclude that there are no unexceptionable principles.

But this objection does not touch Hare's conclusion, for Hare's argument is an a priori one purporting to show that such principles must exist, as a matter of logical necessity. It will hardly do to argue against this from the merely a posteriori claim that as a matter of fact we have not yet discovered them, for Hare's argument is designed to prove that they exist nonetheless, even if, as Hare admits, it may be a matter of the most difficult sort specifying what the "relevant respects" are. At the most, Rosen's argument might prompt us to take a closer look at Hare's reasoning, to probe for possible flaws.

We may begin this reevaluation by noting that strictly speaking, the inference in (a) is not one required by consistency. Consistency in making the claim "This is right" requires only the admission that "Everything exactly like this is right," for there is nothing inconsistent in saying of two actions that they are alike in all other respects, except that one is "right," and the other not. These judgments do not imply any statement of the form "This is both right and not-right," unless "right" is defined naturalistically, to mean "has properties a...n."

Of course, one need not be a naturalist to think that moral "properties" are, in Moore's expression,

"supervenient." But this claim has always seemed to me (and I will not argue for it here) to be nothing more than the assertion that moral language is inherently rational--if I make a moral judgment, I must be prepared to give reasons that refer to some properties possessed by that which is judged. It is together with the assumption that moral language is rational that different moral judgments about exactly similar acts become inconsistent, since whatever reason I would offer for calling the one "right" would apply with equal force to the other.

It is this assumption, then, that authorizes the inference in (a) to "everything like this is relevant respects is right," for a property is "relevant," after all, precisely because it constitutes a reason for calling a thing "right." The consequent in (a) may now be translated to "everything of which I may affirm the same reasons as my reasons for calling this right, is also right."

The abstract principle (b) then becomes

(b') Everything to which reasons A...N apply is right.

This third argument of Hare's, remember, is supposed to prove that behind any particular judgment of rightness must lie a substantive moral judgment of the form "All so-and-so's are right." But once we have determined what (b) means, we no longer have as our conclusion such a schema for a moral principle. (b')



is silent on the question of what sorts of propositions are to count as the reasons for making the judgment-- they may take the form of constitutive principles, but (b') does not require that they do.

Hare's third argument cannot be expected to demonstrate any conclusion about the role of constitutive principles. The argument, after all, simply works through the tautological principle that any rational system requires the consistent application of reasons, a principle that applies to any mode of reasoning, whatever its form or content. The question of constitutive principles is not a question about consistency. What we want to know about principles is not "If we use principles as the only court of appeal in moral reasoning, should we use them wherever applicable?"--that would be quite pointless to ask. Rather, what we want to know is whether we really must use only principles when we give reasons for moral judgments.

Hare's argument demonstrates the necessity of constitutive moral principles only on the assumption that moral reasoning is exclusively deductive. Once that assumption is made, of course, it follows without further argument that universal principles are the evidence on which our moral conclusions must rest, and that principles are the only reasons to which one may appeal.

This argument, then, along with variations on it,<sup>3</sup> shows neither that moral argument requires constitutive principles, nor that it is necessarily or exclusively deductive in form.

## CHAPTER III

### THE SPELL OF THE DEDUCTIVE SYSTEM

As we saw in the last chapter, Hare's argument that principles are the medium through which moral reasoning proceeds itself presupposes that deduction is the only proper mode of reasoning in ethics.

But there is a great obstacle in the way of supposing that moral principles are the foundation of a deductive system of moral reasoning, and that is that the commonly accepted moral principles each incorporate different considerations as relevant to moral decision making. This carries the possibility that in particular situations of choice these different principles will pull us in different directions, and so they will issue no clear directive as to what is to be done. Most interesting and controversial moral questions, it seems, are of this character--as, for example, when as parents we wonder whether we should subject our child to even some small risk by allowing him to participate in a medical experiment that promises to yield knowledge that will benefit others.

If such a moral problem is to be handled within the confines of a deductive system of principles, that system must satisfy two related conditions. First, the principles must all be consistent with one another, not issuing conflicting solutions to the same moral question. In the above example, there seems, at least on the face of it, a conflict between a principle of benevolence obligating us to help others when that is possible, and the duty of parents to look after the welfare of their children. Of course, this example gives rise to such a conflict only because both principles are thought to equally apply to the circumstances--e.g., exposing a child to a statistically small risk not for his own benefit does indeed violate a parent's duty to look after his child's welfare. Thus the second requirement of a deductive moral system is that its principles not be interpretable in such a way that would allow conflicting, but equally valid, applications of them. It is actually this second requirement that is most central. There is no problem in constructing a system of principles which, as propositions, are consistent with one another. The difficulty is not so much reconciling universal principles as it is controlling interpretations or applications of them to particular cases which permit contradictory conclusions to be drawn about those particular issues.

Now if one can devise a moral system whose principles do not conflict in application, then this difficulty no longer stands as an impediment to the view that moral reasoning is deductive, and it might then seem idle to suppose that moral principles were an inadequate foundation for practical decision-making. Even if an a priori argument like Hare's does not work, the successful defense of a coherent system of moral principles would go a long way toward supporting the view that moral reasoning is exclusively deductive.

Alan Donagan, in The Theory of Morality,<sup>1</sup> claims that the principles of "common morality" (the morality of the Judeo-Christian tradition) form such a system, a system that can be derived quite strictly from a single fundamental principle: "It is impermissible not to respect every human being, oneself or any other, as a rational creature."<sup>2</sup>

The precepts so derived need not resolve every moral perplexity that might confront us, he claims. Suppose, for example, that I lie to a friend about my income in order to get him to cosign a loan. If in answer to a subsequent inquiry from the bank, I tell the truth, I will do my friend a grave injustice, for the loan, which I cannot repay, will immediately fall due. Such a situation is a perplexity secundum quid,

"conditional upon some misdeed,"<sup>3</sup> which represents no conflict within the moral system itself, but arises only because the system was violated.

Rather, "a moral system is inconsistent only if it allows the possibility that, without any wrongdoing on his part, a man may find himself in a situation in which he can only escape doing one wrong by doing another."<sup>4</sup> This is what Donagan calls perplexity simpliciter, and what I have described as being pulled in different directions by two equally applicable but irreconcilable principles.

Can the precepts of common morality give rise to perplexity simpliciter? It might seem, especially, that there could be conflicts between those precepts that prohibit an action as violating the respect owed human beings as rational creatures, and those which command policies for the promotion of human good, as for example when we are faced with the choice whether to lie to someone "for his own good." Yet these precepts themselves give us no method or principle by which to choose between them.

Donagan, though, argues that between such "perfect" and "imperfect" duties there can be no conflict. The "Pauline principle," which is part of the system itself derived from the fundamental principle, commands that we are to "Do no evil that good may come."<sup>5</sup> The system, that is, provides its own ordering device, which

permits the promotion of human good only by means of "morally permissible" actions--those that do not violate the fundamental principle. Thus, the system is by its nature consistent, and can provide unequivocal guidance in every situation of moral choice.

One way to attack this solution to the problem of conflicting principles is to attack the Pauline principle directly, by claiming that it is false. Thus, one might say that while it is true that it is wrong to trade gross evils for some good (e.g., killing some innocent to save many lives), it is not so clear that we should not trade a small evil for a great good (e.g., lying to save many lives).<sup>6</sup>

But Donagan has a way of insulating the Pauline principle from this kind of counterexample. As we shall see, what counts as "doing an evil" depends on how one interprets the requirement in the fundamental principle to "respect all human beings as rational creatures." In what follows, I want to show that the prohibition contained in the fundamental principle is interpreted by appeal to a priori convictions about particular sorts of cases, whether they involve the violation of "respect" or not, and that this has unfavorable consequences for the claim that the principles of common morality form a "deductive system."

Donagan's argument that the Pauline principle is a consequence of the fundamental principle goes as follows.<sup>7</sup> Implicit in every imperfect duty (duties to promote a general end, like human good) that is validly derivable from the fundamental principle is the condition that they be discharged through "morally permissible action"; that they not require the violation of the fundamental principle--i.e., the violation of any perfect duty validly derivable from that principle. Therefore, the Pauline principle is "structurally necessary," and is a consequence of the fundamental principle itself.

We may agree that something like the Pauline principle is structurally necessary. If the imperfect duties are entailed by a fundamental principle that is categorical, then it follows by definition of "entailed" and "fundamental" and "categorical" that the imperfect duties cannot require that the fundamental principle be violated, for if F entails D, D cannot entail not-F. But one may make the same observation about the perfect duties, yet Donagan wouldn't think that we had therefore proven that the perfect duties were subordinated to the imperfect ones, and that "The ends justify the means" was thus "structurally necessary." Neither, then, can we show by such an argument that the Pauline principle is entailed by the fundamental principle. This can be



established only by working through the implications of the fundamental principle, specifically concerning ourselves with the meaning and criteria of application for the notion "respect for...rational creatures," to see whether, indeed, there were possible situations in which respect demanded that we pursue the general welfare at the expense of some individual's interests that "respect for rational creatures" also required us to protect. Only when we have shown that there are no such situations, and that "respect" always requires the protection of certain individual interests against the claim of social benefit, will we have shown that the Pauline principle is entailed by the fundamental principle. And it's only after we've thus derived the Pauline principle, or some analog of it, that we can make the claim that the system founded on the fundamental principle is a consistent one and legitimately deductive in nature. The consistency and deductive character of the system of common morality that Donagan argues for, then, is crucially dependent on the way in which the fundamental principle is interpreted.

If we wanted consistency within the system of principles badly enough, we could get it easily, for whenever we come across a case where there seems a conflict of principles, we may modify or interpret one or

the other in such a way that the conflict disappears. So we can, whenever we find a conflict between the obligation to promote good, and the respect due others, declare that respect for others is to dominate; and also, when we find a situation where the good of others requires a sacrifice of some individual interest, we may declare that sacrificing this interest shows a lack of respect for the individual as a rational creature. Thus we can, in the process of specifying what the fundamental principle means, make it fundamental and categorical, and make the Pauline principle a true consequence of it.

The consistency that would result, however, is not the consistency of any system, for the consistency is not the natural product of our principles, but our principles and their meaning are a product of the demand for consistency. That is, we haven't first provided and defended an analysis of the fundamental and subordinate principles and the concepts which they employ, and then determined whether the implications of such a system are consistent. Rather, the goal of consistency has been pursued at any cost, without regard for whether the principles and the concepts in them have been left with any coherent sense at all. This could not properly be called a "deductive system" at all, but would instead be a shapeless collection of meaningless principles the

interpretation of which shifted depending on whatever conclusion one wanted to draw from them.

No one, of course, would try to argue for a system of principles in this grotesque fashion; certainly not Donagan. The point of the exaggeration is that when one's understanding of what is implied by a set of principles is conditioned by prior convictions on what conclusions should be supported by the principles, the claim that the principles "imply" those conclusions amounts to little more than the assertion that  $p$  implies  $p$ . This is a mistake that can be made in more subtle forms, as we'll see by examining some of the "derivations" Donagan makes from the fundamental principle.

Take first his derivation of the precept allowing self-defense as an exception to the general prohibition of killing:

Since respect for human beings as rational creatures entails, in general, treating every normal adult as responsible for the conduct of his own affairs, to interfere by force with anybody else's conduct of his life, unless there is a special and adequate reason, is not to respect him as a rational creature.<sup>8</sup>

Note first that the only "special and adequate reason" consistent with the fundamental principle is that the person who is the subject of force is not a rational creature, responsible for his own conduct. Yet the exception Donagan wants to allow is killing in self-

defense. One may kill those who are not "materially innocent"--i.e., who are attackers, whether they are blameworthy or not.

But what possible implication does material guilt have for being a rational creature, all of whom we are enjoined to respect and each of whom is responsible for the conduct of his affairs? He does not say. But if killing in self-defense is a legitimate exception to the prohibition of force, then it must be justifiable in terms of the very same fundamental principle that prohibits the use of force in general. It's difficult to see how this exception can be explained, for the person attacking me may well be a "rational creature," "responsible for the conduct of his own affairs" (i.e., capable of the rational pursuit of his interests). Of course, it may be "irrational" or "wrong" not to permit self-defense as an exception, but that is what the fundamental principle is supposed to establish. If all we can say in defense of our interpretation of the fundamental principle as allowing self-defense is that self-defense is morally permissible, then there is no meaningful sense in which the exception of self-defense has been "derived" from the fundamental principle.

The same circularity arises in his derivation of the principle of veracity, where he quotes with approval

Kant's solution to the problem whether to lie to a would-be murderer who demands to know where his intended victim is. We may lie, he says, since the criminal "knows full well that (you) will not, if (you) can help it, tell him the truth and that he has no right to demand it of (you)."<sup>9</sup>

It must first be said that this is no solution, since (1) the fact that we would lie does not yet address the question of whether we should; and (2) claiming that the criminal "has no right to demand it" begs the question at issue, for if we should tell him the truth, then he does have the right to demand it.

More to the present point, it is also difficult to see what the connection (or lack of one) is when Donagan says "the principle of respect for man as a rational creature does not require that the truth be told in such a case."<sup>10</sup> But why not? The criminal in such a case may be immoral, and we may well have no obligation to tell him the truth, but unless he is certifiably insane, he is as "rational" as anybody else (possessed of the "practical reason" which Donagan thinks provides the grounds for the fundamental principle).<sup>11</sup>

Donagan might reply that such a person is "not responsible," and so exempted under the language of the principle of veracity, which is taken to hold only between "responsible persons."<sup>12</sup> But far from establishing that respect for rational creatures permits us to lie to the

criminal, this reply exposes the principle of veracity as a precept not derivable from the fundamental principle. It might plausibly be said to be derivable if one takes the expression "responsible persons" to have the meaning which would allow that children and madmen are not due the truth--they are "non-responsible" because they are non- or ir- rational, and for that reason not protected by the fundamental principle. But a sane adult criminal who is "not responsible" is so because he is irresponsible, not because he is non-rational, non-responsible, or blameless. To try to allow beneficial lies to him under the same cover as beneficial lies to children and madmen is to ignore this obvious distinction.

The exception of the irresponsible from the principle of veracity couldn't possibly be a matter of derivation from the fundamental principle. If it were, we should have to be able to say, "Respect for rational creatures doesn't require that one aid wrongdoers by telling them the truth." But then a knowledge of what constitutes wrong-doing will have been made a condition of understanding the notion "respect for rational creatures," which on the contrary is claimed to be the standard by which wrong-doing is judged.

One might suspect that the only reason there is for this exception to the principle of veracity is that

by lying to such a criminal one prevents a great harm from occurring. That is, that here one is doing an evil (not respecting a rational creature by lying to him) so that a good may come, which would be a violation of the Pauline principle mentioned earlier. A more accurate description though is that the perception that telling the truth in such a case would cause a great harm is by itself the reason that telling a lie to a potential murderer is not what is required by the "respect" due rational creatures, and so is not counted as an evil. In this way, the fundamental principle is interpreted to accord with our moral intuitions in cases in which we seem to be preventing harm or pursuing benefit at the expense of some other moral consideration. Thus, the sort of counter-example offered by Gert of doing a "small evil" (telling a lie) to prevent a great harm (a loss of lives) would be cheerfully deflected by Donagan. Telling the lie does not count as an evil, because not prohibited by the fundamental principle's requirement that we "respect rational creatures."

This process can be seen most clearly in his discussion of the prohibition of suicide.<sup>13</sup> After urging that to take one's life "at will" (for no reason) is clearly a violation of the respect due rational creatures, he tells us that nonetheless there are cases where suicide is justified; as in the

example of an overloaded boat, which will sink unless some of its load is jettisoned but the entire load of which is innocent human beings. If nobody can be saved unless somebody goes overboard, and if to go overboard would be suicide, then suicide in such circumstances would certainly not be contrary to the respect due to humanity as such.

This can only be because the respect due to rational creatures permits one to judge that his life has less value than the aggregate value of the lives of those saved by his suicide. Presumably, though, his life in such circumstances is not valueless, so the taking of it is an evil, which is acceptable despite the prima facie prohibition of suicide, because of the good that results. Here then we have the example of a precept which is interpreted so as to admit the doing of an evil so that good may come.

It might be objected that in such a case, suicide is no evil, since in such circumstances it is not wrong to do. But if "evil" in the Pauline principle is taken to mean "wrong to do," then the Pauline principle is true but vacuous, since of course if it is wrong to do something, it is wrong to do, regardless of its good consequences (since they have already been taken into account in deciding whether it is wrong to do).

Strangely enough, Donagan seems to think that there is nothing wrong with the mode of reasoning which I have illustrated and in particular that it preserves



the deductive character of his system. In Chapter 2 he tells us that deriving percepts from the fundamental principle requires "specificatory premises" which "identify a species of action as falling or not falling under the fundamental generic concept of action in which every human being is respected as a rational creature."<sup>14</sup>

After admitting that "virtually all the philosophical difficulties that are encountered in that structure (of common morality) have to do with establishing the specificatory premises,"<sup>15</sup> he claims that nevertheless they can be established with great confidence through "unformalized analytical reasoning in which some concept either in the fundamental principle or in a derived precept is applied to some new species of case."<sup>16</sup> Thus the deductive character of the system can be illustrated with a schema such as

- (1) It is impermissible not to respect every human being, oneself or any other, as a rational creature.
- (2) All actions of kind K fail to respect some human being as a rational creature.
- ∴ (3) It's never permissible to do an action of kind K.

There's no question that the above is a deductively valid argument. But deductively valid arguments, even when sound, don't necessarily constitute a reason for their conclusions. As we've seen, establishing the second premise (identifying what fails to respect

rational creatures) depends largely on settled intuitions on what's permissible in situations of kind K; that is, depends on having the conclusion already in hand.

Such a system could maintain its integrity if premise (2) could be established independently of the conclusion (3). But in order for this to be possible, the specificatory premises must first be established through an analysis of "respect for rational creatures," and that concept must itself be a non-moral one. But Donagan offers no sustained analysis of "respect for rational persons" that could possibly tie together all the specificatory premises that are used, and as the discussion has shown, the notion of "respect" is explicated through the use of our settled moral convictions surrounding suicide, lying, self-defense, etc., which dashes the idea that "respect" is a non-moral concept.

Donagan himself seems to admit this when, in explaining how the concept of respect for rational creatures is understood, he says

...among those who share in the life of a culture in which the Hebrew-Christian moral tradition is accepted, the concept is in large measure understood in itself; and it is connected with numerous applications, as to the different weights of which there is some measure of agreement.<sup>17</sup>

I fully agree that such a concept can be understood only against a background of considered judgments

that serve to interpret it for us, and shall be arguing throughout that such vague moral concepts pervade moral principles and require appeals to our settled convictions. But if this is true, it poses a dilemma for the view that moral argument consists in making a deduction from a moral principle. Every term in the principle that is of moral significance is either more or less vague. The more vague each term is, the less the principle serves to explain or justify moral judgments, for the more difficult it is to know what is supposed to fall under it or how it is to be applied; its criteria of application are problematic. When, on the other hand, we attempt to specify what the principle permits or prohibits in order to eliminate the vagueness of its terms, the less the principle explains or justifies moral judgments, since (1) its meaning becomes a function of convictions about what particular sorts of actions to permit or prohibit, which in turn (2) robs the principle of its desired generality, on which claims of its being fundamental or foundational depend, and threatens the deduction with circularity.

This will be the theme to which I will constantly return, as the difficulty it represents is a pervasive one that infects many other attempts to argue for and use fundamental or general moral principles.

## CHAPTER IV

### THE PRINCIPLE OF UTILITY

In the last chapter, I showed how Donagan's attempt to specify a fundamental moral principle from which we could deduce the right course of action in any circumstance was plagued by problems of vagueness in some of the moral concepts it used, such as "respect for rational creatures," which seemed to require that one already have convictions in particular cases, on what was or was not an instance of the violation of such respect.

One of Donagan's motivations in attempting to establish such a deductive system is to offer an alternative to utilitarianism. Utilitarianism also claims that all moral judgments are ultimately answerable to a single principle--to maximize the good (generally conceived as happiness), doing that which is most conducive to the welfare of mankind. Thus, it claims to present us with a coherent system that ties together in one "principle of utility" all of our various moral judgments, and is able, so it is claimed, to provide us with

unambiguous guidance when we put to it questions about what we should do in particular circumstances.

Although being systematic in this way is not sufficient for being an acceptable moral theory, neither is it a small virtue. Non-utilitarians have accordingly felt the force of utilitarianism's claim to being systematic, and have felt chastened by the charge that non-utilitarian moral theories are a jumble of potentially conflicting principles, which can't support coherent justifications of moral judgments. It is felt that unless one can offer a systematic alternative to utilitarianism, then utilitarianism stands head and shoulders above all other contenders. It is for this reason that Donagan and John Rawls (to be examined in a later chapter) set about the attempt to explicate all of morality, or a significant portion of it, in terms of some single principle or moral concept.

If the problem of vagueness that I am developing attends all such attempts, it might seem to add a new obstacle to the development of non-utilitarianism. But no such support will be forthcoming, I'm going to argue, because utilitarianism is infected with the very same trouble; even the principle of utility, often pictured as the model of clarity for moral principles, is vague and must be interpreted through the use of considered judgments.

There are, as is well-known, different varieties of utilitarianism, according to whether one concentrates on the utility of individual acts, or that of general rules; or whether one interprets the "welfare of mankind" as the sum total of happiness, or as that total divided by the number of persons (which yields "average utility"). I am unconcerned with these doctrinal differences, as the critique I will offer is quite general and unaffected by these distinctions. I will, therefore, consider utilitarianism in its most straightforward form, as act utilitarianism.

What, then, is the act utilitarian principle? J.C.C. Smart asserts that for the act utilitarian, "... the only reason for performing an action A rather than an alternative action B is that doing A will make mankind (or perhaps, all sentient beings) happier than will doing B."<sup>1</sup>

For my purposes, the crucial phrase in this principle is "will make mankind happier." What does this require of us? Suppose that we know all of the consequences of action A for the happiness of every person, and all the consequences of the alternative B. What procedure do we follow in determining which alternative produces the greater happiness for mankind? What, in other words, is the criterion of application for the expression "will make mankind happier?"

The "obvious" answer that is always offered by utilitarians is that we assess the relative amounts of happiness produced by the total situation; i.e., considering the effects on the happiness of each person affected. Not very helpful yet, since we don't know how that "consideration" is supposed to proceed. Some utilitarians allow that we may assign numerical values to the effects of each individual's happiness, and then sum up over all persons who are affected: the numerical total will represent the total happiness to be produced by each alternative. Even those, such as Smart, who shy away from the claim that we can assign numerical weights in this way, agree that the "happiness of mankind" is a function of the total situation in the sense of being conceptually independent of disparities between the degrees of happiness of the different individuals who make up the total situation. Smart allows, for example, that "sending four equally worthy and intelligent boys to a medium-grade (English) public school" might create less total happiness than "leaving three in an adequate but uninspiring grammar school and sending one to Eton."<sup>2</sup> Smart's example loads the dice somewhat, however, since it comforts us with the thought that the three boys won't suffer any worse fate than an "adequate" education. In fact, so far as this conception of the "happiness of

mankind" is concerned, there's nothing to prevent acquiring a greater total happiness by making some people very unhappy, so long as doing so produces a weighty enough benefit for others. If, for example, we subject extremely premature non-viable infants to aggressive "treatment" (vain treatment, in their cases), we may learn how to preserve such infants in the future, contributing positively to the total happiness, since the suffering of the present-day infants and their parents comes to an end; but the benefits to future infants and their parents continue to add up indefinitely.<sup>3</sup> Utilitarians sometimes engage in sophisticated flanking maneuvers around such examples, designed to assure us that in the real world such action would not contribute to the total happiness. The essential point remains, however, that on the usual utilitarian interpretation of "making mankind happier" it is perfectly possible in principle that the actions as described would have a positive effect on the welfare of mankind, and so be morally right.

We still do not know how we get from our knowledge of the effects of each alternative to a decision about which best promotes the happiness of mankind. Smart and others<sup>4</sup> suggest that all we can do is become clear about the facts, and then choose which we prefer. It can't be



mankind" is concerned, there's nothing to prevent acquiring a greater total happiness by making some people very unhappy, so long as doing so produces a weighty enough benefit for others. If, for example, we subject extremely premature non-viable infants to aggressive "treatment" (vain treatment, in their cases), we may learn how to preserve such infants in the future, contributing positively to the total happiness, since the suffering of the present-day infants and their parents comes to an end; but the benefits to future infants and their parents continue to add up indefinitely.<sup>3</sup> Utilitarians sometimes engage in sophisticated flanking maneuvers around such examples, designed to assure us that in the real world such action would not contribute to the total happiness. The essential point remains, however, that on the usual utilitarian interpretation of "making mankind happier" it is perfectly possible in principle that the actions as described would have a positive effect on the welfare of mankind, and so be morally right.

We still do not know how we get from our knowledge of the effects of each alternative to a decision about which best promotes the happiness of mankind. Smart and others<sup>4</sup> suggest that all we can do is become clear about the facts, and then choose which we prefer. It can't be

we have the same understanding of the concept "the happiness of mankind" only to the extent to which we agree, in particular cases, that such-and-such best promotes the "happiness of mankind," or is most "benevolent." The principle of utility would require interpretation through the use of our considered judgments surrounding these concepts.

To leave the matter here, and conclude that the principle of utilities is on a par with all other principles in the necessity of appealing to considered judgments would be misleading. The principle of utility requires an appeal to considered judgments alright, but they're not those connected with "benevolence" and "sympathy."

To understand the nature of the considered judgments utilitarianism does require, let's begin by taking a closer look at the alleged link between benevolence and the principle of utility.

Smart urges that the choice that benevolence should make concerning the happiness of mankind is analogous to the choice that a rational individual would make concerning his own welfare:

...if it is rational for me to choose the pain of a visit to the dentist in order to prevent the pain of a toothache, why is it not rational of me to choose a pain for Jones, similar to that of my visit to the dentist, if that is the only way in which I can prevent a pain, equal to that of my toothache, for Robinson?<sup>5</sup>

But the demands of benevolence are not compatible with the utilitarian conception of the "happiness of mankind," and the requirements of rationality in making prudential choices are nor mirrored in the making of benevolent choices. Benevolence is the inclination to do good for others; for the utilitarian, this means to contribute to their happiness. Benevolence arises out of sympathy, which is the capacity to imaginatively appreciate the happiness and sorrow in the experience of others, just as one feels it in his own. One is sympathetic and benevolent, then, toward other subjects of experience like oneself, of whom it is appropriate to say they are happy or sad, feeling pleasure, or feeling pain.

Yet the notion of "the happiness of mankind" requires that benevolence take as its object something other than a subject of experience, for the good to be considered is that of all mankind, the happiness of the aggregate which is made up of experiencing subjects, but which itself feels no happiness. There is, strictly speaking, no such thing as the happiness of "mankind" as a class, and so benevolence cannot conceivably require that such a good be maximized. The only happiness is that of individual persons, and so it is with individual persons that benevolence concerns itself. (Someone might

suggest that "benevolence" just means "concern with the happiness of mankind," but this both does violence to ordinary language, and makes "benevolence" a technical term of the utilitarian. This in turn would render it useless as an explanation of the notion "the happiness of mankind".)

This means that certain sorts of inequities in the distribution of happiness are not endorsed by benevolence, even though they are part of the utilitarian concept of the "happiness of mankind." Specifically ruled out by benevolence are actions which would place a burden on an individual or group in return for a significantly smaller burden or a relatively insignificant benefit for a larger number of individuals. For example, the notion of the "happiness of mankind" admits in principle the possibility that the tortured death of one individual before many thousands of happily cheering spectators would contribute to the happiness of mankind, for those many thousand small bits of happiness might outweigh the admittedly great suffering of the single person.<sup>6</sup> This would be an act of benevolence only if one could think of "mankind" as an experiencing subject, who was enjoying this vast happiness--but "mankind," of course, is no such thing.

The fact that the abstract class "mankind" is not a subject of happiness, and so not an object of benevolence, brings out as well the difference between a rational prudence and a rational benevolence. An individual might rationally decide to take on a great suffering now in order to avoid a series of smaller sufferings in the future--e.g., he might agree to have all his wisdom teeth taken out at once instead of one at a time over a period of years. And we might, out of benevolence, make such a decision on behalf of an individual--a small child, say. But it doesn't follow that rational benevolence would require us to prefer the suffering of one individual equal to that of having all four teeth pulled to the suffering of four persons, each of whom suffers the equivalent of having one tooth pulled. There is no subject--no being called "mankind"--in whose experience the greater suffering is redeemed, and who consequently enjoys the greater measure of happiness which is the object of benevolence.

What this means is that applying the concept of the "happiness of mankind" requires that we entertain the fiction that mankind is an experiencing subject just like the individual persons that make up mankind. Asking which of alternative actions would produce the greatest total happiness, then, poses a hypothetical question:

if mankind were an individual subject of experience who would feel all the effects of each alternative, which would he prefer? The question asks that we envision the effects of each alternative abstracted from the real individuals who suffer or enjoy them, as if they were components of some one person's experience, analogously to the way in which we entertain alternatives when concerned about our own well-being, when we imagine the various effects different choices will have on us.

But then, just as in our own cases the only test of which effects are preferable for us is to decide which we prefer, so with the "happiness of mankind"--the only way to determine which alternative action is preferable is to decide which we prefer. There is no principle, or set of criteria that will decide for us, and the appeal to benevolence is quite out of place, as we have seen. There are only judgments about particular sorts of cases. For example, that building a ski area will contribute to the total happiness, even though we know that as a result people will be injured or even killed while skiing on it. We can only ask, "If I had the choice of enjoying that much pleasure at the price of that much pain (or risk of death), what would I do?" Just as particular judgments are the basis for understanding Donagan's fundamental principle and the "respect for rational creatures"

that it requires, so other particular judgments are the foundations for our understanding of the "happiness of mankind," and in turn our understanding of the utilitarian principle. We don't justify our preference for building the ski area by appealing to the principle of utility; rather we must already have some such preference, some such conviction about what's permissible to do in the name of human welfare, before we can understand what the principle of utility requires. What's different, though, about the considered judgments required by utilitarianism is that they are relativized to each of our individual systems of values and tolerances for risk-taking, which is likely to lead to much more disagreement about how the "happiness of mankind" is to be understood, making this concept even more vague and indeterminate than the others commonly found in moral principles.

## CHAPTER V

### THE FAILURE OF A PRIORI PROOFS

The pervasive vagueness of general moral principles and the subsequent problem this presents for deriving from them conclusions on particular moral questions are, it may be thought, obstacles that can be overcome by looking to the sort of justification which is offered for the principles themselves. Perhaps the considerations which are the grounds for the principle will supply an interpretation of it that provides unproblematic criteria of application.

This is a hope, however, which is unlikely to be fulfilled, as can be seen by examining an argument by Alan Gewirth that purports to show how we can derive certain significant moral conclusions from non-moral premises.<sup>1</sup>

In making his derivation, Gewirth is ultimately interested in providing the grounds for such particular moral judgments as that we should throw a rope to a drowning man if such is possible, or that we should feed one who is starving, if we have plenty of food.<sup>2</sup> These "are not self-evident," he says, and further, "if they



are to be justified at all they must be derived in some way from other statements,"<sup>3</sup> and these must be non-moral if the need for justification is to stop.

The conceptual starting point for the derivation is "the generic features of action . . . features which are necessarily exhibited by any instance of action";<sup>4</sup> namely, being voluntary and being purposive. By virtue of being voluntary, they are done out of freedom; by virtue of being purposive, they imply that the agent does them for something he considers "good" (in the broadest sense).

Once this groundwork is laid, the argument proceeds in four main steps:

1. "Because freedom and basic well-being are at least instrumentally necessary to all the agent's actions for purposes, from the agent's standpoint his statement, 'I do X for purpose E' entails not only 'X and E are good' but also 'My freedom and basic well-being are good as the necessary conditions of all my actions.'" <sup>5</sup>

2. "Since these are the necessary conditions of all his actions and hence of any purposes and goals he may attain or pursue through action . . . he regards himself as justified in having freedom and basic well-being . . . that he has a right to perform his actions and to have freedom and basic well-being." <sup>6</sup>

3. "He is logically committed to a generalization of this right-claim to all prospective agents and hence to all persons."<sup>7</sup> This is because the "sufficient condition" of his right-claim is that he is a prospective agent with purposes to fulfill, and like all reasons, this one is necessarily generalizable.

4. From this follows the "ought" judgment which is correlative to the right-claim, which "the agent is logically compelled to admit . . . on pain of contradiction." Namely, "I ought to refrain from interfering with the freedom and basic well-being of all prospective purposive agents," or "I ought to refrain from coercing other persons or inflicting basic harm on them."<sup>8</sup>

There are a number of challenges one might make to the adequacy of the argument that leads to this principle. In particular, the inference in step 2 seems suspect. If the agent need make no claims for the objectivity of his judgments of the value of his goals, but only that they are subjectively "good" in the sense of being desired or wanted, then having these goals or purposes assumes nothing about their "justification," since justification presupposes that one is entering his value judgments into the objective realm, claiming that others should agree with them.

But putting aside such challenges to the adequacy of Gewirth's derivation, we may still ask what

guidance the derived principle can offer us. I would argue that by itself, it can offer none at all.

The obligations that the principle imposes are either categorical, and never to be outweighed by other considerations; or they are only prima facie, in which case they have some claim on us, but its force is indeterminate. There is no reason to believe the principle is categorical, since in the first place there is nothing in the derivation to support the contention that these are the only obligations there are or that they are the strongest ones, and since secondly, we know that these are not the only obligations--e.g., sometimes I ought to coerce others or inflict basic harm on them, either for their own good, or for the good of others.

They must, then, be prima facie obligations. But as such they can provide no guidance, for in every situation in which we are confronted with the decision whether to coerce others or inflict basic harm on them, this principle cannot tell us whether or not, or in what circumstances, I can coerce someone in the interests of his own basic well-being, the obligations it imposes are outweighed by other factors, or even what other factors would be relevant.

As it stands, the principle is either unsupportable or false, or useless for moral decision-making.

The whole enterprise can be salvaged only if the principle can be supplied with an interpretation which indicates how it is to apply to particular cases and this is what Gewirth does. What "directly follows" from it (he says) is "the negative duty not to inflict serious gratuitous harm on other persons. There also directly follows the positive duty to perform such actions as rescuing drowning persons or feeding starving persons, especially when this can be done at relatively little cost to oneself," since ". . . to refrain from performing such actions as rescuing and feeding in the circumstances described would be to inflict basic harms on the persons in need."<sup>9</sup>

But we may seriously question whether such duties indeed "follow from" the principle, for it has nothing to say about when harm is gratuitous (i.e., unjustified!), or what counts as serious; neither does it tell us when prudential concerns may legitimately take precedence over concern for others. Of course we may, as Gewirth implies, make such conditions part of what we mean by "inflicting basic harm." But then the principle is both useless and meaningless to us unless we already know that such-and-such is gratuitous harm, or that in such-and-such circumstances one should place concern for others over concern for himself.

But if we are already sure of these more particular moral judgments, and must have them in hand before we can use and understand the general principle, then the general principle does not "entail" them, except in the trivial sense in which they can be said to "entail" themselves. Of course, we all know that one should throw a rope to a drowning person when it is safe to do so, and that we should feed the hungry, when we can do so at little or no cost to ourselves. But that is precisely the point: It is only against the background of such knowledge that we can understand what the principle means to us when it says we should not inflict basic harm on others. If we had no prior moral convictions, the principle would be quite meaningless to us.

## CHAPTER VI

### THE INSEPARABILITY OF PRINCIPLES AND CONSIDERED JUDGMENTS

What the analysis of Donagan, Gewirth, and utilitarianism suggest is that the application of principles to particular moral problems and an understanding of the very meaning and import of principles requires that we already have in hand certain particular moral beliefs which are paradigm instances of the use of each principle and which thus serve to interpret what would otherwise be vague guidance.

If this is true, it means that a moral principle is always conjoined with at least some of those particular moral "conclusions" that would be subsumed under it, those particular judgments, namely, that supply us with the principle's meaning. It would seem to follow from this that it would be impossible to justify or argue for a moral principle considered entirely apart from such particular judgments; e.g., that one could not argue for Donagan's fundamental principle enjoining "respect for all rational creatures" without assuming that lying to a

would-be murderer did not violate the "respect" due him as a rational creature.

To further illustrate and support this claim of the inseparability of principles and particular judgments, I want to examine the arguments developed by John Rawls in his deservedly famous book, A Theory of Justice (TJ).<sup>1</sup> The book is an extended argument for two fundamental principles of social justice--the principle of equal liberty and the difference principle<sup>2</sup>--and for a general conception of justice which holds that "all social primary goods--liberty and opportunity, income and wealth, and the bases of self-respect--are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the least favored."<sup>3</sup>

Rawls recognizes two methods of arguing for the two principles. One is to examine their implications for particular judgments on questions of justice (Rawls' "considered judgments"), to see if they square with our intuitive convictions, and this is what he does in Part II of his book. But Rawls is best known for his other argument for the two principles--that these would be the principles chosen by rational, mutually disinterested persons who were situated behind a "veil of ignorance" that deprives them of all knowledge of individuating particulars, either of themselves or of the class, epoch, or society of which they are members. The

choice of these principles from the "original position" (OP), Rawls thinks, provides a justification of them that is independent of their consequences for our considered judgments.<sup>4</sup> It is true that Rawls believes the principles must pass the test of "reflective equilibrium"--they must both be in rough conformance with our considered judgments, and be justified by the argument from the concept of the original position. But if the original position is to contribute any weight to its end of the balance, it must be an independent form of argument that needn't rely on an appeal to considered judgments. Rawls method of argument from the OP, then, should show how moral principles can be understood and applied without any assumptions about the truth of the considered judgments that would fall under them.

There are a variety of ways in which one might take issue with Rawls--one could claim that the two principles are false or unjust (by, for instance, ruling out communitarian ideals); or that the parties to the OP would choose something other than Rawls' two principles; or that the OP is in some way incoherent; or that the whole enterprise is question-begging. In what follows I will be arguing in essentially the last of these ways, for I will claim that in its very nature the OP assumes what it is supposed to prove: that the structure of the



OP assumes those considered judgments and unifying principles for which the choice in the OP is supposed to provide justification.

The question, then, is how does the choice of the two principles by the parties to the OP justify them as principles of distributive justice? It is not the mere fact of agreement (the fact that these are the principles that all persons so situated would choose), as both Thomas Nagel and Ronald Dworkin have pointed out.<sup>5</sup> There must, in addition, be some feature of the original position in which this agreement is reached that makes such a choice morally compelling for us.

It seems to me that there are two possible sorts of features of the OP which would give the choice of the two principles some moral force. The first would be the embodiment by the OP of purely formal criteria for justice or rightness--e.g., features of moral principles like universality and generality which Rawls discusses as "the formal constraints on the concept of right."<sup>6</sup> But such constraints on the principles to be chosen do not seem to require major characteristics of the OP like the veil of ignorance. Indeed, Rawls acknowledges that the only conceptions of "justice" ruled out by the concept of right as such are egoistic ones.

An alternative conception of the OP which would make it morally relevant would be to have it embody or

exemplify some non-formal, substantive moral claims which are, however, uncontroversial, or at least less controversial than the two principles themselves.

In his article, "The Original Position," Ronald Dworkin has attempted an explication of Rawls' argument along just such lines.<sup>7</sup> Behind the OP, he explains, lies a "deep theory," committed to the "right to equality," which is the right of all persons to "equal concern and respect." This is a "highly abstract right," whose meaning is determined by the device of the OP, being "assumed . . . in its design," and thus conditions the choice of the two principles, which therefore specify what the abstract right must entail for just social institutions. The OP, then, is not morally neutral, but the right that is assumed in it is abstract and uncontroversial enough to mediate between the competing conceptions of justice.<sup>8</sup>

In one of his few replies to critics, Rawls seems to agree with this interpretation of his argument.<sup>9</sup> The OP is not morally neutral, he tells us: "Rather, it is intended to be fair between individuals conceived as moral persons with a right to equal respect and consideration in the design of their common institutions."<sup>10</sup> This right is central in Rawls' list of what he calls "formal features of a well-ordered society," these

features setting "formal constraints" on admissible conceptions of justice, and the well-ordered society being the proper object of social organization. The description of the OP is "designed to incorporate" these features.<sup>11</sup>

It is, then, fair to say that it is this "right to equal respect and consideration" that gives the decisions made in the OP their moral relevance, and thus that it is their relation to this abstract right, as mediated by the OP, that provides support for the two principles independent of their relation to considered judgments.

Now I will argue that this right, as it is incorporated into the OP by means of the crucial veil of ignorance, cannot provide support for the two principles independent of an appeal to considered judgments. The trouble with appealing to this right as a justification for the conditions of choice in the OP is just that problem that arises for Donagan and Gewirth--general principles such as the right to equal respect and consideration are intolerably vague, since the crucial moral terms in them are neither transparent in meaning nor clearly definable, and so can only be understood against a background of considered judgments. I will try to show that to the extent that the principle is kept vague, it places no constraints on the conditions of the OP; and

to the extent that its meaning is specified, its interpretation presupposes those very principles and considered judgments that are supposed to be independently justified by the device of the OP.

The right to equal respect and consideration can be construed in two distinct ways, depending on who the "persons" are who are owed this right. For we may say that it is owed all actual persons, persons as concrete psychological and physical beings, with all their different capacities, goals, life-plans, and conceptions of the good and right. Viewed this way, there is no morally relevant difference between "persons" and "personalities," and the right is equivalent to the right of all conceptions of the good, life-plans, etc., to the equal opportunity to flourish. Or it may be owed all "moral persons," where equal respect and consideration is given not to persons in all their particularity, but in and to their capacities as moral agents. It is important to recognize that these two versions of the right are distinct claims. We may argue that the latter is necessitated by the very idea of a moral system (and is thus a formal constraint), moral systems presupposing beings capable of making rational choices out of duty or principle which may conflict with the dictates of self-interest. But the former does not follow from this, for we may respect the

capacity within persons without respecting the choices they make, for they may exercise their capacity poorly, or not at all, making choices that are morally wrong. Indeed, to respect a capacity is not per se to respect any actual person or being, for it may be that no one possesses the capacity.

Either interpretation of the right might be thought to be incorporated by the veil of ignorance. We may deny the parties all knowledge of their individuating features either (1) to insure that all actual interests, conceptions of the good, etc., will be, so far as possible, provided for by the principles and corresponding institutions which are chosen, since each of the parties must consider that he may be any possible person in any possible society--thus incorporating equal respect and consideration for all personalities; or (2) to insure that each party, in making his choice, will be making it as an equal moral person with all the others, with no undue consideration of his own advantage or preferred conception of the good.

If the OP incorporates the first version of the right, then the two principles gain no independent justification, for so construed the right is just as debatable as the principles themselves. A utilitarian, for example, would deny that all persons and their interests are due equal consideration, and for the same reason that

he would dispute the difference principle--interests are to be considered equally only insofar as consideration of them contributes to increased utility. The right to equal respect and consideration of actual persons implies a prima facie obligation to treat all equally; but the utilitarian has no reason to give even prima facie weight to equal distribution as a moral imperative. That he would count each person's happiness equally (not leaving anyone out) does not imply that utility would be best served by making each person equally happy, and so if the OP is based on the right to equal respect and consideration of actual persons, the utilitarian would not accept such a device, as it begs the question against him.

Rawls may well agree that the right to equal respect and consideration would be question-begging if interpreted as in (1): "the notion of respect or of the inherent worth of persons is not a suitable basis for arriving at these principles. It is precisely these ideas that call for interpretation."<sup>12</sup> Interpreted as owed all actual persons, the right to equal respect and consideration is much too rich a conception, which already contains within it the general conception of justice and the exemplifying principles which Rawls hopes to establish. The general conception and the two principles may indeed be true, but they cannot be argued for in this fashion.

The second version of the right is a much more likely candidate, given Rawls' claim, cited earlier, that the right is among the "formal" requirements of a well-ordered society. Also, in both TJ and in "Fairness to Goodness" Rawls says that the right is owed "individuals conceived as moral persons."<sup>13</sup> In TJ there is some argument to the effect that one need only have the potential to be a moral person to have the right; one need not have a fully developed moral personality.<sup>14</sup> But this complication may be passed over, since nothing I will say tramples on it, and since the right of potential moral persons is parasitic on the right of fully developed moral persons, anyway.

Well, what about the right to equal respect and consideration due all moral persons? How does the OP incorporate this right into its design? First of all, we may note that though the parties may be "moral persons" in some attenuated sense, they are not acting as moral persons, since they are not acting out of any principle other than to advance their best interests when selecting the principles of justice. Of course, the obvious reply is that we cannot have the parties acting out of moral principles when the purpose of their deliberations is to select the principles they should act out of. If my thesis of the inseparability of principles and particular

judgments is correct, however, it is impossible to argue for principles independently of some moral convictions by which those principles are given content; it will be shown that this is true of the parties to the OP.

As I said, the principle of equal respect and consideration for moral persons is built into the OP, apparently, through the device of the veil of ignorance, whose purpose, Rawls tells us, is to insure that morally irrelevant considerations do not influence the parties' deliberations; that they be able to give no undue consideration to their own advantage when selecting the principles that will govern their association. But under the "thick" veil of ignorance Rawls prescribes, consideration of any individual or group interests, life-plans, conceptions of the good, etc., is ruled out as "undue" and morally irrelevant. That is, the respect and consideration that is due moral persons is being interpreted as respect for actual persons with all their different interests, etc., for the purpose of determining the thickness of the veil of ignorance; respecting someone "as a moral person" requires, it seems, giving equal consideration to all of his interests, etc., relative to the competing interests of others.

As Dworkin says, equal respect and consideration for moral persons is an "abstract notion, and this fact



betrays the weakness in appealing to it as a justification for the conditions of choice behind the thick veil of ignorance. Its very abstractness means that it needs an interpretation if we are to have any idea of what obligations it entails for our treatment of others. It has already been pointed out that respect for the moral capacity in persons by itself entails nothing about the respect due the actual range of choices people make; nothing has been said to indicate when, if ever, people properly exercise their moral capacities, or how moral capacity reveals itself in the choices that are made, and so when the use of the thick veil of ignorance implies that "respect for moral persons" requires consideration of all interests, conceptions of the right, etc., this is an interpretation that must flow from a background of considered judgments, since it doesn't follow from an analysis of "respect for moral persons."

Rawls may well be right that justice requires equal consideration of all interests, etc., in the design of social institutions, but this cannot be well argued for by pointing to the right of moral persons to equal respect, for the desired conclusion of such an argument is what supplies the abstract right with substantive content. Unless we already accept Rawls' view of what justice requires, either we will not think that the

right appealed to implies that view, or we will simply be unsure what the right means at all.

In this respect, the right of all moral persons to equal respect and consideration is just as problematic as its ancestors, the Kantian injunctions to treat rational nature as an end-in-itself and to never use a rational creature merely as a means. Such principles seem to capture something essential to morality, but they carry no clear criteria of application, and so knowing when they have been violated runs the risk of amounting to nothing less than knowing in specific sorts of cases what it is right to do.

The fact of the matter is that specifying what counts as "equal respect and consideration" of moral persons requires an already developed conception of the right, which includes such considered judgments as "A person's sex is irrelevant to the share of social goods that person should receive," and so on through all morally irrelevant individual attributes. Under Rawls' interpretation of the right to equality of moral persons, as revealed in the construction of the veil of ignorance, this includes all individual attributes. This means, for instance, that Rawls' interpretation of the right depends on the considered judgment that "A person's talent is morally irrelevant to the share of social goods that person should receive." The situation of

choice in the OP, far from being neutral enough to decide between competing conceptions of justice (as Dworkin claims it is), begs the question against some of them.

The suggestion is that the right of moral persons to equal respect and consideration necessarily presupposes the very sorts of principles which appeal to it is supposed to establish. This happens not only in the construction of the device of the original position, but in the reasoning of the parties from behind the veil of ignorance.

Remember that the parties to the OP have no idea of what their conception of the good is (except that they want as much of the "primary goods" as they can get). Nor do they have a conception of the right. They are completely ignorant of what sorts of persons they will be once the veil is lifted--they may be saintly and benevolent, or sadistic and intolerant, or have any other combination of the innumerable moral attributes. This uncertainty is crucial to Rawls' argument for the choice of the two principles over the principle of average utility, and he cannot limit the uncertainty without jeopardizing this argument.

Thus, one of the possibilities, given the thick veil of ignorance, that each party must consider in his deliberations is that he will be a person whose conceptions of the good and the right compel him to suppress

those who are "immoral," from his point of view. Historically, this is a fairly common conception--one thinks of the various holy wars and persecutions (though intolerance need not be of the religious sort). If each party must consider this possibility alongside all the others in choosing those principles which are most likely to best promote his self-interest, for what reason could they choose the principle of equal liberty, knowing that a society which incorporated this principle would not allow them to pursue their conception of the good, should they be "born" intolerant and despising the principle of equal liberty, as Rawls conceives it?

Rawls takes up the question of the choice of equal liberty in Sections 33ff. There he tells us "that equal liberty of conscience is the only principle that persons in the original position can acknowledge. They cannot take chances with their liberty by permitting the dominant religious or moral doctrine to persecute or suppress others if it wishes."<sup>15</sup> But, as I have just implied, they take a chance with the principle of equal liberty as well.

Rawls goes on to argue that though they run a risk with the principle of equal liberty, it is not as great a risk as with other possible principles. The parties can know that under the principle of equal liberty, their freedom of conscience will be respected,

at least within a certain limit--i.e., until its exercise threatens equal liberty itself; in other words, until the intolerant sect gains a significant following.<sup>16</sup>

But is the risk really least with the principle of equal liberty? Why should the parties to the OP choose the principle of equal liberty, with its limiting condition on the behavior of the intolerant, over the principle that minority moral and religious views shall be given all and only those liberties that do not threaten to overturn the majority's control over political institutions and social life? This principle would protect the intolerant "born" into a tolerant society to the same extent as the principle of equal liberty would. Unlike the principle of equal liberty, however, it would permit the suppression not only of threats from the intolerant, but of threats from the pacifistic, or the vegetarian, or whatever, relative to the dominant moral or religious codes of the society to which it was applied. From the point of view of a party to the OP, who cannot know what his conceptions of the good or right will entail, or whether his views will be those of the majority or not, the chance of having his liberty to prosyletize and organize circumscribed is equal between these two principles. Therefore, there are no grounds under the conditions of the OP for the choice of the principle of equal liberty over all other principles.

There are no grounds, that is, unless the parties may assume that the intolerant are justly suppressed in certain circumstances, whereas the suppression of other deviant moral views is less justifiable, if at all. This indeed is what they are allowed to assume. The "intolerant" are not just those who suppress another point of view when it threatens them; otherwise, those who defend equal liberty against oppression would be "intolerant." Rather, the intolerant are those who take their point of view to be binding on everyone, when the question of its truth is a matter of serious debate. Speaking of religious toleration, Rawls says that what motivates the selection of equal liberty rather than a principle to the effect that all are to follow the true faith is that in the OP "no particular interpretation of religious truth can be acknowledged as binding upon citizens generally," so that "each person must insist upon an equal right to decide what his religious obligations are."<sup>17</sup> It is the same with opposing moralities--each party must insist on equal liberty of conscience, since moral disagreements are such that "no one knows how to reconcile them by reason."<sup>18</sup> Thus, the parties know that the suppression of any deviant moral view is an intolerance; that is, is indefensible on rational moral grounds. But this is nothing but the knowledge of the truth of the principle of equal liberty itself.

The right to equal concern and respect, then, can only be understood against the background of a conception of the right that fills in what the right is to amount to. The failure to appreciate this explains the logical blunder Ronald Dworkin makes when he claims that the embodiment of the right in the OP does no more than make "the minimal assumption that political arrangements that don't display equal concern and respect are those established for the benefit or enhancement of certain groups, or talents, or ideals."<sup>19</sup> If this is only the assumption that one of the necessary conditions for violating a right to equality is that someone is treated unequally, then it is indeed a minimal one. But the veil of ignorance, which rules out as morally irrelevant any consideration of differences between people, makes the logically distinct assumption that any such inequalities are violations of the right to equality. It does not follow that because all violations of the right to equality are inequalities or benefits to special groups, all inequalities and special benefits are violations of the right. Nothing substantive will follow from any such purely abstract and formal conception of the right to equality. If we want the right to mean anything at all, we must already have some conception of it that includes a notion of what the consequences of such a right are. Understanding the right to equality requires

knowing through particular moral judgments what constitutes violations of it.

Rawls has set an impossible standard when he says that the thick veil of ignorance is justified because to include any more information would be to import into the OP "the very moral considerations (or aspects thereof) that we seek to understand in the light of other and more basic notions."<sup>20</sup> What Rawls fails to appreciate is that excluding information also imports a significant moral conception into the OP--namely, the very same egalitarianism that he wants to establish in the two principles. Importing moral conceptions into the OP is unavoidable, for the basic notions themselves have meaning only in the light of particular moral judgments.



## CHAPTER VII

### VAGUE PRINCIPLES AND DEDUCTIVE REASONING

It has been shown by illustration that the justification and use of general moral principles is plagued by what seems to be their pervasive vagueness, which can only be removed with the help of the very sorts of particular judgments which the principles themselves are supposed to establish as true. For example, we can appreciate what the injunction against inflicting "basic" harm requires only if we already know that we should toss a rope to a drowning person from the safety of shore.

A more general case can be made for this claim by exposing the inadequacies of the alternative picture of moral principles, which views them as sets of sufficient conditions which license an inference from some finite set of non-moral properties to a moral conclusion. This is the sort of moral principle that Hare and others have in mind when they claim that any particular judgment that X is wrong presupposes a principle of the form "All actions with properties a...n are wrong," a...n being the features by virtue of which X is

classified as the kind of wrong defined by the principle, and which features it shares with all other particular wrongs that fall under the principle.

That this is a false picture of the complexity of moral judgments can be seen by examining the concept of "murder."<sup>1</sup> Murder is by definition wrong, being an unjustified homicide. If the present view of the matter were correct, there should be some similarity between all murders such that they are all homicides that are unjustified. But in fact, there is no similarity that can be pointed to beyond the "fact" of being unjustified. To kill someone because he is a Jew is a murder; so is killing someone for his watch. These are both unjustified homicides, but there is no further similarity between them that makes them both murders. As further examples would illustrate, there is an indefinite number of morally insufficient reasons for killing someone, and so there is no common thread running through and identifying unjustified homicides.

Nor is this simply because "insufficient reasons" are all those that are not good enough, so that the indefinite number of bad reasons can be demarcated by reference to the few good reasons for killing someone. The number of good reasons for killing someone is also indefinite; i.e., "justified homicide" is no better off than

murder in lacking pervasive non-moral similarities between cases.

Let us say self-defense is the only justification for a homicide. But to say that all justified homicides are alike in being cases of self-defense is not to mark any factual similarity between them, for one of the concepts crucial to the use of self-defense as a justification is that of "legitimate force"--homicide done out of self-defense is justified only if killing is a legitimate response to use against a particular attacker. Whether it is or not depends first on the relative capacities of the attacker and the victim. Is it reasonably possible to escape by some other means short of lethal force? This is not a question that can be answered by a formula accounting for different heights, weights, strengths, training in self-defense, etc., not only because the list of relevant factors is indefinitely long, but because judging their relative weights is a matter of an assessment of reasonableness, not a matter of applying the criterion for reasonableness, for there is no such thing. Whether killing is a legitimate response depends as well on whether the "victim" is himself justly killed; whether, for example, he has forfeited or waived his right to life (the bank robber who in the course of a shoot-out kills a policeman cannot claim self-defense as an excuse). Applying the concept of self-defense, no less than

applying the concept of murder, is itself a moral, evaluative activity that calls for our assessment; it is not reducible to the application of a mechanical decision procedure.

Thus, in deciding whether a particular homicide is excused because done out of self-defense, it is not enough that I know what the facts of the case are. What I must do in addition is make an assessment of the relevance of those facts for a determination of the reasonableness and legitimacy of the use of lethal force, and there is an indefinite number of factors relevant to that determination. So neither "justified homicide in self-defense," nor "murder" can be given some closed set of criteria of application that can be relied upon to properly decide all cases that may arise. We cannot specify a true principle such as "All homicides which are X are murders," for there is no non-moral, non-normative property X shared by all murders.

What is true of "murder" is true of moral concepts generally, including the most general of them, "right," "wrong," "good," "bad." Using the criteria of application for a moral concept requires making moral judgments. There is no morally neutral criterion that tells us what "respect for persons" requires; we just know, for example, that to break a promise to someone merely for our own convenience shows no respect for him

as a person; or that to let an adult win a game of tennis under the false impression that you were trying your best shows him no respect, and so on, through an indefinite number of considered judgments which specifies for us what "respect for persons" involves, but which, like cases of unjustified homicide, are not united under a set of non-moral criteria for the concept.

General moral principles that make use of moral concepts or terms of assessment are, then, irreducibly vague, since they can't be supplied with a closed set of criteria of application. It is a consequence of this that moral argument cannot be exclusively deductive, for where the terms of the premises of an argument are vague, they don't yet offer a reason for the conclusion. Take as an example the argument

- A. 1. Mort has far less hair than average.
- 2. Men who have far less hair than average are bald.
- ∴ 3. Mort is bald.

This is deductive in form, but it is not yet an argument that supports any conclusion, for if "far less hair than average" has a different denotation in 1 than in 2, then the premises could be true, but the conclusion false. Therefore, before the inference can safely proceed, it must be determined that the denotation is the same for both occurrences of the expression.

How can this determination be made? Not by appeal to some closed set of criteria of application, for these are by definition impossible to specify for vague expressions like "far less hair than average" and "bald." Of course, even though we cannot apply this description according to some definitive standard, we often enough do know that someone is bald--e.g., Telly Savalas is bald--and that someone else is not. It is against a background of such cases, then, that we understand the use of the premise in 2, and know whether it applies to Mort.

But if this is true, then an argument for the conclusion of A cannot be simply deductive, as can be seen with the help of the following dilemma:

1. Either this case (the state of Mort's hair) is exactly like one of the cases by which 2 is understood, or it is not.
2. If it is exactly like one of them, then a proof that Mort is bald is superfluous.
3. If it is not exactly like any of them, then it is an open question whether Mort's case falls within the scope of the principle in A-2, a question not decidable by appeal to a principle, and hence not decidable by deductive argument.
4. A deductive proof (i.e., the appeal to a principle) that Mort is bald is either superfluous, or insufficient.

This is precisely the dilemma faced in moral reasoning when one appeals to a moral principle to settle some particular issue and that principle contains vague terms of moral assessment. For instance:

- B. 1. I made a trivial promise which I must break to avoid a great harm.  
 2. One may break trivial promises to avoid great harm.  
 ;, 3. I may break this promise.

Of course, the principle in B-2 is true; but uselessly so, so far as the purpose of appealing to it is to settle the question whether I may break this promise. The substantial issues in such a case are whether the promise was indeed "trivial," and whether the impending harm was significantly "great." Again, though, these are vague concepts which have no clear criteria of application, and which are understood only relative to the clear cases of their proper use, which are in the form of considered judgments like "One may break a tennis date to attend an accident victim with arterial bleeding."

Since there is no standard for a vague concept like "triviality," settling the question whether this promise I made was a "trivial" one cannot proceed deductively. The only avenue of reasoning left open is a comparison of the nature and circumstances of this promise with those of the clear cases of "triviality" by which the notion is understood. There being no finite set of properties which such cases share in common, the relevant similarities between this case and the clear cases will be matters of degree: it will neither be exactly like any one of them, nor exactly like the composite of them, but will show some resemblances to some cases, none to others.

The sort of argument that this suggests is a kind of argument by analogy that compares the case at issue to the considered judgments that supply the principle appealed to with its content. Precisely what kinds of analogical argument these are, and by what rules they are governed, I am not now prepared to say. At this point, I can only hope to illuminate their nature by example, several of which I will offer in the next chapter.

I can, though, suggest that such arguments cannot be formal in the way deductive arguments are, by use of an analogy which brings out in another way the defect of thinking of deduction as the sole paradigm of sound moral reasoning.

Considered judgments, such as those that lie behind the principle in B-2, are to an understanding of moral concepts, what a recognition of the truth of "If all men are mortal, and Socrates is a man, then Socrates is mortal" is to an understanding of the logical operators "if...then," and "all." Anyone who disagrees that a promise to play tennis is trivial relative to the duty to render aid shows thereby that he has a defective understanding of the concept of triviality, (in the sense of not knowing how to apply the term; he may--trivially--"understand trivial" in the sense of being able to supply a synonym like "insignificant" without yet being able



to apply either word properly) just as someone who claims that the conditional above is false shows that he misunderstands the proper use of "if...then," and/or "all."<sup>2</sup> One who believes that tennis dates are solemn promises will not be moved by appeals to the principle in B-2; and whoever thinks it does not follow that Socrates is mortal will not be persuaded by referring him to a logic text.

The difference between moral judgments and logical ones is that the truth of the latter depends solely on their form. Once the truth of a logical claim like that just given is acknowledged, we may ignore the nature of its content (that it is about Socrates, and man, and mortality) and abstract from it its logical form, which then becomes the basis for the formation of rules of inference by which other particular arguments are passed on as valid or invalid, like or unlike paradigm cases of good deductive reasoning. Whether a particular logical judgment is an instance of a logical rule of inference is a question that is independent of the particular subject matter of that judgment.

Considered moral judgments and principles, however, are not related to one another in this way; their truth is not "formal," but inescapably material. The truth of the judgment that tennis dates are not solemn promises is purely a matter of the content of that judgment and the particular nature and circumstances of

tennis dates. One can abstract from a logical judgment to form a principle without loss of clarity or meaning, because logical judgments are abstract by nature, not being about anything. But when one abstracts from considered moral judgments to form a principle, there is inevitably a loss in clarity and meaning, since the truth of such judgments is not independent of their material content.

The mistake, then, of the view that moral argument properly proceeds solely by the use of principles lies in the failure to recognize that moral arguments are not formal--hence the inadequacy of moral principles like "one may break a trivial promise to avoid great harm." They are necessarily material, and this means that moral argument is bound to the particularity of considered judgments, and must in some way use them directly in moral reasoning.

I want to stress that I do not want to claim that principles play no part in moral reasoning, and I have been careful to avoid that implication. Though they cannot be the only elements in moral reasoning, they serve the important function of codifying, even if vaguely, which sorts of considerations are relevant and must be taken into account in making moral evaluations, and the relative weights of some of these.

Moral principles can perhaps themselves be reasoned about in some sense independently of considered judgments,

though the application of the results of such reasoning cannot be made independently of them. We may take as an illustration of how this is possible a recent article by Norman Daniels,<sup>3</sup> where he offers an elaboration of Rawls' notion of "wide reflective equilibrium" as an account of the logical forces that are brought to bear in moral arguments, which he takes to be analogous to the process by which scientific theories are tested. I will not address the question whether wide reflective equilibrium is significantly like the scientific method. What I will focus on instead is the implied claim that wide reflective equilibrium provides a way of testing and criticizing moral principles independently of their consistency with considered judgments (this consistency is "narrow reflective equilibrium").

Wide reflective equilibrium is a method that seeks coherence between (1) considered judgments, (2) a set of moral principles; and (3) a set of background theories. The coherence is achieved by moving down from the set of background theories through philosophical argument intended to decide between competing sets of principles each of which has a rough, though not identical, fit with our considered judgments. Significantly, the background theories may themselves be constrained by some set of considered judgments. So long as this set is not identical to the set that constrains the moral principles

being evaluated, the method is not viciously circular, and provides an argument for the truth of moral principles independent of their respective considered judgments.

Daniels does not provide any detailed example to clarify how this process is supposed to work. Let me offer the following, suggested by Martin Benjamin in discussion. Suppose that before us we have the following, which is by hypothesis a considered judgment:

1. (Considered judgment) It is wrong to kill a human being in an irreversible coma.

And suppose that the principle offered to explain this as well as our other considered judgments concerned with murder is

2. (Principle) It is wrong to kill an innocent, helpless human being.

If this is all we have, and this principle fits our considered judgments as well as any other, then Daniels would say that there is no way to evaluate the truth of 1. But we may widen the scope of the equilibrium we seek, to include, in the present case, theories of personhood and of the nature of rights, and by philosophical arguments establish the following:

3. (Background theories) A "person," as a minimum necessary condition, has interests which he wants protected. And because a "right" is erected as a protection to an interest, it is persons who possess rights (such as the right to life).

When these background theories are brought to bear on 2, they require that it be amended to

2'. It is wrong to kill an innocent, helpless person.

since a consequence of the background theories is that it is only persons who can be wronged by being killed.

Now Daniels' position seems to be that once the principle has been critiqued and amended, we can apply it straightaway to the case of the comatose to show that the considered judgment in 1 is false, at least once neurophysiology can reliably identify the irreversibly comatose and the state of their brains.

But this is false, for the question whether the comatose have "interests" of a kind and degree sufficient to call them "persons" with a right to life is not a psychological issue, but a moral one. We must determine whether the comatose can have interests which deserve protection, and which can be protected only by not killing them. This can be done only through a comparison of the interests they can be said to have with the interests which paradigmatic persons have which are protected by the right to life.

Clearly, the comatose do not have interests which require a felt desire or a sense experience--they cannot have an interest in thrilling to the sight of the sun rising tomorrow morning. However, these are not the only

sort of interests which the right to life protects for even the ordinary person; one may, for example, have an interest in working long enough to send a child to college, an interest that requires that one continues to live, but which does not seem to have anything much to do with the content of one's felt experience. One need not even still be living when the child starts school.

The comatose may also have interests of this sort--in not having their bodies mutilated or invaded, or in a child's receipt of an inheritance that requires the continuance of life for a certain period. These and other examples must be evaluated and compared to the interests of persons which are acknowledged in considered judgments to be protected by the right to life to determine whether they are legitimate interests which command and require the right to life.

Thus, no matter how useful philosophical arguments may be in the justification and amendment of principles, those principles may still remain crucially vague, and need interpretation through the use of considered judgments to decide on the application of the principle to questionable cases.

## CHAPTER VIII

### ANALOGICAL REASONING: SOME ILLUSTRATIONS

In broad terms, my argument so far has been that principles alone can play no decisive role in moral reasoning about particular cases, because any moral rule sufficiently general to be plausible a priori or to be thought "basic" to our moral understanding employs vague moral concepts that can be themselves interpreted only by means of our prior decisions in considered judgments.

Thus, we are thrown back upon reliance on a form of argument by analogy that argues more or less directly from our considered judgments that interpret a principle to a decision whether and how that principle applies in the situation confronting us. It is still unclear precisely what such an argument would look like though, so in this section I will show by example (1) what this form of argument involves in practice; (2) that it is actually a not uncommon form of moral argument; and (3) that it is persuasive. These illustrations will take the place of any attempt at a general account of analogical

reasoning, being merely suggestive of a direction for further study.

### Familial Obligations

In "What Do Grown Children Owe Their Parents?" Jane English attempts to show that grown children do not owe their parents a debt of gratitude for the sacrifices their parents made on their behalf; when parents are in need, their children have no special obligation to "repay them" by extending help.<sup>1</sup>

Her method consists first in contrasting our considered judgments concerned with the duties incurred by requesting and receiving favours, and considered judgments concerned with the duties that follow sacrifices made out of love or friendship (if there are any such "duties").

For example, if I ask my neighbor to collect my mail while I am away, I create a debt to her that she has the right to have repaid--i.e., I would be obligated to collect her mail at a later date, should she ask. Not every sacrifice is a "favor" in the relevant sense, however. If my neighbor mows my lawn while I am away, I have no duty to reciprocate. Her mowing my lawn might create in me a feeling of friendship or affection, English points out, but then anything I subsequently do



for her is properly said to be done out of friendship, not in repayment.

The principle governing favors, English suggests, is reciprocity--one should return in kind equal to what was given. The principle governing friendship is mutuality--"friends offer what they can give, and accept what they need, without regard for the total amounts of benefits exchanged."<sup>2</sup> So, for example, if a man asks his date for a goodnight kiss in repayment for the meal he paid for, this is evidence that the meal was not a gesture of friendship, but part of what he took to be a contractual agreement. Moreover, after a friendship ends, the duties of friendship end. If a man donates blood for his wife's operation, and they are later divorced, he has no right to expect that she donate blood for him in return so long as the original donation was supposedly out of love and concern.

It is important to realize that these principles, even if true, do not settle the issue by themselves. We must also determine whether or not the sacrifices of parents are "favors," or whether instead they are properly assimilated to the sorts of things friends do for one another out of love and concern. That is, we can settle the issue whether grown children owe their parents help only by making analogical arguments from our

considered judgments on the duties of favors and the duties of friendship.

English's argument recognizes this, for she goes on to urge that the relation between parents and children is not one characterized by the granting of favors, but one that more closely resembles the relation between friends. The sacrifices of parents are not favors that carry duties to reciprocate simply because, unlike the favor of collecting the mail, the child did not ask that these sacrifices be done for him. In addition, since parents naturally do things for their children out of love and concern, this is a positive reason for thinking that it is the mutuality of friendship that is involved in the duties between parents and children. Therefore, once there is no longer love between parents and children, there are no longer any special duties of mutual aid beyond those owed to everyone else.

By pointing out these features of the parent-child relation, English has tried to show how that relation is more like that between friends when they have duties of mutual aid than like contractual exchanges governed by reciprocity, and has thus provided an argument for her claim that the moral principle of reciprocity does not apply to the parent-child relation; an argument, moreover, that is an argument by analogy which notes the

similarities and dissimilarities between this relation and those other relations described in our considered judgments concerning the principles of reciprocity and mutuality.

Yet the conclusion that English reaches is unsatisfactory, as a pursuit of some other comparisons with considered judgments will show. We may start by pointing out that there is a significant difference between the parent-child relation and both friendship and favor-granting--the sacrifices a parent might make (at least some of them) are obligatory. The child, or someone acting in his behalf, may justly make demands (for food, clothing, minimum medical care, minimal education and nurturance, etc.), not just ask for help that the parent may or may not give, at his discretion. This is a further reason that these sacrifices of parents are not favors of the kind described. But in addition, the duties of parenthood survive the loss of mutual love and affection between parent and child, unlike the duties of friendship. The parent who comes to dislike or feel indifferent toward his child is not thereby freed of all responsibility toward it. Children "depend on" their parents in a different sense in which friends "depend on" each other.

So at least while the children are dependents, these obligatory sacrifices made by parents are properly seen neither as favors nor as simply the obligations of friendship. So far though, this analysis just provides a different line of reasoning for the same conclusion, for so long as the parental sacrifices are those required by duty, they create no obligation to reciprocate, just as in the following example: A policeman, in the ordinary course of duty, recovers a sum of money, but is wounded in the process. The person whose money is retrieved is under no obligation to compensate the policeman for his pain and suffering, though he may deserve acknowledgment in the form of thanks.

It has been clearly established, then, that the obligatory sacrifices of parents do not create for the child any obligation of repayment; at least not under the principles of reciprocity or mutuality. But what of those sacrifices that a parent makes out of generosity, that go "above and beyond the call of duty"? English thinks it is enough that these sacrifices were not requested to show that they create no obligations to repay. But there are different kinds of "favors," and they create different sorts of obligation. That is, the principle of reciprocity must be more broadly interpreted than English allows, as an appeal to additional considered judgments will show.

Imagine a postman making his normal rounds in a wealthy neighborhood who comes upon a burning house. The owner is sobbing hysterically because her beloved dachshund is trapped inside, sure to suffer a terrible death. The postman dashes in and saves the dog, but is badly burned. He recovers, but his insurance will not cover all his bills, which the wealthy dog-owner can easily afford to pay. Now there is an important sense in which he did her a "favor," not because she asked for his sacrifice, but because he did much more than duty required out of concern for her. Because the sacrifice was made out of concern (the sole motive for many other sorts of favors, by the way), not in expectation of repayment, it would perhaps be inappropriate for him to demand compensation. Still, there is more than the duties of friendship involved here, because he nevertheless deserves compensation from the dog-owner. If she declines to give it, she has done something wrong; she has not discharged her debt of gratitude. Her obligation to help him is not the obligation imposed by a claim of a right to compensation; we may not coerce her into payment. But neither is her obligation to him merely the general obligation to help someone in need. It is greater than that, and so she does owe him help as required by the principle of reciprocity.

This example, along with the others, illustrates how considered judgments are indispensable to the interpretation of a principle like that requiring that favors be reciprocated. The principle is vague, because it does not tell us what counts as a "favor" of the sort that carries duties to reciprocate. We know that my neighbor collecting my mail at my request counts as such a favor. We know that her unexpectedly mowing my lawn does not count as one, even though it may properly be called a "favor." We know that the postman's heroism counts as one. There may yet be other examples of favors distinct from these. We can already appreciate that there is no set of non-vague necessary and sufficient conditions for being a "favor" of the required sort. Both mowing the lawn and dashing in to save the dachshund are supererogatory, but one is a "favor" in the original sense, and the other not, so being supererogatory is not a sufficient condition. The difference between them is one of the degree of sacrifice and the degree of significance of the benefits given, which are not differences of kind. Deciding where other cases belong is not a matter of the application of a principle or rule so much as making an assessment of where along these continuums each case falls, relative to the fixed points provided by our considered judgments. Being supererogatory may not be a necessary

condition either, or at least not a useful one, for what is required by the duties of the relation that one happens to be in is often itself not definitively determinable. The duties of parents are to see to the welfare of their children, but what counts as being in the child's welfare, and what extent of welfare the parent is obliged to promote when this obligation conflicts with other obligations either to himself or others, may again not be settleable by principle, but known only by example.

A decision about the duties created by acts of parental generosity requires that we make a direct comparison between particular sets of parental generosity and those considered judgments (such as the postman example above) that serve to specify for us the sorts of "favors" that create reciprocal duties. The case of a parent who works 70 hours a week or goes deeply into debt to send a child to college might be a comparable sacrifice to that of the postman. Such parents are sacrificing above and beyond the call of duty, and what they do is out of love or concern for another's welfare. So it may remain true of them, too, that they have no right to expect a "return." It is only to this degree that their sacrifices for their children are like sacrifices friends make for one another. But if their

sacrifice is sufficiently extraordinary, and the benefit to the child sufficiently great (cf. the postman example), the child who abandons them, perhaps now deeply in debt, has done something wrong, for he is being ungrateful. He owes them at least his thanks and his acknowledgment of the generosity of their sacrifice, and if, as in the above example, their current need is a result of sacrifices made for him, he owes them his help.

In summary, the principle that favors, but not acts of friendship, create duties to reciprocate, can be understood and applied only in light of the considered judgments that specify how the key notion of a "favor" is to be taken. The appeal to these judgments is not merely suggestive, so that we may devise a more sophisticated principle that sidesteps the necessity of relying on considered judgments. The principle that supererogatory acts involving great sacrifice create duties is certainly a specification of a particular kind of favor, but it has not avoided the need for considered judgments--they are still required to fix what is meant in the principle by "great sacrifice," and perhaps also, in particular circumstances, what sorts of acts are "supererogatory."



### Limits of the Right to Life

My other example of analogical reasoning will come from Judith Thomson's famous article, "A Defense of Abortion."<sup>3</sup> In that article, Thomson assumes for the sake of discussion that the fetus is in the full-blown moral sense a "person," with a right to life. She argues that even so, abortion may be permissible, since even those with a right to life do not thereby have a right to the use of another's body against her will, and so abortion is not murder.

One of her targets in the article is the view that abortion cannot be performed even to save the life of the mother because this would be the direct killing of an innocent person (the child), which is murder.

But, Thomson replies, an innocent may be an attacker who is justly killed, at least by the one who is threatened, for imagine a woman who is trapped in a house with a rapidly expanding child who will crush her to death unless she kills him first. Perhaps bystanders can do nothing, "but it cannot be concluded that (she) too can do nothing, that (she) cannot attack it. . . ." <sup>4</sup>

And this case shares the following features with the situation of a woman threatened by the growth of her fetus:

. . . there are only two people involved, one whose life is threatened, and one who threatens it. Both are innocent: the one who is threatened is not threatened because of any fault, the one who threatens does not threaten because of any fault.

"In sum," Thomson claims, "a woman surely can defend her life against the threat to it posed by the unborn child, even if doing so involves its death,"<sup>5</sup> apparently on the strength of the principle that an innocent person may be an attacker who is justly killed by one who is also "innocent"--i.e., "is not threatened because of any fault."

Assuming that this principle is a proper interpretation of the lesson of the expanding child judgment, we must still mount some reasoning that shows when (if at all) a woman who is pregnant is not so because of any "fault." The principle itself does not specify what counts as being "faultless" so that one is justified in killing a "faultless" attacker. All we know so far is that the woman in the house with the child is "faultless," and that so is the child, even though it is the potential cause of an evil (the death of the woman). This is a considered judgment that serves to interpret the sense of the vague moral concepts that occur in the principle. In order to successfully apply this principle to the case of a woman threatened by her fetus, we must show that the principle properly applies to that

situation as well. The principle itself cannot decide this question. We can only see how the situation of the pregnant woman compares with the case of the expanding child or other considered judgments which exemplify the principle. If there are significant differences between the circumstances surrounding abortion and these considered judgments, then the principle is no defense of abortion, however true it may be when applied in other circumstances.

One possible difference, for example, is that the person in the house is in no way responsible for the presence of the expanding child, having had nothing to do with its getting there; but a mother who voluntarily engages in intercourse does have something to do with the fetus' presence, and so is responsible for it, so that its right to life does guarantee it the use of its mother's body. The principle justifying killing the innocent does not apply. This is indeed a difference, but it is a difference whose significance Thomson must deny. And she denies it not by resorting to a theoretical argument on the difference between causal and moral "responsibility," but by dealing directly with the sort of case that such a theory would depend upon, and which is also an exemplification of the principle in question. There are, she asks us to imagine,<sup>6</sup> people seeds that drift about in the air, which will take root and grow

in carpets and upholstery. If, not wanting children, one puts the best available screens on the windows to keep the seeds out, one is not in the appropriate sense "responsible" or at "fault" if a seed gets through a defective screen and takes root, and so would be justified in killing it if it threatened one's life.

This of course is meant to be analogous to the situation of the woman who, despite the use of the best available contraceptives, finds herself pregnant. She too, then, is not morally responsible to the child, and it has no right to the use of her body.

A response to this defense which Thomson acknowledges would be to say that one is after all responsible for the person seed, for one could have lived with the windows closed and without carpets and upholstery; its presence could have been prevented. Thomson's reply is to say that "by the same token anyone can avoid a pregnancy due to rape by having a hysterectomy," but our considered judgment is that a woman who is raped is not responsible for the fetus that results, even if she could have taken such a drastically certain precaution. That is, this is an example of a circumstance in which one is "responsible" for the fetus in the sense of "could have prevented" its conception, but which is still one in which the woman is "faultless" with respect to the

fetus' presence, and also permitted to kill it should it threaten her.

The usefulness of this latest considered judgment (setting aside the complication of the case's being one of rape) for applying the principle permitting the killing of innocent attackers to abortion depends on the assumption that the sacrifice of having a hysterectomy is no worse than the sacrifice of living with the windows closed for much of one's life, or in turn, abstaining from intercourse when one is unprepared to assume the risk of pregnancy. We must ask what degree of preventive measures are required in order to be "faultless" in the sense that the principle requires. Is abstention too much to ask?

A whole new range of argumentation opens up now, concerned with the issue of whether abstaining from intercourse is a reasonable precaution to require of those who want no responsibility toward any fetuses, so that those who "indulge" in intercourse are not "faultless" and so have no right against the claim of any resulting fetuses to the use of their bodies. The method of analogical reasoning would address this problem by asking whether abstention is more like cases of reasonable precautions, or cases of excessive precautions, that ask too much. (For example, we do not think that a water skier who runs over a passenger who falls

out of a passing boat is in any morally significant sense "responsible" for the death. It would be inappropriate to blame him, saying that he should have foregone the pleasure of water skiing, knowing that such unusual accidents have happened in the past.)

It is in this way that arguments by analogy with considered judgments require that the whole range of our considered judgments be coherent with one another, the question of whether an abortion is murder, in the present instance, requiring a look at the sort of precautions that excuse one from moral responsibility for the lives of others, and whether and how abstention from intercourse fits into our considered judgments in this area.

I would like to offer one final example of the use of analogical arguments requiring the appeal to considered judgments, which will serve at the same time to reinforce my claim that such appeals are an unavoidable part of moral argument.

One possible objection to the use of considered judgments is that they are relativistic. An argument like Thomson's will be persuasive only to the extent that we share the considered judgments that it hinges on. But if we don't, if we don't agree, for example, that the woman in the house with the expanding child has a right to kill it, then the argument comes to a grinding halt.

It can't tell us which of these reactions should be taken as the morally correct one.

This is essentially the criticism that R.M. Hare makes of Thomson's argument in his article "Abortion and the Golden Rule":<sup>7</sup>

Her examples are entertaining, and help to show up our prejudices, but they will do no more than that until we have a way of telling which prejudices ought to be abandoned.

That method being, Hare goes on, "a theory of moral reasoning that will determine which arguments we ought to accept."<sup>8</sup>

To the charge of relativism, I plead guilty, but it is no special objection to the use of considered judgments in analogical reasoning. If two people disagree over one of the crucial judgments in an argument like Thomson's, and further they can't agree on any other judgments that might serve as a starting point for discussion, then there is no basis for their coming to a rational decision on the difference between them. First, though, I don't think that this is likely in fact to happen, that two people differ in all of their considered judgments. But even the mere possibility of it happening is an objection to analogical argument only if there is some other method of argument in ethics that can avoid the risky appeal to considered judgments. If the proposed alternative also requires the appeal to considered

judgments, then it is no more secure against the charge of relativism than the method of analogical argument which I have been outlining.

Hare thinks that the ultimate court of appeal in settling moral disputes is the realm of moral theory; it is only at this level that we can decide between the conflicting "prejudices" people might have as expressed in their considered judgments. The theoretical principle that Hare takes to be central is the "principle of universalizability," which requires us to make the same moral judgments about relevantly similar cases. For the purposes of his discussion of abortion, Hare embodies this principle of reason<sup>9</sup> in a form of the Golden Rule: "we should do unto others what we are glad was done to us."<sup>10</sup> One can question whether the principle of universalizability implies this version of the Golden Rule. The implication seems to require the additional assumption that we are obligated to do whatever it would be beneficial to do, an assumption one might not share if not a utilitarian like Hare. Be that as it may, Hare then employs this principle in the following argument:

if we are glad that nobody terminated the pregnancy that resulted in our birth, then we are enjoined not, ceteris paribus, to terminate any pregnancy which will result in the birth of a person having a life like ours.<sup>11</sup>



The critical question for my purposes is whether Hare can make this argument work without any appeal to considered judgments; if he can't do without considered judgments, then the "theoretical" moral argument he advocates is vulnerable to the same charge of relativism that he brings against Thomson.

There are certainly other uses of the Golden Rule in his article that do require the appeal to considered judgments. Specifically, what does the Golden Rule tell us to do when it's impossible for us to do to everybody implicated in our decision what we would be glad was done for us; for example, when the choice is between aborting a fetus and causing its mother great psychological trauma? According to Hare, "we are to do to the others affected, taken together, what we wish were done to us if we had to be all of them by turns in random order."<sup>12</sup> As I pointed out in Chapter IV, this takes the aggregate to be an "other" like ourselves, which it is not, not being a subject of experience. So it's not clear how the Golden Rule would require or even permit that we conceive of our duty in the way Hare describes. But even if we do treat the aggregate as if it were a subject of experiences, there is no principle that could guide us in making the choices that we would make if we were it. What principle will tell me whether I would

prefer the miserable life of the mother together with the happy life of the fetus (if the pregnancy is not terminated) to the happy life of the mother together with no life for the fetus (if it is terminated)? There is no principle; I just have the preference that I do. And how do we decide between people with conflicting preferences? Certainly not by appeal to the Golden Rule itself, when the issue between them is what the Golden Rule requires when faced with such a choice. All we can do is appeal to other sorts of preferential choices that we agree upon, and argue that similarities with these favor one preference or the other in the present matter of abortion. It is only against a background of such judgments that it would be possible to understand the application of the Golden Rule to the hypothetical choice situation (choosing as if we were the aggregate) that Hare asks us to envision. Indeed, he himself says quite confidently and without argument that "in the kinds of cases in which (liberals) would approve of termination, the interests of the mother will usually be predominant enough to tip the balance between those of the others affected, including potential persons."<sup>13</sup> If it weren't possible or acceptable for us to make some such judgments with confidence and without argument, the Golden Rule would be useless in settling conflicts of interests.

Still, even if this use of the Golden Rule requires the appeal to considered judgments, it might perhaps be thought that Hare has at least shown that there is a prima facie obligation not to abort the fetus, the potential happy person, and that he has done this with an argument that requires no appeal to considered judgments:

1. We should do unto others what we are glad was done to us.
2. We are glad nobody terminated the pregnancy that resulted in our birth.
3. This fetus will become a person glad that nobody prevented his birth only if we don't abort it.
4. We should not abort this fetus.

This use of the Golden Rule requires, as Hare acknowledges, that among the "others" with whom the rule is concerned are "nonactual persons," since if the fetus is aborted, no person comes into existence. Yet, we have somehow wronged him (the non-existent person) nonetheless by preventing him from enjoying the kind of life that we enjoy. How can the Golden Rule obligate us to non-existent persons in this way? Why should we think that the "others" we should be concerned with include merely possible persons? Clearly this is not a question

we can expect the Golden Rule to answer, since it concerns how the Golden Rule is to be understood. Yet these are moral questions, and so on Hare's account, should be addressable by the basic principle, the Golden Rule.

The method of analogical reasoning would approach this question by comparing cases in which the Golden Rule clearly does apply, to the situation presented by abortion, where the application of the Golden Rule to merely possible persons is problematic. To the extent that it can be shown that this questionable application is similar to acceptable uses of the Golden Rule, then we have all the arguments we can ask for for allowing Hare's application of the Golden Rule to abortion.

This, in fact, is the sort of argument that Hare employs on this crucial question. He attempts, for example, a comparison with our considered judgment that we can wrong presently non-existent persons--i.e., future generations--by, for example, "using up all the world's resources or by releasing too much radioactive material."<sup>14</sup> The disanalogy with the case of abortion, of course, is that harming future generations in the ways which Hare describes results in the future suffering of some actual person, in the form, say, of genetic disease. But who is there who is harmed by abortion? There doesn't seem to be any subject to whom the Golden Rule would apply.

The question whether the non-existent person is harmed by abortion is really the other side of the question whether our existence is a benefit given to us, for Hare is right that "if it would have been a good for him to exist...surely it was a harm to him not to exist."<sup>15</sup> If I would benefit a being by bringing him into existence, then it follows that I've harmed him (by denying him) if I refuse to bring him into existence. But why should we think either of these is true? Because, Hare tells us,

We do commend actions which resulted in our existence--every Sunday in thousands of churches we give thanks for our creation as well as for our preservation and all the blessing of this life; and Aristotle says we ought to show the same gratitude to our early fathers as 'causes of our being'.<sup>16</sup>

Hare thinks that all he's doing by appeals to these considered judgments is establishing that it's meaningful to talk about harming (or benefiting) the non-existent. But Hare's argument doesn't merely require that such judgments be sensical, but that they be true. Without such judgments, the argument can't even get started. It's useless to his purposes pointing out that we regard our lives as a benefit to us unless he can also say that life is in fact a benefit to the non-existent. Only if we agree with these considered judgments will we agree with Hare that the benefit of our lives was a good "done for us" under the meaning of the Golden Rule.

Hare has not succeeded, then, in giving a "theoretical" argument against abortion, free of the considered judgments he decries as mere prejudices. His own argument, in particular the meaning of the Golden Rule within it, depends essentially on his reader's agreement with certain considered judgments. He is, therefore, no more free of relativism than Thomson.

I might add, as a concluding remark, that the considered judgments to which he appeals are likely a good deal less defensible than hers. Unless we entertain some notion of an eternal or at least an immaterial soul that could or could not be embodied or given life, the judgment that life is a benefit that can be given or denied to us bears no resemblance to our other (and more central) judgments concerning the situations in which we speak of benefits being conferred or denied, and is therefore isolated and ad hoc, no matter how strongly some people might hold it. The price of coherence with other considered judgments is the burden of a very problematic metaphysical doctrine, which can itself lay no very great claim upon the belief of others.

## CHAPTER IX

### SUMMARY AND CONCLUSIONS

We saw in the first chapters that the appeal to moral principles to settle moral issues was limited in its usefulness by the vagueness of the moral terms which the most general of principles use. In examining the arguments of Donagan, Gewirth, and Rawls, it was shown in particular that what is often taken as the foundation of the non-utilitarian elements of morality, the injunction to respect other persons, is vague with respect to just what constitutes "respect," and what aspect of being a person calls for that respect. Neither, it was seen, is utilitarianism free of the problem of vagueness. This vagueness, it was argued, was and must be in all cases removed by appeals to considered judgments which exemplify for us the meaning of such a principle.

Of course, principles need not employ vague terms of moral assessment. "A person is justly punished who is found guilty by a jury of his peers" contains no such terms. "It is wrong to kick old ladies in the shins for entertainment" is another example. But such cases are called "principles" only by a courtesy, for while they

are universal in the sense of containing no proper names or definite descriptions, they are so particular in their content as to represent moral conclusions or isolated judgments, not reasons which might serve to unite otherwise disparate individual judgments with one another under one heading, as when the requirement to respect persons is used to account for prohibitions against telling lies as well as prohibitions against paternalistic interventions in the affairs of others.

The price of the desired generality is a loss of specificity in the criteria of application of systematic moral principles. In Chapter VI, I argued that moral concepts were "open-textured," having no closed sets of criteria of application. As moral principles increase in their generality, the more various are the sorts of cases and judgments subsumed under them, and the more open-textured are the moral concepts which are used to unite those different judgments.

Another way of putting this same point is to say that the making of moral or normative judgments is an unavoidable part of using moral principles. General principles which unite more particular judgments under such terms as "respect," "trivial," "harm," and so on, can only be fruitfully applied if we can answer the question whether these concepts apply in the appropriate way



to the problem before us. Why does lying to a person with an incurable disease about his illness violate the "respect" due him? This question is a moral one that calls for our assessment, not for the processing of some set of non-moral necessary and sufficient conditions. It is not just that we haven't yet been imaginative enough to specify when "respect" has been violated, or that Moore's Open Question argument has made us skeptical of naturalism in ethics. It is rather that moral principles and the terms they employ are necessarily abstractions of a moral landscape which is textured in an indefinite and open-ended way. The contrast that was drawn earlier with logical principles is instructive. Logical principles are abstractions which are relatively easily applied because the logical features of particular sentences which they require us to pick out are few in number and, usually, readily identifiable. In morality, it is not the same, for here we must decide whether an action is "right" or "wrong," violates the "respect" due others, is a "harm," and so on, all things considered. All of the morally relevant features of an action and its consequences must be considered, and weighed relative to one another. There is no more a decision-procedure for this weighing process (which is the judgment of the relative strengths of competing evidence and

considerations) in ethics than there is one in science or in the assessment of factual arguments generally.

In any moral issue of any significance, there will be reasons on both sides of the question. In some cases, the competing reasons will take the form of competing principles, as in the ubiquitous conflict between principles of autonomy and respect for the individual, and principles of beneficence. In other cases, which are the sort with which my thesis is concerned, the competition will be between conflicting interpretations of the same principle, as in debates over the question of whether preferential hiring of women or minorities violates the respect due the better-qualified who is passed over. In each of these sorts of cases, we are called upon to make an assessment, or decision, which cannot itself, in the nature of the case, be the application of a principle, though it is a decision that affirms one principle, or interpretation of principle, over another.

This "leap of reason," I argued, since it is unguided by principle, is outside the bounds of deductive inference. If we accept the limitations of Hume's fork, which declares that reason is either deductive or inductive, the conclusion drawn is that morality in the final analysis is arational, and that between ideologues who have leapt in different directions at the crucial point, there is no reasoning.

I have urged, to the contrary, that the limits of deduction and principle in morality are not the end of reason, at least when the dispute is over the proper scope of a moral principle. Here one can appeal to the considered judgments which exemplify for us the demands of the principle in question. Such judgments serve as fixed points to which one may refer in deciding whether a particular situation falls within the scope of the principle involved.

They are fixed points, of course, only insofar as the parties to a dispute agree that they are. If two people disagree whether the act of falling upon a live hand grenade to save the lives of one's comrades is consistent with "respect for persons," a judgment about that specific act could hardly serve as a fundamental element in the resolution of any other disagreement about the principle.

Just so long as there are such fixed points, we may argue from them in defense of an interpretation of a moral principle. The argument cannot be a deductive one that proceeds through the application of principle, however, for the considered judgments do not supply a definition of "respect for persons," for example, in a set of necessary and sufficient conditions. Rather, they are elemental moral commitments which in a variety of

different ways and senses exemplify what "respect" requires; they are the moral decisions which are an inextricable part of understanding the meaning of the moral principle itself.

The argument over whether a particular case properly falls under a moral principle consists of attempting to place it within a family cluster formed by the considered judgments which exemplify the principle's use. So, for instance, the debate over capital punishment is in its overall structure an argument about the analogies traceable between judicial killing and killings of other sorts, both unjustified, and justified (e.g., done in self-defense). Proponents point out that by murdering another, the criminal has forfeited his own life, so that the state may justly kill him just as his victim or some bystander would have been justified in doing, except the lives which are being defended are those of potential victims (through the deterrent effect), or the "life" of the social order which must support the prohibition against killing another as the worst evil. Opponents to capital punishment generably argue that it is disanalogous to self-defense, since it does not defend the victim (it will not "bring him back"), and the motives that support it among the general public are those of unreasoning revenge,

which makes capital punishment like any other murder done out of hatred.

The structure of the debate, then, is an extended argument by analogy between capital punishment and certain of our considered judgments involving homicides, chiefly those concerned with self-defense. The substance of the debate will often be, however, over the factual presuppositions behind the claims for similarities (e.g., especially the question of the deterrent effect), for it is by defeating those that we are often best able to weaken the analogies of one side or the other, and so shift the balance of evidence clearly to the other side of the question.

Analogical arguments with considered judgments need not proceed exclusively in the factual domain, however, as was shown in the example from Judith Thomson. There it was pointed out that whether the fetus has a claim to the use of its mother's body depends on whether it is a guest, or an uninvited intruder; whether the mother who used contraceptives is in the appropriate sense "faultless" of its being there, and that this was a matter of comparing the precaution of contraception with other sorts of precautions, both those that are reasonable and those that are not, taken for the sake of pursuing a pleasure. Some of these comparisons may be

along factual lines (e.g., comparing degrees of risk), but others will not be--for instance, those that involve an assessment of the importance of sexual relations in our lives.

All I've been able to provide are sketches of the ways in which analogical arguments proceed. A body of rules governing them will probably be as hard to come by as they have been for inductive arguments, since in neither case is there any necessary or formal relation between premises and conclusions that can be abstracted into a principle of reason. Even after the appropriate analogies have been thoroughly argued, the proper conclusion will not fall into place of its own accord; there will still be competing evidence to be weighed. Analogies with considered judgments will not eliminate the necessity for making the leap of reason in moral arguments. But they will provide a better view of the choices below.

NOTES TO THE TEXT

## NOTES TO THE TEXT

### Chapter I

<sup>1</sup>David Hume, A Treatise of Human Nature, Book I, Part III, Section XI (of the probability of chances).

<sup>2</sup>J.P. Schneewind, "Moral Knowledge and Moral Principles," in Knowledge and Necessity, Royal Institute of Philosophy Lectures, Vol. 3 (1968-69):255-262.

<sup>3</sup>This seems to be Bernard Rosen's view in his criticism of Hare. See Chapter II.

<sup>4</sup>William David Ross, The Right and the Good (Oxford University Press, 1930), pp. 19-20.

### Chapter II

<sup>1</sup>See R.M. Hare, Freedom and Reason (London: Oxford University Press, 1963), Sec. 2.2; The Language of Morals (London: Oxford University Press, 1952), p. 60ff.; and "Principles," The Presidential Address, Aristotelian Society Proceedings, 1972-73.

<sup>2</sup>Bernard Rosen, "Rules and Justified Moral Judgments," Philosophy and Phenomenological Research, March 1970, pp. 436-443.

<sup>3</sup>See, for example, William K. Frankena, Ethics, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1973), pp. 24-25.

### Chapter III

<sup>1</sup>Alan Donagan, The Theory of Morality (Chicago: The University of Chicago Press, 1977).

<sup>2</sup>Ibid., p. 66.

<sup>3</sup>Ibid., p. 144.

<sup>4</sup>Ibid., p. 145.



<sup>5</sup>Ibid., p. 149.

<sup>6</sup>See Bernard Gert's review of Donagan's book in The Journal of Medicine and Philosophy, Vol. 2, No. 4 (December 1977): especially 416-417.

<sup>7</sup>Donagan, pp. 154-155.

<sup>8</sup>Ibid., p. 82 (*italics mine*).

<sup>9</sup>Ibid., p. 89.

<sup>10</sup>Ibid.

<sup>11</sup>Ibid., see Chapter VII.

<sup>12</sup>Ibid., p. 88.

<sup>13</sup>Ibid., pp. 77-79.

<sup>14</sup>Ibid., p. 68.

<sup>15</sup>Ibid., p. 71.

<sup>16</sup>Ibid., p. 72.

<sup>17</sup>Ibid., p. 71.

#### Chapter IV

<sup>1</sup>J.J.C. Smart and Bernard Williams, Utilitarianism: For and Against (Cambridge University Press: Cambridge, 1973), p. 30.

<sup>2</sup>See Robert and Peggy Stinson, "On the Death of a Baby," The Atlantic Monthly (April, 1979).

<sup>3</sup>Ibid., pp. 35-36.

<sup>4</sup>James Griffin, "Are There Incommensurable Values?" Philosophy and Public Affairs Vol. 7, No. 1 (1977):39-59.

<sup>5</sup>Smart, p. 37.

<sup>6</sup>I owe this gruesome example to Martin Benjamin.

Chapter V

<sup>1</sup>Alan Gewirth, "The 'Is-Ought' Problem Resolved," Proceedings and Addresses of the American Philosophical Association 47 (1974):34-61.

<sup>2</sup>Ibid., p. 35.

<sup>3</sup>Ibid, p. 37.

<sup>4</sup>Ibid., pp. 46-47.

<sup>5</sup>Ibid., p. 52 (*italics mine*).

<sup>6</sup>Ibid., p. 53.

<sup>7</sup>Ibid., p. 54.

<sup>8</sup>Ibid., p. 57.

<sup>9</sup>Ibid., pp. 57-58 (*italics mine*).

Chapter VI

<sup>1</sup>John Rawls, A Theory of Justice (TJ) (Cambridge: Harvard University Press, 1971).

<sup>2</sup>Ibid., see p. 302 for final formulations of these principles.

<sup>3</sup>Ibid., p. 303.

<sup>4</sup>Ibid., p. 152.

<sup>5</sup>See Thomas Nagel, "Rawls on Justice," and Ronald Dworkin, "The Original Position," both in Reading Rawls, ed. Norman Daniels (New York: Basic Books, 1974).

<sup>6</sup>Rawls, Sec. 23.

<sup>7</sup>Dworkin, "The Original Position."

<sup>8</sup>Ibid., pp. 46-53.

<sup>9</sup>Rawls, "Fairness to Goodness" (FG), Philosophical Review (October, 1975):536-554.

<sup>10</sup>Ibid., p. 539.

- <sup>11</sup>Ibid., pp. 547-549.
- <sup>12</sup>Rawls, A Theory of Justice, p. 586.
- <sup>13</sup>Ibid., pp. 180, 505; and Rawls, "Fairness to Goodness," p. 539.
- <sup>14</sup>Rawls, A Theory of Justice, p. 505ff.
- <sup>15</sup>Ibid., p. 207.
- <sup>16</sup>Ibid., Sec. 35.
- <sup>17</sup>Ibid., p. 217.
- <sup>18</sup>Ibid., p. 220.
- <sup>19</sup>Dworkin, p. 50.
- <sup>20</sup>Rawls, "Fairness to Goodness, p. 538.

#### Chapter VII

- <sup>1</sup>The following discussion is taken from J.M. Brennan, The Open-Texture of Moral Concepts (New York: Barnes and Noble, 1977), p. 57ff.
- <sup>2</sup>For a similar view of the relation between rules of inference and particular logical judgments, see Nelson Goodman, Fact, Fiction, and Forecast, 2nd ed. (Bobbs-Merrill, 1965):63-64.
- <sup>3</sup>Norman Daniels, "Wide Reflective Equilibrium and Theory Acceptance in Ethics," Journal of Philosophy (April, 1979):256-282.

#### Chapter VIII

- <sup>1</sup>Jane English, "What Do Grown Children Owe Their Parents?" in Having Children: Philosophical and Legal Reflections on Parenthood, O'Neill and Ruddick, eds. (New York: Oxford University Press, 1979):351-356.
- <sup>2</sup>Ibid., p. 353.
- <sup>3</sup>Judith Thomson, "A Defense of Abortion," Philosophy and Public Affairs, Vol. 1, No. 1 (1971):47-66.

<sup>4</sup>Ibid., p. 52.

<sup>5</sup>Ibid., p. 53.

<sup>6</sup>Ibid., p. 59.

<sup>7</sup>R.M. Hare, "Abortion and the Golden Rule," Philosophy and Public Affairs, Vol. 4, No. 3 (1975); as reprinted in James Rachels, Moral Problems, 3rd. ed. (New York: Harper and Row, 1979):151-173.

<sup>8</sup>Ibid., p. 152.

<sup>9</sup>See my discussion, Chapter II.

<sup>10</sup>Hare, p. 158.

<sup>11</sup>Ibid., p. 159.

<sup>12</sup>Ibid., pp. 163-164.

<sup>13</sup>Ibid., p. 165.

<sup>14</sup>Ibid., p. 172.

<sup>15</sup>Ibid.

<sup>16</sup>Ibid., p. 171.

MICHIGAN STATE UNIV. LIBRARIES



31293101024937