# Rules and Consequences as Grounds for Moral Judgments

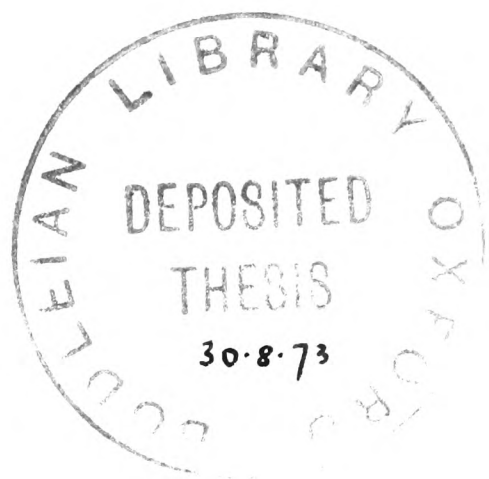
by

Raymond G. Frey, Jr.

A Thesis submitted for the degree of

Doctor of Philosophy in the

University of Oxford




St. John's College,
Oxford

January, 1973

ACKNOWLEDGMENTS

I have profited greatly from discussions with my supervisor, Professor R. M. Hare, whose patient and detailed criticisms of earlier drafts of this thesis have helped immensely to clarify my thinking on the issues involved. His time and efforts on my behalf are deeply appreciated.

I have discussed the subject matter of this thesis with many people, and they have certainly influenced the ideas contained in it. Some of these individuals will be acknowledged in footnotes, to particular points in the text; but I must mention here the many discussions I have had with Professor Henry R. West, whose thinking on utilitarianism has so often stimulated my own.

Of course, in Oxford just now, practical moral philosophy is once more being done in earnest; and I am especially grateful for the opportunities I have had to participate in, and to read papers to, seminars, discussion groups, etc., devoted to the issues of this thesis. In particular, a seminar on Utilitarianism given by Professors R. M. Hare and J. J. C. Smart in 1970 was especially helpful.

Lastly, there are my debts of stimulation in philosophy. There are many to whom I am indebted on this score, but to none more than to Dr. P. M. S. Hacker and Dr. D. A. J. Richards.

## ABSTRACT

My aim in this essay is wholly constructive: it is to present the lines along which a satisfactory utilitarianism may be developed. Such a theory is satisfactory in respect of its being able to over_come or evade objections to previous utilitarianisms, specifically, to previous act_utilitarianisms; I have picked several of these objections to form the rock upon which the strength of a utilitarian is to be tested. The objections in question all center around the question of whether, given his consequential account of rightness, an act_utilitarian can support the useful social rules and institutions of our society; or whether his position, because of its conse_quential account of rightness, commits him to acting in such a way as to undermine these rules and institutions. I shall argue that a new form of act_utilitarianism, which I call tempered act_utilitar_ianism, can both retain its consequential account of rightness and yet can (a) accommodate these rules and institutions within it, (b) allow its proponents on act_utilitarian grounds to advocate ad_herence to them, as providing us with the best chance of doing the right or optimific thing, (c) cater to the views of the 'plain man' in this important respect, and (d) achieve all this without re_course to rule_utilitarianism.

In Chapter One, I distinguish the components of the utilitarian theme and proceed to emphasize the consequence component. The consequence component of all act_utilitarian theories is conse_

quentialism, according to which acts are right or wrong solely in virtue of their consequences; thus, the tempered act-utilitarian is a consequentialist. Two groups of issues arise in connection with his acceptance of consequentialism: one group is concerned with consequentialism itself, with defending it against its rivals, and this is my aim in Chapters II-V; a second group is concerned with the problems, including those about support for useful social rules and institutions, that arise for an act-utilitarian as a result of the inclusion of consequentialism in his theory, and I consider some of these problems from Chapter VI onwards. In Chapter Two, I distinguish two opponents of consequentialism, the internalist, who thinks that some acts are intrinsically right or wrong, and the absolutist, who thinks that some acts may never be done, whatever consequences threaten. Employing a scheme for the redescription of acts, I argue that consequentialism and internalism are not two fundamentally different and logically distinct ways of assessing the rightness of acts; and I suggest that certain of our findings in this chapter cast doubt upon the validity of the absolutist's central claim. In Chapter Three, I consider several internalist moves by which to challenge the consequentialist's use of the scheme for the redescription of acts, but I conclude that none of these moves compels us to give up our reconciliation between internalism and consequentialism. I return to the absolutist in Chapter Four, and argue that several hard cases make an impression against his central claim.

I consider in some detail ways in which an absolutist might try to avoid the impact of such cases, but I express serious doubts about their success. In Chapter V, I consider the doctrine of double effect, a principle of conduct to which, in certain of these hard cases, an absolutist may resort; but I conclude that we have reasons for rejecting the doctrine and so reasons for thinking the absolutist's reliance upon the principle to evade the impact of hard cases to be in vain. I have here indicated only the negative side of the consequentialist's defence of his position against his rivals; but many positive features are of necessity discussed, including his method for allowing circumstances to figure in his consequential account of rightness.

In Chapter VI, I begin the discussion of issues which arise as a result of the inclusion of consequentialism in an act_utilitarian theory, and I discuss the fears of the 'plain man' in this matter. I chart the way forward both positively _ I indicate how a form of act_utilitarianism can take steps to relieve the fears of the 'plain man' _ and negatively _ I indicate that a powerful utilitarian argu_ ment will be brought against untempered act_utilitarianism, so that only by taking the aforementioned steps to meet the 'plain man' will a form of act_utilitarianism in practice remain viable. In Chapter VII, I exhibit these steps an act_utilitarian can take to removes certain fears of the 'plain man', and I argue that these

steps mark off a new species of act-utilitarianism. The tempered act-utilitarian's rejection of optimificity as a thesis of justification, his distinction between rightness and justification, and his very great emphasis upon practice (upon equipping ourselves for the sorts of cases that we are likely to meet with) are discussed in this chapter. In Chapter VIII, in an attempt to bar any relapse from a tempered into an untempered species of act-utilitarianism, I offer a utilitarian argument to establish the disutility of untempered species. I argue that the general practice of untempered act-utilitarianism will undermine the useful social institution of promising and, by implication, other useful social institutions; and I conclude that the general practice of such a theory produces disutility of a degree that is not negligible. In Chapter IX, I consider a defence to the utilitarian argument, a defence that I show to rest very firmly upon features, particularly the thesis of linearity, that figure very prominently in David Lyons' well-known argument to demonstrate the extensional equivalence between forms of utilitarian generalization and forms of act-utilitarianism. I provide two arguments to show that this defence may be breached; they show that the utilitarian argument of Chapter VIII may be maintained against the untempered act-utilitarian whether or not linearity obtains between the effects of a single act and the effects of a series of such acts. I examine in more detail, by going more deeply

into Lyons' Forms and Limits of Utilitarianism, the theses of linearity and non-linearity in Chapter X; and I maintain there that Lyons has by no means made out his case in respect of linearity. I also indicate further difficulties in Lyons' general position on utilitarian consequences and on threshold effects, on their production and distribution; and I use distinctions between causal and contributive properties, causal and contributive consequences, and individual and collective acts to build the case of the non-linearian against Lyons. I conclude the chapter by finding in its discussion of these topics additional grounds for thinking tempered act-utilitarianism attractive.

In Chapter XI, I once again take up the question of justifying exceptions to useful social institutions. I provide the tempered act-utilitarian with a means for distinguishing justified from unjustified claims to exceptions, which he requires, since he rejects any reliance upon optimificity as a thesis of justification; and I conclude the chapter with a discussion of the demands of logical consistency and role-playing in this matter of exceptions to institutions. Finally, in Chapter XII, I say something more on the rules to be found in tempered act-utilitarianism - for throughout I shall be trying to make plain the characteristics of these rules; I emphasize that the utility of the rules to be found in tempered act-utilitarianism resides in their generality, and that degree of generality or, conversely, specificity, in rules may be a point of

division between tempered act-utilitarianism and forms of rule-utilitarianism.

A tempered act-utilitarian, then, will emerge in these pages both as a consequentialist and as an act-utilitarian, though with differences from the usual sorts of each; and he will be seen to be a champion of the rules and institutions of the 'plain man' at the same time.

CONTENTS


| | | |
|---|---|---|
| I. | THE UTILITARIAN THEME | 1 |
| | 1. Objections to AU theories | 1 |
| | 2. Complexities in utilitarianism | 4 |
| | 3. The utilitarian theme | 11 |
| II. | CONSEQUENTIALISM, INTERNALISM AND ABSOLUTISM | 17 |
| | 1. Distinctions | 17 |
| | 2. Re_description of acts | 22 |
| | 3. Internal and external consequences | 26 |
| | 4. Deontology and teleology | 32 |
| III. | ACT DESCRIPTION AND INDIVIDUATION | 33 |
| | 1. The nugget theory | 33 |
| | 2. Sameness of act | 37 |
| | 3. Abstract act_descriptions | 45 |
| IV. | ABSOLUTISM AND HARD CASES | 49 |
| | 1. Hard cases | 49 |
| | 2. Distinctions | 50 |
| | 3. Breaking down and harm befalling others | 61 |
| | 4. A verbal ploy | 70 |
| V. | DOCTRINE OF DOUBLE EFFECT | 73 |
| | 1. Conditions for its use | 73 |
| | 2. Criticisms, cautions and clarifications | 77 |
| | 3. Control and responsibility | 88 |
| | 4. Acts of omission | 92 |
| | 5. Acting and refraining | 98 |
| | 6. Positive and Negative duties | 101 |
| | 7. Place in the causal series | 104 |
| | 8. Legal Doctrine of Necessity | 108 |
| VI. | ACT_UTILITARIANISM AND THE HARD CASES ARGUMENT | 113 |
| | 1. The 'plain man's' fear | 113 |
| | 2. Impasse | 118 |
| | 3. The way forward | 120 |
| VII. | TEMPERED ACT_UTILITARIANISM | 122 |
| | 1. Philosophy and practice | 122 |
| | 2. Ignorance, weakness and optimificity | 125 |
| | 3. Rightness and justification | 128 |
| | 4. Essential rules | 133 |
| | 5. Moral rules: some doubts | 135 |

VIII. THE DISUTILITY OF UNTEMPERED ACT_UTILITARIANISM 140

1. Hodgson's argument 141
2. A preliminary to the argument 145
3. A utilitarian argument 151

IX. A DEFENCE TO THE UTILITARIAN ARGUMENT 158

1. UAU and Lyons' argument 159
2. Reply to the defence 165
3. The late ascription of threshold effects 167

X. UTILITARIAN CONSEQUENCES 170

1. Lyons and linearity 170
2. Causal and contributive consequences 172
3. Individual and collective acts 175
4. The non_equalizer 179
5. In prospect and in retrospect 186

XI. JUSTIFICATION AND EXCEPTIONS 192

1. Justification and consistency 192
2. Discharge restrictions and roles 196

XII. GENERAL RULES 204

1. RU and TAU 204
2. Cooperation 206
3. Cooperation and rules 209
4. Generality in rules and exceptions 213

FOOTNOTES 215

BIBLIOGRAPHY 241

# I

## THE UTILITARIAN THEME

After a long and more or less exclusive diet of meta-ethics, the search for an adequate normative ethical theory has been renewed in earnest, with the result that a good deal of attention has once again been focused upon the subject of utilitarianism and upon attempts to develop a satisfactory utilitarian theory. My aim in this essay is to articulate such a satisfactory form of utilitarianism; my method consists simply in examining a few objections to traditional act-utilitarian theories, and then presenting a type of utilitarianism, what I call tempered act-utilitarianism, which evades these objections. The only other remark about method that warrants a mention is that the objections to be canvassed and the views to be discussed are not approached in their historical setting: so much has become familiar in this area in the past few years that very little is lost to comprehension through leaving aside historical concerns.

### I.1 Objections to AU theories

In order to work out a satisfactory utilitarianism, it is first necessary to understand what considerations have rendered previous forms unsatisfactory. In much recent work, the alleged defects of classical- or act-utilitarianism (AU) have received a thorough airing, and their removal has spurred the somewhat bewildering proliferation in types and versions of types of rule-utilitarianism (RU) and utilitarian generalization (GU). Part of the motivation behind these developments has been the conviction, widely held (or so one gathers

from much of the writing in this area), that if only we continue to explore different types and versions of utilitarianism, eventually we shall work out a theory which (a) meets many of what can only be described, because of their wide exposure and pervasive influence, as the 'standard' objections to AU and (b) contains, therefore, none _ or very few _ of the defects ascribed to AU. With the objections met and the defects purged, the theory finally worked out will be put forward as a satisfactory utilitarianism and, if it meets certain other require_ ments, as an adequate normative ethical theory. This is the conviction, and, too, the explanation of why recent utilitarianisms take their start from criticisms of AU. At present, however, all this is conviction only; the fact of the matter is that the continuing development of new utilitarian theories and the reformulation of old ones are undoubtedly an all too accurate confession that no theory, whether of the AU, RU or GU variety, has yet been worked out which overcomes the 'standard' ob_ jections. Moreover, the very severity of the attack upon these new theories, that they are open to the old objections or that they give rise to new ones, that they 'collapse' into AU or that they are 'extensionally equivalent' to it, suggests that we might at this stage in the search take another look at AU and its alleged defects, to see if a form of AU might not be worked out which is free of the old defects and not open to the old objections. Nor is such a proposal out of the question: the measure of the satisfactoriness of a utilitarian theory has been its degree of success in meeting what I am calling the 'standard' objections, and there is at least no logical bar to a form of

AU achieving such success. Indeed, I think a form of AU can be worked out/meets those 'standard' objections, at least those that we shall consider.

For the list of significant objections to AU is far too lengthy to consider them all; instead, I shall select one or two that have become increasingly important in the search for a satisfactory utilitarianism and examine them. These objections, therefore, are to form our obstacle course, for I shall assume that no version of utilitarianism will ever be a satisfactory one, however it fares at the hands of other criticisms, unless it meets these objections. On the other hand, should a version be worked out which does meet them, then I shall assume that we have a powerful reason for adopting it. Of course, we may demur until it is tested in the face of other objections; but the ones I shall consider, in order of importance as criticisms of AU and as problems for any type of utilitarianism, stand in the forefront. The difficulties at issue center around the question whether, given his consequential account of rightness, the AUian can support the useful social rules and institutions of our society; or whether his position commits him to undermining them.

It should be obvious that there are many aspects to and difficulties with the multifarious forms of utilitarianism that are to be ignored here. My justification for this neglect rests in part on the importance to utilitarian theories of the issues to be considered, and in part on a certain shift of emphasis that has crept into recent work on utilitarianism and with which our central concerns are in line. The basis of the shift is the realization that 'utilitarianism' does not refer to a single theory

but to a group of theories which are variations on a theme, the components of which may be distinguished. The shift has been away from one component in favor of another, and the concerns of this essay arise when emphasis is placed upon this other component. It will be useful to fill in the details of this shift.

## I.2 Complexities in utilitarianism

The term 'utilitarianism' has recently come to be applied to a number of different ways of assessing the rightness of acts; David Lyons, for example, in his Forms and Limits of Utilitarianism[1], distinguishes sixteen different versions, though these are usually divided into three main types. First, there is AU, according to which an act is right if, and only if, its consequences are at least as good as those of any alternative and, as a corollary, that an act ought to be performed if, and only if, its consequences are better than those of any alternative. Then there is RU, according to which an act is right if it conforms to a rule the general following of which would have (or has) good conse_quences; clearly, the whole thrust of RU is to de_emphasize the role of consequences and to emphasize the role of rules in determining the rightness or wrongness of acts. Finally, there is utilitarian general_ization or GU, according to which an act is right if it is of a kind which would have (or has) good consequences if acts of that kind were performed generally; the whole thrust of this type of theory is to de_emphasize the role of actual consequences and to emphasize the role of the consequences in a hypothetical situation in which an act of that kind is being generally performed. In fact, whether GU is a distinct type of

utilitarianism, or whether it is a version of RU, is a moot point, and one which, I suspect, is largely verbal; in any event, nothing of any significance in this essay turns upon the outcome. The important thing is that what have been given are general characterizations of types of theories, and that these can be qualified in certain respects to generate different versions of each type. For example, the general characterization of AU may be qualified in (at least) two ways, in respect of consequences and of standards of goodness; and in respect of consequences, the following are among the more important qualifications an AUian might wish to make. (1). The view that consequences make acts right or wrong may be formulated in terms of actual or probable, expectable or intended consequences. (In avoiding actual consequences, an AUian might hope to avoid the objection that, since we cannot know all the actual consequences of any act, we cannot know whether any act is right or wrong.) (2). The act's consequences considered relevant to its rightness may be confined to those that affect human beings, so that the effects of our acts on animals will not be considered. Or perhaps a weighing system is preferable, whereby affected humans count for more than affected animals. (3). For the purposes of determining rightness, an act's consequences may be restricted in another way: limits may be imposed by specifying where in the causal chain of consequences those considered to be relevant begin and end, so that only those which fall within these limits will figure in the assessment of rightness.[2] (In this way might an AUian hope to avoid the objection that, because the consequences of an act extend into the indefinite future, we cannot tell at the time of doing it whether

any act is right or wrong.) There is a look of arbitrariness about this maneuver: why begin just here in the causal chain? And why end just there? And would any AUian refuse to include in his assessment of consequences those of immense utility or disutility which fell just outside the limits imposed? (4). Failing the imposition of limits, an AUian may adopt a 'ripples in the pond' postulate, according to which the remote consequences of acts tend to diminish to zero, and so may be left out of our assessment.[3] However, this postulate has not been proven: it is an assumption[4], the only visible ground for which is that more often than not remote consequences do not affect decisively our previous assessments of rightness. Whether this is the case is a factual matter, and one's opinion may be colored by the range of examples one considers. Then, too, as it stands, the postulate is ambiguous: the assumption that remote consequences <u>tend</u> to diminish to zero does not make clear whether this is a judgment of frequency _ that in N percentage of cases of <u>all</u> types remote consequences approach zero _ or a judgment about certain kinds of acts _ that with certain kinds of acts but not others remote consequences approach zero. And the fact that we can often predict the remote consequences of acts, and so take them into account, also weighs against the postulate. (5). The thesis that consequences make right acts right, instead of being formulated in terms of the production of good consequences, may be put in terms of the avoidance of bad consequences, and one may go on to generate a kind of negative utilitarianism.[5] Certainly, this course is superficially attractive: we may feel that we are under a duty to avoid

inflicting pain and suffering but not to increase the amount of pleasure (happiness, utility, etc.) in the world; and there is obvious merit in fostering a concern with minimizing suffering. There have, however, been few attempts at developing a negative utilitarianism, one explanation for which may possibly be found in R. N. Smart's contention that, if a negative utilitarian is to maximize the minimization of human suffering, he seems committed to the extermination of the human race.[6] (6). Again, the thesis about consequences and rightness may be formulated in a comparative (better or worse) rather than in a non_comparative (good or bad) sense. This maneuver may allow the AUian to deal with those situations described as 'choosing the lesser of two evils', in which none of the courses open to the agent have good consequences, though one may have better consequences than the others. These, in short, are but a few of the possible qualifications in respect of consequences to the general characterization of AU that can result in different versions.

It is also possible to generate different versions of AU through employing different standards of goodness, at least one of which is required in any AU theory, since an act is right only if its consequences are at least as <u>good</u> as those of any alternative. Very broadly, there are hedonistic and non_hedonistic or ideal standards of goodness, though further distinctions must be drawn within each camp. For example, Bentham's hedonism does not only include physical pleasure, for he gives the word 'pleasure' a broad meaning, so as to include among other things, such (admittedly slippery) notions as benefit, advantage and happiness.[7] As is well_known, both Mill and Sidgwick subsequently adopted pleasure_

standards, but with differences: Mill distinguishes between qualities of pleasure[8], whereas Sidgwick introduces a principle of justice in addition to and independent of the pleasure-standard, so as to evidence in his theory a concern with the distribution as well as the maximization of pleasure.[9] Differences also exist in the non-hedonistic or ideal camp. Stace, for example, formulates his standard of goodness in terms of happiness, which includes but is not exhausted by pleasure[10], whereas both Rashdall[11] and Moore[12] hold that a number of things - not simply pleasure and happiness - are good-in-themselves. More recent writers have by and large followed Moore in thinking that various things besides pleasure and/or happiness are good-in-themselves, hedonism, for a combination of reasons, having fallen into disrepute.[13] On two important issues, however, the hedonistic and ideal camps agree. First, whichever standard of goodness is adopted by AUians, it will be a standard of <u>intrinsic goodness</u>; thus, Stace's view is that all and only happiness is intrinsically good or good-in-itself. Accordingly, while AU theories offer a consequentialist account of rightness, they offer a non-consequentialist account of goodness. Of course, this conviction that some things are good-in-themselves will come under stress to the extent to which the concept of intrinsic goodness itself comes under attack. For example, both John Dewey in his <u>Theory of Valuation</u>[14] and Monroe Beardsley in a paper on 'Intrinsic Value'[15] seek to establish that there is no such thing as intrinsic value and that there is nothing in the world that is intrinsically good, and either thesis, if established, would bear upon the AUian's thesis that acts are right or wrong in virtue

of the intrinsic goodness or badness of their consequences. Second, the two camps also agree over the identity of those things that are intrinsically good. The view which has underlain nearly all formulations of the principle of utility, beginning with Bentham and Mill, is that the only things intrinsically good are certain states of mind, which are severally referred to as 'pleasure', 'happiness', etc. There may be disagreement over which states of mind are intrinsically good, but that the principle of utility is concerned with the production of intrinsically good or bad states of mind, and that states of mind are the only things good_in_themselves, has amongst utilitarians rarely been a matter of controversy. Nevertheless, one may object _ an objection to which I am sympathetic _ that to confine utility merely to a concern with states of mind is too narrow, and that the notion should be expanded to include, for example, the satisfaction of interests.[16.] One advantage to formulating one's standard of goodness in terms of satisfying interests is that it broadens the AUian's net, and enables him to include within his theory the suggestion that, at least with ordinary people, satisfaction of interests is a factor in determining the rightness of acts.

What is true of AU is equally true of GU and RU: different versions of each may be developed through qualifying their respective general characterizations. Of course, much of what was said of AU applies to RU and GU, i.e., different standards for evaluating the goodness of consequences yield different versions of each theory; but there are some important differences. One might hold a type of GU according to which an

act is right if it is of a kind which would have (or has) good consequences, if acts of that kind were performed generally (by people in general, though not by each and every one) or, on the other hand, if acts of that kind were performed universally (by people individually, by each and every one). Likewise with the general characterization of RU: different versions may be generated depending on whether one refers to the general or universal following of a rule. Still additional versions of RU may be developed by making rightness depend not upon the utility of a rule's being followed, but upon that of its being recognized, and here too one might distinguish between general and universal recognition.[17] Then, too, one may distinguish those versions of RU which are formulated in terms of a set of rules which would maximize utility if they were in force, sometimes called ideal-RU, from those versions formulated in terms of a set of rules actually in force, sometimes called conventional-RU.[18] And one may want to distinguish from the others those versions of RU formulated around the notion of a practice, i.e., those which distinguish between justifying a practice - promising and truth-telling are the usual examples - and justifying acts falling under the practice.[19] Of course, the practice may be an ideal one, one which would maximize utility if it were adopted (or recognized) by a society, or an actual one, one that is presently operative in a society[20]; and to both of these it must be added that radically different analyses of what a practice is may be and have been given. Finally, if instead of discussing the adoption or recognition or following of a rule or set of rules or a practice by a society one focuses upon the adoption

or recognition or following of a rule or set of rules or a practice by an individual, other versions of RU may be generated, some of which are set out in D. H. Hodgson's Consequences of Utilitarianism.[21]

The above, then, are but a few of the complexities that have been introduced into the discussion of utilitarianism; yet even these few are sufficient to illustrate how different are the views labelled utilitarian. In fact, one may say that the term 'utilitarianism' refers not to a single, widely accepted theory but to a cluster of theories that has grown up out of alleged shortcomings in AU and of attempts to put these right. So there is nothing which is utilitarianism per se; there is simply a cluster of theories that are variations on a theme. This is the point from which the quest for a satisfactory utilitarianism (and thereafter any defense of a utilitarianism as an adequate normative ethical theory) must begin; and if one is to call oneself a utilitarian, it is necessary to articulate and defend the particular version that one espouses.

## I.3 The utilitarian theme

What, then, is this utilitarian theme? At the very least, it involves three components: (i) that the rightness or wrongness of acts is tied in some way to (the production of) good and bad consequences - which I shall call the consequence component; (ii) that the goodness or badness of consequences is to be evaluated by means of some standard of intrinsic goodness - which I shall call the value component; and (iii) that it is the act's consequences as affecting everyone and not just the agent himself that are to be considered in determining whether an act is

right or wrong. Something like (iii) is necessary, for not every theory which contains consequence and value components of the sort sketched is utilitarian in character. So far as I can see, the only major difference between ethical egoism and AU is that the ethical egoist is concerned with maximizing utility in his own case, so that only consequences that affect him bear on the rightness or wrongness of the act; the two theories are alike in having consequence and value components, and may even share identical particular components. Whereas in a utilitarianism everyone affected by the act is to be considered, and, on the classical utilitarian assumption that each is to count for one and no more than one, to be considered equally.

In the past, the value component occupied utilitarians and their critics, whereas today the selection and formulation of a standard of intrinsic goodness or principle of utility no longer excites very much interest. Of course, a good deal of past concern with the value component stemmed from the hedonism of the classical utilitarians, and the fact that ethical and psychological hedonism have fallen on hard days has surely contributed to this decline in interest. Nevertheless, what this decline has meant, and the reason it is important here, is that a whole series of issues has ceased to form the testing-ground of the satisfactoriness of a utilitarianism. Arguments, for example, (i) about hedonism, and whether we should strive to live the life of a pig or of Socrates; (ii) about whether there are qualities of pleasures; (iii) about the formulation of ideal standards of intrinsic goodness; (iv) about how

many constitutes the greatest number; (v) about how one selects the group within which one will maximize the happiness _ or pleasure or whatever _ of the greatest number[22]; (vi) about whether we should strive for the greatest total or greatest average happiness; (vii) about what 'happiness' or 'pleasure' or 'benefit' or 'utility' mean; (viii) about whether there are units of happiness or pleasure or whatever that enables one to calculate; (ix) about whether and how interpersonal comparisons of utility can be made[23]; (x) about whether calculating with or without a ready calculus can be done at all in the case of intrinsic goodness; (xi) about how one finds out which things are intrinsically good; (xii) about what is meant by speaking of something as 'intrinsically good' _ these arguments and a host of other issues surrounding the value component no longer form the main battleground for determining the acceptability of a utilitarianism. Curiously enough, the reason for this cannot be traced to any agreement on solutions to these problems, i.e., to any agreement upon, e.g., the greatest total or greatest average principles, upon a standard of goodness or upon an accepted formulation of the principle of utility: there is no formulation, whether in terms of pleasure, happiness, minimizing suffering, etc., considered free of difficulties. Rather, the shift of emphasis away from the value component has stemmed from the realization that, even if we came to agree upon a standard of goodness by which to evaluate consequences, the consequence component itself may prove unacceptable. For it, too, can appear the source of problems; and it is <u>these</u> problems that have come to supplant those surrounding the value component as the testing_ground of

the satisfactoriness of a utilitarianism.[24]

As one would expect, these new problems have arisen mostly over the consequence component found in AU theories. Phrased in a general way, that the rightness of acts is tied in some way to (the production of) good and bad consequences, the consequence component is vague; for so long as discussion remains on a general level, it seems absurd to deny that consequences have something to do with rightness.[25] But what have they to do with rightness? It is over the reply to this question that the AUian differs from other utilitarians and from many non-utilitarians; for the AUian response has been that acts are right or wrong solely in virtue of their consequences. That is, an act's consequences make it right or wrong. A consequence component of this sort is strong, since consequences alone determine rightness, and to align myself with current practice, I shall call the view expressed by this strong consequence component consequentialism. It is most important to be clear on the fact that consequentialism is not utilitarianism nor utilitarianism consequentialism. One may be a consequentialist without being a utilitarian, as is the case with an ethical egoist, and, I shall argue later, with some forms of deontology; and one may be a utilitarian without being a consequentialist, as is the case with any RUian or GUian, neither of whom think that acts are right or wrong solely in virtue of their consequences. Because of the shift of emphasis to the consequence component of the utilitarian theme, of the problems that arise in connection with it having come to form the testing-ground of the satisfactoriness of a utilitarianism, and of the attention thereby drawn to consequentialism as

an important element of AU, some AUians today may tend to blur the distinction between consequentialism and utilitarianism. For example, J. J. C. Smart begins his paper on 'Extreme and Restricted Utilitarianism' with the statement that 'utilitarianism is the doctrine that the rightness of actions is to be judged by their consequences'.[26] This is a mistake: if the passage is intended as a statement of what AU consists in, then it overlooks the necessity for a standard of goodness by which to evaluate consequences; if it is supposed to refer to <u>any</u> utilitarianism, then it is incorrect in the cases of RU and GU; and if it is intended to distinguish utilitarianism as a type from other normative ethical theories, why then it overlooks ethical egoism and some forms of deontology, which can have the same view about rightness and consequences.[27] On the contrary, the truth of the matter is this: a <u>consequence component</u> _ that the rightness or wrongness of acts is tied in some way to (the production of) good and bad consequences _ is one component of the utilitarian theme and as such will be a feature of all utilitarianisms; a <u>strong consequence component</u> is the consequence component of AU and, amongst utilitarianisms, AU only, for only AUians think that acts are right or wrong <u>solely</u> in virtue of their consequences. Accordingly, consequentialism is <u>not</u> to be identified with the utilitarian theme or, for that matter, with AU.

In this essay, I align myself with the shift of emphasis away from the value component of the utilitarian theme: the issues I shall consider all arise in connection with the consequence component, specifically, with

the strong consequence component (consequentialism) to be found in AU theories. For, as I have already remarked, I believe a species of AU can be defended against some of the 'standard' objections to AU, and the objections I am concerned with arise in connection with its consequence component; so it is with the issues surrounding this component that I must deal. These fall into two groups. First, there are those issues that arise in connection with consequentialism quite independently of its inclusion in a utilitarianism and of a kind that even a non-utilitarian consequentialist would have to consider. In essence, they are concerned with defending consequentialism against its rivals, i.e., other accounts of what makes right acts right and of what may be done; and it is with this aim that I shall consider them in Chapters II-V. They will to some extent be discussed for their own sake, but their effect on the utilitarianism in which consequentialism is a component will always be in the background, as it must, if consequentialism is the consequence component of any and all AU theories, including tempered act-utilitarianism. Second, there are those issues, which include questions of support for useful social rules and institutions, that arise in connection with AU theories as a result of their inclusion of consequentialism, and these will be discussed with the aim of vindicating our new species of AU, from Chapter VI onwards. Most importantly, our new species of AU will be seen not to be susceptible to the utilitarian argument we shall bring against previous species.

# II

## CONSEQUENTIALISM, INTERNALISM AND ABSOLUTISM

Two supposedly different and incompatible views of the role of consequences in determining the rightness of acts are *consequentialism* and what I shall call *internalism*. Consequentialism (C), to recall, is the view that acts are right or wrong solely in virtue of their consequences; internalism (I), on the other hand, denies the truth of C and affirms instead that some acts are intrinsically right or wrong, right or wrong in themselves, apart from their consequences. In this chapter, I shall argue both that C and I are *not* two fundamentally different and logically distinct ways of assessing the rightness of acts and that a means exists by which to reconcile I with C.

### II.1 Distinctions

C and I cannot stand alone: neither, of course, is a complete normative theory, but nor are they complete accounts of what makes right acts right; each requires supplementation. In the case of C, a means of assessing or evaluating consequences is required, which may be provided through the inclusion of C in AU theories; however, a standard for evaluating consequences, though a necessary feature of the utilitarian theme, forms no part of C. All that is required to make one a consequentialist (*Cist*) is that one maintain that rightness is solely a function of consequences. I also wants supplementation: the claim that some acts are intrinsically right or wrong requires a *means* of identifying *which acts* these are, and such a means may be provided through the inclusion of I into deontological normative

theories. In the deontology of Ross[1] and Prichard[2], this means is provided by their intuitionism. However, intuitionism is no part of I: all that is required to make one an internalist (Iist) is that one maintain that some acts are right or wrong in themselves, apart from their consequences; how one tells which acts these are, whether by reference to intuitionism or perhaps natural law, is another matter.[3] Thus, just as C can be treated independently of AU theories, of which it is a component, so I think I may be treated independently of deontological theories, of which it is a component.

An Iist rejects the thesis that rightness is a function only of consequences, and may hold either (a) that rightness is a function of a number of things, including consequences, or (b) that it is a function of considerations other than consequences. Of these two variants, the first has been by far the more popular and may be identified with Ross; for today it is usually rendered in terms of the distinction between acts right prima facie (or prima facie duties) and right absolutely (or absolute duties).[4] Importantly, this first variant does not sever completely the link between rightness and consequences; and in The Right and The Good, Ross allows that some acts, perhaps the majority, are right or wrong consequentially, and that even those acts which fall under the seven types of prima facie duties, and so are not right or wrong consequentially, may on occasion, e.g., in conflict situations, be judged right or wrong on the strength of consequences.[5] The less popular, second variant of I, according to which rightness is a function of con-

siderations other than consequences, should be restricted in application, to appear plausible, to only some acts, so that only some acts are held to be right or wrong independently of consequences. Thus restricted, this second variant is sometimes confused with what I shall call absolutism (A), according to which there are some acts (or kinds of acts) it would always be wrong to do, no matter what the consequences.[6] The confusion has arisen because those who have been absolutists (Aists) have also been among those who have held the second variant of I; for example, in his essay 'On a Supposed Right to Lie from Altruistic Motives', Kant appears to hold both[7], as do Professor Anscombe, in her article on 'Modern Moral Philosophy'[8] and her pamphlet on 'Mr. Truman's Degree'[9], and Professor Geach, in his paper on 'The Moral Law and the Law of God'[10]. Nevertheless, an adherent to the second variant of I is not logically compelled to adopt A: from the fact that one maintains that some acts are right or wrong independently of their consequences, it does not follow that these acts may never be done or are proscribed absolutely; for this to follow, some statement to the effect that acts right or wrong independently of their consequences are proscribed absolutely must be inserted, and such a statement forms no part of either variant of I. Put differently, if one is an Aist and holds that some acts are proscribed absolutely, one will almost certainly hold the second variant of I; but if one is an Iist of the second variety, one is not logically committed to A. I in both its varieties is a thesis about what makes right acts right; A is a thesis about what may be done;

and A clearly goes beyond I, since it adds the thesis that some of those acts right or wrong in themselves are also ones that may never be done, whatever consequences threaten. This distinction between I and A is important here for two reasons. First, though the Cist has a disagreement with each camp, they are not the same disagreement. To repeat, what is at issue between the Iist and Cist is what makes right acts right; what is at issue between the Aist and Cist is whether some acts (or kinds of acts) are such that they may never be done. Second, as a result of this difference, the reconciliation of I with C that I shall offer will not resolve the dispute between A and C; but I think it will cast doubt upon the truth of A, at least on what I regard as the central claim of A.

What I take to be the essence of A, though to be found elsewhere, e.g., in R. W. Beardsmore's 'Consequences and moral worth'[11], may best be seen in Professor Anscombe's seminal paper on 'Modern Moral Philosophy'. She objects there to 'every one of the best known English academic moral philosophers', for, according to her, they all teach a philosophy that would leave open to question the judicial execution of the innocent; and 'though the present Oxford moral philosophers would accord a man permission to "make it his principle" not to do such a thing, they teach a philosophy according to which the particular consequences of such an action could "morally" be taken into account by a man who was debating what to do...'.[12] (Professor Anscombe returns to this theme in her radio talk 'Does Oxford Moral Philosophy

Corrupt Youth?'[13]) In this passage, it is C to which Professor Anscombe is objecting and in opposition to which she mentions the 'Hebrew-Christian ethic':

> ...it has been characteristic of that ethic to teach that there are certain things forbidden whatever consequences threaten, such as: choosing to kill the innocent for any purpose, however good; vicarious punishment; treachery (by which I mean obtaining a man's confidence in a grave matter and then betraying him to his enemies); idolatry; sodomy; adultery; making a false profession of faith. 14

She then goes on to state what I take to be the central claim of A - if you like, a more precise way of putting the view that there are some acts (or kinds of acts) it would always be wrong to do, whatever the consequences:

> the prohibition of certain things simply in virtue of their description as such-and-such identifiable kinds of action, regardless of any further consequences, is certainly not the whole of the Hebrew-Christian ethic; but it is a noteworthy feature of it.... 15

The view that there are descriptions of acts which, merely being the descriptions they are, eliminate as morally significant, in trying to decide what to do, any other possible description of those acts is rejected by the Cist, for a reason shortly to emerge. The conclusions of this chapter will not be sufficient to dispatch this view; but they do suffice, I think, to render it suspect.

## II.2 Re-description of acts

Internalists (Iists) have talked of acts as right or wrong 'in themselves', as 'intrinsically right', as right 'apart from their consequences'; and it is largely through their efforts that the expressions 'the nature of the act' and 'the description of the act' have gained a currency. However, what these phrases mean, particularly the last two, is unclear, in the light of the fact that of any particular set of events it is possible to offer innumerable descriptions, all different, yet all true. For example:

(1) John tensed his forefinger
(2) John pulled the trigger of a gun
(3) John shot Jim
(4) John killed Jim
(5) John murdered Jim
(6) John widowed Jim's wife.

Descriptions (1) - (6) all describe the same set of events, what John did, and it would be quite arbitrary to select one and say that it is 'the' description or the 'correct' description of what John did, just as it would be arbitrary to say of any one of them that it captures the 'nature' of the act. (Indeed, what is meant by 'the' act?) Nevertheless, I believe these typically Iist expressions can be given a meaning, consistent with the possibility of redescribing the same set of events.

The scheme I have in mind is this. If someone referring to the events in question picked (5) - 'John murdered Jim' - as what John did, as the act he was referring to, then (6) - 'John widowed Jim's wife' -

will be a consequence of the act, while descriptions (1) - (4) will in this scheme be part of the circumstances of the act, whereby I mean to include, perhaps unusually, the conditions that surrounded the act's performance at a particular time. Likewise, if one selects (3) - 'John shot Jim' - as what John did, then descriptions (4) - (6) will denote consequences of the act, while (1) and (2) will denote part of the circumstances or conditions that obtained at the time the act was performed. In other words, we are able to draw the line between act and consequence at different places in the causal chain, and what one is doing in using the expression 'the description of the act' is selecting from among the variety of possible act-descriptions that act-description under which one will refer to what happened. In selecting <u>this</u> act-description and referring to it as 'the description of the act', one is locating 'the act' in the sequence of circumstances and consequences, as with 'John murdered Jim'. The act-description selected ties what we are to take as 'the act' to a <u>particular</u> set of circumstances and consequences, and the expression 'the nature of the act' is simply a way of referring to <u>this</u> description and to <u>this</u> set of consequences to which it is tied. In selecting a particular act-description, one is <u>picking out as important</u> certain features of the situation; what counts as important, what features are picked out, depends upon one's aim or purpose in trying to find out what John did. Leaving aside non-moral aims, the

usual aim is that of assessing responsibility for what happened, and what is picked out as important will be those features of the situation one thinks relevant to this aim. Which features these are will vary from case to case: sometimes, it may be thought relevant to mention tensing the forefinger or flexing certain muscles; more often, however, we want and require more than this.[16] In John's case, to describe what he did as 'tensing his forefinger' or 'pulling the trigger of a gun' would not be as complete an act-description as could be given, or as probably would be demanded. As a rule of thumb, probably the wisest course for assessing responsibility is to cast one's net as widely as possible, and so to select a description later instead of earlier in the chain.

The fact that the act/consequence line can be drawn at different places, in different ways, according to what one desires to pick out as important can evoke the fear, as A. B. Cody has suggested[17], that wrongful deeds will go unpunished, that those who perpetrate such acts will evade responsibility for them through redescribing them, in such a way that what formerly appeared as bad now appears as good or morally neutral. In my view, this fear is groundless; it arises only on the mistaken assumption that the line between act and consequence is the <u>same</u> as the line between the morally relevant and the morally irrelevant: if one is responsibile for one's <u>acts</u>, one may possibly evade responsibility for an evil one by redescribing it in

such a way as to render a consequence what was previously (a part of) the act. It is a mistake, however, to think that the act/consequence line separates morally relevant from morally irrelevant features of the situation; for one can be held responsible (and also blamed) for bringing about a particular set of consequences. A terrorist who opens fire on a crowd, in the midst of which is an enemy leader, and who kills the leader and six other people, can still be held responsible (and also blamed) for these six other deaths. Even if one is inclined to say that the act/consequence line separates the moral from the non-moral features and that one is responsible only for one's acts, it is still the case that responsibility for bringing about a set of consequences can be allotted. Moreover, if in contemplating his act the terrorist foresees this set of consequences or these deaths coming to pass, then that is at least a reason for his not doing it, and a reason that is surely morally relevant in his deciding what to do. In other words, considering the series of events which together comprise act and consequences, we may say, with H. L. A. Hart[18], that a man is responsible not only for those consequences he intends but also for those he foresees. (Of course, we may also say that, even if a man did not foresee what was to happen, he should have.) Nor can the terrorist escape responsibility for the six deaths by arguing that he did not

intend to kill the six since he did not want to kill them: what the agent wants need not be relevant to what he intends[19], and though the terrorist in no way wants to kill the six (if by some chance they had escaped, he would not have tried to kill them), still, if they are killed, he is responsible. To be responsible for their deaths, it is enough that he foresaw that they could be killed and opened fire. (I consider intention and foresight, direct and indirect intentions, and senses of responsibility in Chapters IV and V.)

## II.3 Internal and external consequences

It is sometimes held that there are two conditions necessary for something to count as a consequence: first, the state of affairs or event which is the alleged consequence must be subsequent in time to the act, and second, this state of affairs or event must be causally related to but logically distinct from the act.[20] It is this second condition, with its requirement of logical distinctness, that a Cist wants to reject, since it rules out as consequences/items he wishes to include. For example, he wants to include as consequences of the act described as 'opening the window' both that there are now flies in the room and that the window now stands open, of the act described as 'John killed Jim' both that Jim's employer has a vacancy to fill and that Jim ends up dead. However, whereas consequences of the former sort, which

I shall call external consequences, are logically distinct from the act, consequences of the latter sort, which I shall call internal consequences, are logically dependent upon the description under which the events in question are being referred to.[21] That is, which consequences are internal depends upon how the events are described: change this description and the internal consequences change. For a Cist, then, consequences may be either internal or external; both are to be considered in assessing rightness. Moreover, it obviously becomes pertinent to consider a number of alternative descriptions of a proposed act, since in this way one is able to assess both the internal and external consequences of the act under different descriptions; and the reason why it is morally relevant to consider various possible, alternative act-descriptions is because the internal and external consequences will differ under each description, and an agent will want to be sure just what these differences come to when he is deciding what to do.

There must be a means of evaluating these consequences, so let us assume our Cist is also an AUian, in which case a right act for him is one that produces a greater balance of utility over disutility. With the act described as 'John killed Jim', the Cist/AUian will say that, unless the external consequences contain a greater balance of utility than the balance of disutility contained in the internal consequences, the act is wrong. And, just

as an Iist, he may say this not only about a particular act of killing but also about the class of such acts; acts of this class (i.e., under the description 'killing') will be attended by certain internal consequences the disutility of which will have to be outweighed on any given occasion by a balance of utility in the external consequences, if he is to call them right. Therefore, I do not believe an Iist can account for the wrongness of a type of act, e.g., killing, which the Cist/AUian cannot. Moreover, unless there is on any given occasion a balance of utility in the external consequences, acts of this type might even be said to be 'intrinsically' wrong, meaning that the disutility in the internal consequences is yet to be outweighed. One thing, however, the Cist/AUian cannot say is that, e.g., murder, in advance and quite categorically, is wrong in every instance, which is what the Aist would have him say; he cannot say this, since he cannot be certain in advance that in some future instance a balance of utility in the external consequences will not outweigh the balance of disutility in the internal consequences. However, this, too, is quite consistent with I, particularly the first variant, for neither does Ross think it would always be wrong, e.g., to break a promise or tell a lie, and the same, I should think, would apply to murder.

Even the expression 'right in itself, apart from its consequences' may be accommodated in the present scheme; for if one distinguishes the internal from the external consequences of acts of murder, one may say of this class of acts (under this description) that they are right or wrong, apart from their external consequences. And it is this fact, about external consequences, that Ross usually insists upon. Nevertheless, one must be careful of the phrase 'right in itself', for this can lead one to believe that the act/consequence line is not flexible but rigid or fixed, a view previously rejected.

The second variant of I, according to which rightness is a function of considerations other than consequences, lends itself to the view that the difference between C and I is that the Cist makes rightness a function of something external to the act, whereas the Iist insists on making it a function of something internal to the act. For 'external' and 'internal' here one may also read 'extrinsic to the act' and 'intrinsic to the act'.[22] What is objectionable about this view is that the expressions 'external' and 'internal' to the act suffer from the same disability as 'the nature of the act' and 'the description of the act', viz., that the possibility of providing alternative descriptions of the same act leaves their meaning in doubt. To know what is external

or internal to the act one must know what 'the act' is, but what 'the act' is and what the 'consequences of the act' are depends upon what description is employed. In a word, what is external or internal to the act will depend upon the description under which the set of events in question is being referred to, so that one must speak of what is 'external' or 'internal' to 'the act under a given description', which is, expressed differently, the distinction between external and internal consequences and as such will not mark out a fundamental difference between C and I. Moreover, provided one accepts that different, true descriptions can be given of the same act, it appears that what is internal or intrinsic to 'the act' under one description may be external or extrinsic under another, and this too weighs against any attempt to distinguish I from C along these lines.

In short, on the strength of the above considerations, we may say that I and C are not two fundamentally different and logically distinct ways of assessing the rightness of acts, and that the list position on rightness can be reconciled with C and the expressions so characteristic of I accommodated and given a meaning within C.

On the other hand, nothing so far compromises A; yet I do think its central tenet is at least called in question. For if the considerations of this and the preceding section are correct, then

in trying to decide what to do, one should consider as many alternative descriptions of the act as one thinks likely to be morally relevant. Only in this way will one be able to ascertain both the internal and external consequences of the act under different descriptions; since these consequences will be different with each description, the rightness of the act may be affected accordingly. Therefore, the claim that there are descriptions of acts which, merely by being the descriptions they are, eliminate as morally significant, in trying to decide what to do, any other possible description of those acts _ this claim, though not refuted by our considerations, is rendered suspect. Of course, though I have given reasons for preferring it, an Aist may come to reject the view that different descriptions may be given of the same act and argue instead that different descriptions yield different acts; indeed, I should expect the Aist at this juncture to begin to argue in this way. Nevertheless, it is a striking feature of the work of some Aists that the central tenet of A is held in conjunction with the view that the same act admits of a number of different, true descriptions.[23], and the puzzle is, in the light of the present chapter, how an Aist can ride both horses at once. To this question, I return in Chapter IV, where I consider but reject what I think to be the only way to perform this particular

equestrian feat.

## II.4 Deontology and teleology

The traditional division between deontological and teleological normative ethical theories has often been thought to be of considerable ethical importance; in my view, it is a false dichotomy, at least if it is presumed, as it so often has been, to mark off two distinct conceptions of the role of consequences in determining the rightness of acts.[24] In a word, to the extent that the division rests upon a supposed (radical) difference between I and C, then to that extent it fails to divide.

Of course, the division may be drawn on other lines, e.g., on the status of moral rules[25], on whether the end justifies the means[26], on the dispute between A and C[27], and it may be that irreconciliable differences which neatly fall into two camps may be found in these areas. Even so, it is difficult to avoid the conclusion that the words 'deontological' and 'teleological' are mere labels for a number of problems (and their solutions), with different problems emphasized by different writers; so that, to label a normative ethical theory such as AU 'teleological' does not really distinguish it from deontological theories, without some indication as to which problems (and their solutions) are to be taken as defining the division. And once we look into the defining problems (and solutions), we may find, as with I and C, that the ground of the division is laid in sand.

# III

## ACT DESCRIPTION AND INDIVIDUATION

I want in this chapter to consider several moves which an Iist can make to the Cist's use of the redescriptions of acts; my purpose is not to resuscitate I, but rather to fill out the re-description thesis itself.

### III.1 The nugget theory

In a paper on 'The Ethical Dimensions of the Concept of Action', J. Ladd makes the following observation:

> Ethically it makes a great deal of difference whether a certain segment of the act-consequence series is to be put under the category of action or under the category of consequences.

Ladd is mistaken here: as we have seen, the act/consequence line does not separate morally relevant from morally irrelevant features of the case. But he continues:

> This is because these categories function differently in ethical discourse. Ethical discourse, and moral evaluation in particular, focuses on action, and what it has to say about consequences is relevant only through the concept of action. In other words, moral good and evil are properties of action as such, and hence, whenever we whittle down an action to its bare minimum or, on the other hand expand it to encompass various consequences, we directly affect the moral evaluation of the action thereby. 1

Two features of this passage are noteworthy. First, and rather curiously, Ladd completely fails to distinguish between an action and a description of an action; for example, it is not an action that is compressed or expanded but a description of an action, a

phenomenon that has been dubbed the 'accordion_effect'.[2] Second, the whole passage suggests, indeed, seems dependent upon, the view that there is something, i.e., 'the action', which exists independently of all descriptions of it, and which, if we start compressing and expanding it, changes in moral tone. I shall call this the nugget theory, according to which there are 'acts as such', and all else in the act_consequence series is consequence; against it is pitched the view of the preceding chapter, according to which what counts as 'the act' depends upon what description one selects under which to refer to the events in question. On that view, any contrast between 'the act' and possible, alternative descriptions of it is a false one: it is not as if an act exists independently of all descriptions and that one sets about to find the description that best characterizes it, or is, so to speak, a perfect or near_perfect 'fit'; only the events or pieces of behavior exist independently of the descriptions. This view and the nugget theory, then are opposites.

The advantage for an Iist of embracing the nugget theory is that its adoption will block the Cist's use of the redescription thesis against I; for if there are 'acts as such', then the line between act and consequence is fixed or rigid, and, therefore, not flexible, as is required for the redescription thesis and so for the Cist's reconciliation efforts to get off the ground. The

disadvantage is that he must come up with the nugget, an 'act as such', one that is clearly discriminated and specified; and this is no easy task.

There are two notions with which it may be tempting to identify 'acts as such'. One is Prichard's notion of 'setting oneself' to do something.[3] Here, the nugget theorist, going further and further back in the causal chain, terminates his search in the will, in the 'setting oneself' to do; so if asked for what John did, it is the 'act as such', the 'setting himself' to do, and not the actual doing, that is strictly appropriate. It follows that all ordinary act-descriptions, e.g., 'John killed Jim', will not capture the 'act as such', since they will include a significant chunk of consequences within them. Now this Prichardian view can be countered by a Cist most simply by pointing out that we can 'set ourselves' to do acts under different descriptions, and that therefore the indeterminate boundary of act-descriptions remains to plague the List even if he makes this move. Of course, it is also peculiar to say that what John did was 'set himself' to do; this view, however, is part of a larger theory about the will and willing, one that is pre-Wittgensteinian in character and which has come under serious and sustained attack in the recent past.[4] A consideration of this theory would take us beyond our immediate concerns; suffice it to say, therefore, that an attempt to identify 'acts as such' with Prichard's

concept of 'setting oneself' to do will require a defense of an already suspect theory of the will and willing, a defense, I feel sure, a contemporary List is unlikely to undertake.[5]

A second, and much more likely candidate with which to identify 'acts as such' is Arthur Danto's concept of a 'basic action'.[6] A basic action is one not done by doing something else, and Danto's examples are, typically, movements of limbs, e.g., raising of an arm; here, Danto argues, we have actions the causes of which are not other actions of the same individual. Writing a cheque, on the other hand, is not a basic action because of a different action, viz., signing one's name, or making certain letters and figures with the pen, which stands in the relation of cause to the first action. Put differently, writing a cheque is something one does by means of signing one's name and making certain letters and figures with the pen. Now while Danto's basic actions may be of use in a theory of action, especially for explaining the use of 'cause' in the explanation of human actions, they are of little assistance to an List, and for an obvious reason. Actions like raising an arm or laughing hardly lend themselves to a theory about what makes right acts right, for it is difficult to understand how raising an arm or laughing could be right or wrong in themselves, apart from their consequences. Yet, any further information, any expanding of the description 'raising an arm' to include consequences, and we cease to

have a basic action before us. So an Iist has little use for a nugget of this sort.

In fine, I suspect that all Iist attempts to sustain the nugget theory places the Iist in one of two unenviable positions. Either he is unable to come up with a nugget, a clearly discriminated and specified 'act as such', in which case the nugget theory, and the Iist's opposition to C, collapses _ whatever the reason for his failure to find a nugget, e.g., that it would require the defense of a suspect theory of willing and of acts of will; or else the nugget the Iist finally sifts out, as in the case of a basic action like 'raising an arm', requires the addition of consequences before it becomes of interest to those concerned with what makes right acts right _ in which case it will not be a legitimate instance of an act right or wrong apart from its consequences.

III.2 Sameness of act

A second move the Iist can make against the Cist's use of re_descriptions stems from the Cist's rejection of the nugget theory. For if there is no nugget act, which we redescribe and to which we 'fit' descriptions, then how are we to tell that it is the same act that has been redescribed? In other words, if the nugget theory is rejected in favor of the Cist's redescription thesis, what, then, is the criterion of identity for sameness of act? Failure to provide such a criterion will be held to undermine the redescription thesis

itself, and so to affect the Cist's reconciliation efforts. Once again, a detailed consideration of this move would carry us too far from our immediate concerns; but we can, I think, indicate in general terms the criterion of identity for sameness of act the Cist might provide.

Of (a) 'John wrote a cheque' and (b) 'John scribbled some letters and figures and signed his name', I want to say that these describe the same act, and further, that because the notion of an act seems to involve the notion of the body, of physical or bodily movement, sameness of act is determined by sameness of bodily move_ment. While (a) and (b) are different descriptions of what John did, John's bodily movements are the same. Likewise with Professor Anscombe's case of the man pumping poisonous water into a house[7], what enables us to know that moving his arm, operating the pump, replen_ishing the water supply, and poisoning the inhabitants are different descriptions of the same act is that each description makes reference to the same set of bodily movements. And sameness of bodily move_ments, of course, is more often than not determined by observation, by looking and seeing, sometimes with but more often without the aid of technical instruments. Another way of approaching the problem of the criterion of identity for sameness of act is this: the redescript_ion thesis is not about providing different descriptions of the same act but about providing different, true descriptions of the same act.

Thus, of (c) 'John turned on the light' and (d) 'John alerted the prowler', what we need to ask is not, as Michael Cohen suggests in his paper on 'The Same Action'[8], 'What does it mean to say that both descriptions describe the same action?', but rather, 'What does it mean to say that both descriptions are true?' This cannot mean 'Do both descriptions hold true of the act in question?', since what we are trying to find out is whether there is one, or two, acts; what it does mean, I think, is 'Do both descriptions hold true of the same set of physical or bodily movements?'. In the case of John, or Professor Anscombe's poisoner, it is only because we know that the different descriptions do hold true of the respective sets of bodily movements that we speak of the descriptions as true; how we know they hold true of the respective sets of bodily movements is for the most part through observation, sometimes with the aid of instruments (in order to detect the poison?). It is tempting to say, then, that it is by physical or bodily movements that we individuate human acts[9]; so that, if one of three descriptions does not involve essentially a reference to this set of bodily movements, it is not a description of the same act as the other two.

Of course, I have so far confined myself to descriptions and re-descriptions of overt behavior, and it may be asked whether this begs any questions. I do not think so, but I am not at all sure that I can establish that it does not; the following considerations do,

however, support my view. First, we are almost always in moral philosophy concerned with overt behavior, and in discussions between the List and Cist of what makes right acts right, the acts referred to are almost always pieces of overt behavior. Second, there is a problem about establishing the truth of any purported redescription in the case of what might be pieces of covert behavior, e.g., thinking, intending, etc. As we have seen, with a description and a purported redescription, we shall only know that they describe the same piece of behavior if we know them to be true, but in the case of covert behavior, particularly in the case of first-person psychological statements, we have only the agent's word that the two describe the same piece of behavior. Indeed, we do not even know of the original description, e.g., 'I was intending to go home', whether it is true or not, since we do not know that this was what the agent was intending; and a purported redescription of this piece of intending will suffer from the same deficiency. Moreover, we cannot check the agent's claim that two descriptions of the first-person psychological type describe the same piece of intending, and unless the agent tells someone, others will never know whether either of the two descriptions is true. On the other hand, we can check in the case of overt behavior whether two descriptions describe the same piece of behavior or the same act, since, on my view, we can ascertain whether they refer to the same set of bodily movements. Third, of the

following sets of descriptions,

(1) I intended to kill Scott
(2) I intended to kill the author of Waverley

(3) Oedipus killed a king
(4) Oedipus killed his father
(5) Oedipus killed a haughty stranger,

even if we suppose that both sets are instances of genuine re-descriptions, we must still reckon with the fact that they are not complete but only partial redescriptions, in that the difference among the members of the two sets is the substitution of expressions in the object position that refer to the same individual, i.e., the substitution of referentially identical expressions. (Of course, an individual ignorant of the fact that Scott was the author of Waverley or of the Oedipus story would still not be in a position to judge whether the purported redescriptions describe the same piece of intending or the same piece of killing; so they are only going to be even partial redescriptions to someone with certain bits of knowledge.) In order to obtain a complete redescription, I think a minimal demand is that the verb be altered; what is needed, therefore, is a description, true yet different from the others and with an altered verb. In the Oedipus case, such a description can be given, viz.,

(6) Oedipus caused a drought.

Now the question is whether there can be a complete redescription

with altered verb in the first set of descriptions, a complete redes_ cription in the case of 'I intended to kill Scott'. I am not certain of the answer, but I do think there are in this case two arguments against complete redescription that must be answered. First, we intend acts under certain descriptions, so if the description is different, the intending is different. This argument would carry no force against <u>partial</u> redescriptions in the case of psychological verbs, but with <u>complete</u> redescriptions, e.g., 'I decided to kill Scott', 'I planned to kill Scott', the outcome is not so clear. Second, if 'intend' (or 'think') is an intentional verb, can sameness of intention be retained through alteration of the verb in the purport_ ed complete redescription? In particular, can it be retained through alteration in verb and verb phrase, e.g., from 'I intended to kill Scott' to 'I decided to shoot the author of <u>Waverley</u>'? The point may be put in terms of intentional objects: if 'intend' has an intent_ ional object, will it not follow that if the object is different, the intending is different? If it does follow, then will it not also be the case that, if the <u>description</u> of the object is different, the intending is different? I suspect, without being very sure, that it <u>will</u> be the case, so that a complete redescription is not possible in this, and similar, cases. It is for these three reasons, then, that I do not <u>think</u> any questions are begged in confining ourselves to overt behavior.

Let us turn to the question of whether (6) 'Oedipus caused a drought' describes the same act as, e.g., (3) 'Oedipus killed a king'. It may be thought that, if sameness of act is determined by sameness of bodily movements, then to redescribe in such a ways as to change the verb from one with an immediate reference to a set of bodily movements, e.g., 'killed', to one which appears to lack such immediate reference, e.g., 'caused', makes it impossible to tell whether there is one, or two, acts involved. However, this is not the case, as may be seen if we raise the question 'By doing what did Oedipus cause a drought?' The answer is, by killing a king (his father, a haughty stranger), and this returns the verb to one with an immediate reference to a set of bodily movements. Again, if we ask, 'By doing what did John cause Jim pain?', the answer will be, by hitting Jim in the face, and we have reference again to a set of bodily movements. Accordingly, I think (6) does describe the same act as (3). In principle, I can see no reason why the question 'By doing what did he...?' cannot be asked of any proffered redescription (ones that are different yet true) in the Oedipus or any other case involving overt behavior.

However, will not the asking of this question lead to difficulties? For our opponent will now insist that we carry on asking the question, e.g., 'By doing what did Oedipus cause a drought?', 'By doing what did he kill a king?', 'By doing what did he plunge a knife into the king's

breast?', and so on, further and further back until we come to an act not done by doing something else, i.e., to something like a basic action. My objection to this enterprise is that to go very far back in the causal chain - back beyond observable, discriminable bodily movements - is to go beyond what we would say was an _act of Oedipus'_. For though one may go on to say something about the altering of certain nerve fibres in Oedipus' arm or even about certain neuron firings in his brain, these are not really things that Oedipus _does_ in the pertinent sense. The altering of the position and tension of the nerve fibres in Oedipus' arm is something he undergoes, not something he does; and while it may be difficult to delineate precisely what this sense of 'do' is[10], I doubt whether neuron firings and altering nerve fibres are hard cases. Nor, _contra_ D'Arcy's _Human Acts_[11], do I think that the question 'What is he doing now?' picks out any identifiable group of things one may be said to be doing in the relevant sense of 'do'. For example, 'frothing at the mouth' and 'having a fit or seizure' are just as natural responses to this question as are 'walking to work' and 'cutting the grass'; yet, as with 'altering his nerve fibres', I should argue, first, that these are not things one _does_ but rather things one undergoes, and second, that at the very least they are not things one _does_ in the typical sense in which walking to work or cutting the grass are. Moreover, I do not think one can distinguish frothing at the mouth or

having a fit or sweating from altering one's nerve fibres by claiming that the last is voluntary while the others are involuntary: many film actors are masters of the 'cold sweat' and feigning seizures, and epileptics themselves can induce fits. Of course, a special story might be told so that we would accept altering his nerve fibres as an act of Oedipus', but such special circumstances are required for us to do so. So I do not think we leave ourselves open to the difficulty envisaged, if we employ the question 'By doing what did he...?' in the Oedipus case.

One final point. If (a) 'He wrote a cheque' and (b) 'He signed his name' describe the same act, and it is true that (a) entails but is not entailed by (b), then one may be tempted to conclude that the act under description (b) is more basic than under description (a).[12] The Cist is not troubled by this conclusion, since the view that (a) and (b) do describe the same act allows his reconciliation efforts to get under way; and the additional fact that he takes into consideration a number of different descriptions of the act, in order to assess its internal and external consequences under different descriptions, and not merely what is on the present view the more basic description, renders this version of the basic/non-basic distinction harmless.

III.3 Abstract act-descriptions

We are now in a position to consider a third, and quite ingenious, move of

the List against the Cist's use of the redescription thesis. In a paper entitled 'Consequences', James Griffin propounds the following argument against the Cist:

> His position has these two features. First it requires supplementation with a standard of intrinsic goodness or badness. Second, it makes this the fundamental moral standard; that is, all judgments that acts are right or wrong derive from judgments that situations are good or bad. Thus, C says that from certain judgments about situations being good or bad we may derive judgments about acts' being right or wrong. But then from his own judgment that certain situations are intrinsically good or bad we can derive the judgment that certain acts, namely, the acts of producing these situations, are intrinsically right or intrinsically wrong. If, for example, C chooses pleasure and pain as the only things good or bad in themselves, then he should also accept that giving someone only pleasure is right in itself, and hurting someone unnecessarily is wrong in itself. Certainly, C cannot object to speaking of the 'act' of producing happiness; it may be an abstract act description, but it is not improper for that reason. 13

The List, then, employs the redescription thesis against the Cist by redescribing the same act in such a way, e.g., as 'an act of producing pain', as to make it right or wrong apart from its consequences.

Now the important thing to be said here is that the reason an act of producing pain is wrong apart from its consequences is because the consequences have been included in its description; that is to say, if this new move brings victory to the List, then it is surely a Pyrrhic victory, since it is achieved only by incorporting the consequences

into the description of the act. This can be established as follows. Of an act described as 'an act of producing pain', we may ask, once again, 'By doing what did John produce pain?', and the answer will be in terms of hitting Jim, yelling at him, etc. This use of the question 'By doing what did he...?' shows that what Griffin is calling an abstract act_description is a stand_in for other act_descriptions, at least one of which can always be given, viz., that one which says through what or by doing what pain was produced. If we consider <u>this</u> act_description, one which says by doing what pain was produced, e.g., 'John hit Jim', then causing Jim pain will figure in the external consequences of the act under that description, and Griffin's ploy will be seen to amount to incorporating these external consequences into a subsequent redescription, i.e., 'an act of producing pain', of the same act. In fine, in response to the charge that an act of producing pain is wrong apart from its conse_quences, the <u>Cist</u> will point to the fact that 'an act of producing pain' is a stand_in for at least one other act_description, one which says by doing what pain was produced, and that of this act_description, pain figures in its external consequences. So, if there is a victory for the <u>List</u> here, it is a Pyrrhic victory: for it is only in a strained sense that 'an act of producing pain' is wrong <u>apart</u> from its consequences.

---

In summary, then, the foregoing attacks by the Iist on the redescription thesis may be withstood by the Cist; and in the absence of a reason for thinking otherwise, we may conclude that the reconciliation of I with C remains in force.

# IV

## ABSOLUTISM AND HARD CASES

Since the means of reconciliation between I and C casts doubt upon but does not refute the central thesis of A, which is, loosely, that there are some acts (or kinds of acts) it would always be wrong to do, whatever the consequences, the dispute between A and C persists. Moreover, its persistence would seem inevitable, for the question of whether some acts are proscribed absolutely appears to make sense only against a backcloth of acts wrong apart from their internal and/or external consequences, and the Aist and Cist are completely at odds over whether there are any such acts.[1] On another line, however, the Cist may be able to tighten his case against A. Accordingly, in this chapter I consider several hard cases and one or two ways in which an Aist may hope to deal with them; then, in the subsequent chapter, I examine the doctrine of double effect, a principle of conduct to which, in certain of these hard cases, an Aist may resort.

### IV.1 Hard cases

If an Aist identifies more than one kind of act as proscribed absolutely, it is possible that conflicts will arise; and there appears to be no happy solution for an Aist to conflicts among his absolute rules, duties or proscriptions. For example, he cannot weaken an absolute rule without its ceasing to be absolute, and the same is true of absolute duties or proscriptions. Nor is

there any hierarchy of absolute proscriptions, with some more absolute than others: sodomy is proscribed as absolutely as murder, so that in any conflict the latter does not take precedence in degree of proscriptiveness over the former. Moreover, though the problem of conflicts may be got round by holding a theory of absolute rules or proscriptions in which there is but one such rule or proscription, Aists invariably identify more than one kind of act that may never be done. Thus, Professor Anscombe, as we saw in Chapter II, includes within this category killing the innocent for any purpose, treachery, idolatry, sodomy and adultery; and in his book Medico-Moral Problems, Gerald Kelly, S.J., adds to this list:

> There are some actions, as we know, that are intrinsically evil: blasphemy, perjury, masturbation, and murder. Such actions are always forbidden, no matter what beneficial consequences they might have. 2

(The fact that this list appears very strange, with masturbation grouped together with murder and perjury, I shall overlook here.) In sum, with no obvious solution to hand, the problem is one of how the Aist proposes to deal with conflicts between absolute pro-scriptions.

## IV.2 Distinctions

There are subtle ways in which the Aist may try to evade having to address himself to the problem of conflicts, and these we may hope

to deal with by drawing certain distinctions.

At one point in her paper on 'Modern Moral Philosophy', Professor Anscombe writes:

> But if someone really thinks, in advance, that it is open to question whether such an action as procuring the judicial execution of the innocent should be quite excluded from consideration _ I do not want to argue with him; he shows a corrupt mind. 3.

The first distinction to be observed is that between contemplating judicial murder in an imaginative way, for the purposes of philosophy (e.g., in order to comment on the adequacy of certain normative views about the rightness or wrongness of a particular kind of act), and contemplating it in an actual case. Prima facie, there is no reason to think the philosopher shows a corrupt mind, nor is there any reason why he should not contemplate what he does; indeed, because normative ethical theories are sometimes tested against hypothetical as well as actual examples, the skilful use of imaginative cases can be a helpful tool in evaluating those views. On the other hand, for a judge in an actual case to contemplate judicial murder would certainly be unwise, not because a situation in which it would be right to commit judicial murder could never arise, but for the rather obvious reason that, having to decide amidst intense pressures (e.g., mob action, and demands to find those responsible for a particularly heinous crime) and having no

clear view of the consequences of his decision, and with every temptation to engage in special_pleading, he is apt to convince himself that this is one of those rare cases in which judicial murder is right. And if this is true in the judge's case, how much more will it be so for those even less equipped to consider the situation intelligently: here, it is a positive hazard to consider performing such an action, and there can be little doubt that it is for the best in nearly all actual cases that one dismiss completely from one's mind any thought that this might be one of those exceptional cases. This line is perfectly consistent with C. (Indeed, I make a good deal of it later, and on this score the tempered act_utilitarian goes some way towards accommodating the Aist; the reader may go forward to Chapter VII.1 for the first stages of this subsequent discussion.) The differences between the cases of the philosopher and the judge lie in the circumstances or situation in which they contemplate judicial murder; whereas Professor Anscombe seems of the opinion that both exhibit corrupt minds, and ought not to contemplate what they do.

To the above passage, Professor Anscombe appends a footnote:

> In discussion when this paper was read, as was perhaps to be expected, this case was produced: a government is required to have an innocent man tried, sentenced and executed under threat of a 'hydrogen bomb war'. It would seem strange to have

> much hope of averting a war threatened by such men as made this threat. But the most important thing about the way in which cases like this one are in_ vented in discussion, is the assumption that only two courses are open: here, compliance and open defiance. No one can say in advance of such a situation what the possibilities are going to be _ e.g., that there is none of stalling by a feigned willingness to comply, accompanied by a skilfully arranged 'escape' of the victim. 4

(It is worth pointing out, that if a feigned willingness to comply is successful and the victim does escape, it may still be true that in adopting this course the Aist will have been forced to break other rules that are absolute or to perform other acts that are pro_ scribed absolutely, e.g., lying.) A second distinction to be observed is that between a logically possible case and one that is never likely to be met with in our ordinary experience. Unless the Aist is prepared to say that a situation wherein it would be right to procure the judicial execution of an innocent man could never arise, it will remain possible, with careful attention to detail, for the philosopher to spell out a case in which it would be right; and if the Aist does argue that such a case is logically impossible, then apart from trying to unravel the puzzle of why this is so, we may observe that it is he who runs afoul of Professor Anscombe's remark that the possibilities cannot be told in advance.[5] A much more attractive position, and one I espouse, is that in our every_

day lives we are never likely to come across a case in which judicial murder is right, so that, with a view towards adopting rules, etc., by which to act, we should pay attention only to likely cases, to those that our rules will be called upon to deal with. This view, however, is perfectly compatible with C: the fact that we are never likely to come across a case in which judicial murder is right does not exclude the possibility that such a case could arise, and that, if it did, it would be in virtue of its consequences that it would be right or wrong.

In the above passage from Professor Anscombe, the case produced for consideration was that of a government 'required to have an innocent man tried, sentenced and executed under threat of a "hydrogen bomb war"'. A third distinction to be observed is that between the skilful and helpful use of imaginative examples and the production of bizarre or fantastic ones. Here, we may tie together our preceding remarks (though the reader is referred to Chapter VII for the use to which they will be put). The skilful use of imaginative cases can be an aid to theory-constructing and -testing, and I think that how a theory deals with such cases is in part (but only part) a measure of its acceptability; on the level of theory, then, I should wish to press the judicial murder case and conflict situations against the Aist. On the level of practice, however, we are not doing philo-sophy, but rather are attempting to find the best way, the best

guides, for doing the right thing; and reliance upon imaginative cases, unless they closely resemble actual ones, ones that we are likely to meet with and with which our guides will have to deal, may lead us to select guides that let us down at crucial points, e.g., by equipping us for cases we shall not meet, and not equipping us for cases that we shall. Clearly, reliance upon bizarre or fantastic cases will be even more likely to equip us inadequately; the situations to be met with by ordinary people are already complex enough without adding bizarre features to them. What I am suggesting, in short, is that there is a place for theory_constructing and so for imaginative examples, a place for the judicial execution and hydrogen bomb case and all the other cases of this sort that fill a philosophy seminar; I do not, therefore, think that desert_island promises, death_bed promises, etc., have no place in the critique of classical utilitarianism, any more than I think that the judicial murder example and conflict situations are out of place in a critique of A.[6] A theorist tries to bring as many and as varied cases as he can under his theory, and tries also to find out why some will not fit or fit less well than others. But, to repeat, when it comes to selecting those guides by which to act, it is important that we consider likely cases; since we are never likely to come across a case in which judicial murder is right, and since we are always sus_ceptible to special_pleading, it is for the best on this practical

level that we dismiss such a case from our minds completely. Thus, to gloss over the real difference between the philosopher's use of the judicial murder case and the contemplation of judicial murder by a judge in an actual situation, as the Aist does, leads to the objectionable view that, if a case can be described as 'procuring the judicial execution of the innocent', then he who contemplates it exhibits a corrupt mind, whatever the circumstances or situation in which he does so, whatever his reason or aim in doing so, and whatever other descriptions the events in question may bear. So, the Cist does teach a philosophy that leaves open to question the prohibition on the judicial execution of the innocent, in the sense that he explicitly allows the possibility that a case could arise in which he would, on the strength of consequences, judge judicial murder right; but he neither urges us to conduct our lives along lines that raise in each and every case, the prospect that this is one of those rare cases in which it is right - indeed, he can consistently urge just the opposite; nor does he advocate that we consider such cases in select-ing those rules, principles, etc., by which we act - indeed, once again, he can consistently advocate precisely the opposite.

One final, though related, point. For the Aist, however many details about circumstances or consequences are forthcoming, and however many other possible descriptions may be produced, if the proposed act can be described as 'procuring the judicial execution of

the innocent', then it is proscribed absolutely. Embedded here is a way in which an Aist may hope to embrace the redescription thesis and yet to retain the central thesis of A, viz., that there are descriptions of some acts which, merely by being the descriptions they are, eliminate as morally significant, in trying to decide what to do, any other possible descriptions of those acts. In brief, the Aist will agree that the same act can be described in different ways, but will go on to urge that we look at our acts under their worst possible descriptions, not under their best; and that the act under every description must be permissible. This will be so, moreover, whatever the example (or act). The difficulties with this suggestion are, first, that we are given no explanation of why we should always choose to look at our proposed acts only under their worst descriptions and never under their best, and second, that we are given no criteria for choosing 'worst' descriptions. I cannot select 'the ugliest philosopher in England' without some indication of what I am to look for; the same is true of the 'worst description of my proposed act'. An interesting problem crops up here, for the most obvious Aist reply as to how we pick the 'worst' descriptions of our acts lands him in difficulties. That is, he may argue that, prior to the consideration of other possible act-descriptions, he has a list of kinds of acts that are absolutely forbidden, and when con-fronted with a case, he runs over the possible descriptions of the

act he thinks likely to be morally relevant: if one of these descriptions matches one of those on the list of forbidden acts, then the act, however else it may be described or whatever its beneficial consequences, is impermissible. 'Worst' description, then, will be defined extensionally, by the list of forbidden acts. Now the difficulty Jonathan Bennett finds with this reply is that the Aist can 'safely be assumed to have delivered himself over to a moral authority and to have opted out of moral thinking altogether'[7]; but this is a bit too general to pinpoint the main trouble. The important criticism of this Aist reply is not that a list of immutable prohibitions serve the Aist as a moral authority, but rather that the uncomplicated act-descriptions that fill this list, e.g., 'lying', murdering', etc., make no allowance whatever for the full circumstances in which such acts are contemplated, for the degree of complexity of situation to be commonly met with.[8] Frequently, merely to describe a case as one of lying is not to say all that there is to say about it, since such a description leaves out any mention of further circumstances or additional features of the situation, e.g., that the lie is contemplated to save a dying patient further mental anguish; and it is precisely because the Cist can allow circumstances or situation into his theory, by allowing them to figure in different redescriptions of an act and so to be taken into account (since the internal and external consequences change and these changes can be

assessed in the determination of rightness), that the theory may be of important assistance in helping us to decide what to do in often complicated situations.[9]

Now to leave out any mention that a lie is contemplated in order to save a dying patient further mental anguish, and so on, masks the real differences among cases. But how can an Aist include these complicating factors, or fuller specifications of circumstances, in his theory? If he both allows these sorts of things to figure in different descriptions of the same act (e.g., 'lying', 'lying to save a dying patient mental anguish') and for these descriptions to be morally relevant, he violates the central thesis of A, since he will have allowed further descriptions of acts on his list of forbidden acts to be morally relevant to deciding what to do; if, on the other hand, he allows such factors to figure in different descriptions of the same act, but excludes these descriptions as morally irrelevant, in the case of acts on the list of forbidden acts, then he excludes as morally irrelevant (part of) the circumstances in which acts are contemplated and performed. If the list of forbidden acts simply cannot be made to take account of the situations in which we actually find ourselves, then apart from any glow of unworldliness it may take on, its practical usefulness must surely be diminished. I regard it as a virtue of C that fuller specifications of circumstances, through inclusion in redescriptions of the same act, with resultant changes

in the internal and external consequences of the act under different descriptions, are able to be taken into account in deciding what to do.[10]

This is a significant point in C's favor, for, prima facie, one might well think that the Cist errs in placing so much weight upon an act's consequences, unless he makes some provision for including the circumstances of the act, in the determination of its rightness. The Cist might hope to argue that, in those cases where circumstances might seem to make a morally relevant difference to the rightness of an act, this difference will reappear in the consequences; but this is too simple: it seems to commit the Cist to the view that any morally relevant change in an act's circumstances will result in, bring about or cause a change in the act's consequences, without there being any ground for believing this view correct. The more plausible line here is the one I have taken, viz., that the Cist allows circumstances to be included in the redescriptions of an act, which in turn means that internal and/or external consequences of the act change; these changes in consequences may then be taken into account for the purpose of assessing rightness. In other words, which consequences are internal and which external to an act depends, as we have seen, upon the des_ cription of the act; change this act_description and these consequences change. On reason for expanding or changing an act_description is to include more of the circumstances surrounding what happened; once in_

cluded, the act-description is different, and so then are the internal and/or external consequences. Nothing similar is possible in the case of A: lying is forbidden, whatever the consequences, whatever the circumstances. To repeat, then, there is no way to take account of the complexity of the situation in which one finds oneself on such a view, and this must surely diminish its practical usefulness.

## IV.3 Breaking down and harm befalling others

Having clarified the distinctions in terms of which examples may be produced against the claims of the Aist, and the level on which examples are pressed against the Aist, we may turn to another Aist reply to conflicts. In a recent paper 'On Consequentialism', J. Cargile provides us with the tools out of which we may hope to fashion an Aist solution to them.[11]

Cargile distinguishes two types of A:

> ...besides the absolutism which holds that some kinds of action are always wrong, we have the absolutism which says that some kinds of actions could not be performed by a good man. Such actions would be physically within the power of a human body. But the agent who could bring himself to perform them could not be good. 12

Let us call the former type of absolutism A1 and Cargile's new type A2. According to Cargile, A2 is a stronger opponent of C than is A1, and there are considerations which may be thought to support this view. For example, it may be that a number of peasants will not be needlessly shot if the hamlet chief, who has cooperated in a government pacification

program, is surrendered to the revolutionary cadres; and one of the peasants may reason that it is better that one should be murdered than many, with the result that he may come to think that the right thing to do here would be to hand over the hamlet chief to his enemies. But what, then, of the moral worth of the peasant who could bring himself to do this? Here, our peasant may reason that a good man could not betray his friend and confidant, the hamlet chief; and he may go on to think that it would be a portent of character if one could bring oneself to act in this way.

However, the support this sort of story gives Cargile's view is not at the expense of C; that is, it does not support the view that A2 opposes C. Three observations are relevant in this respect. First, A2 is not A1; for A2 does not claim that there are descriptions of acts which eliminate as morally significant any other possible descriptions of those acts. Moreover, there are no absolute proscriptions, no immutable prohibitions, no list of forbidden acts, no acts it would always be wrong to perform. What in A1 has produced clashes with C is dropped in A2. Second, the contention that there are some acts which no good man could bring himself to perform need not be at odds with C. For there is no reason why on A2 acts may not be judged right or wrong solely in virtue of their consequences, as, for example, in the case of the peasant. C is a view about what makes right acts right, not about what makes good men good; and so long as A2 allows

rightness to be consequentially determined, there is no reason for it to be at odds with C. Third, as for goodness and badness, if our Cist is also an AUian, he may point to a long-standing distinction within AU theories between the rightness or wrongness of an act and the goodness or badness of an agent, and the considerations relevant to each; for example, while he may consider intention, motive, character, etc., in assessing goodness, he considers consequences in determining rightness. There is no reason why A2 should blur this distinction, nor in the peasant case is it so blurred; and if it is not blurred, there is no reason to think A2 incompatible with it.

However, Cargile's distinction between A1 and A2 does provide us nevertheless with the tools for constructing a possible Aist solution to conflicts. Let us suppose that the Aist is faced with a situation in which he has been told something in confidence by a friend and then in the witness box is asked a question to which he can only give a truthful answer by breaking the confidence: the Aist, if he is a good man, could bring himself to do neither, and he would, so to speak, simply break down. If he goes ahead and does one or the other, he cannot on Cargile's view be a good man, and furthermore, Cargile suggests, it must surely be a sign of his character that he could bring himself so to act. Moreover, it makes no difference on this view whether the good man thinks that, e.g., perjurying himself is right in this case; he still could not bring himself to commit perjury. (In truth, however, Cargile does not cite

perjury as the sort of act a good man could not bring himself to perform; he deals rather with cases of murder, torture and sadism, a point to which I return below.) So, in a conflict situation of this sort, the Aist, the good man, breaks down, and by doing so will be seen to remain firm behind his conception of what a good man can and cannot bring himself to perform. To this proposed solution to conflicts, there are a number of Cist replies that deserve mention.

(1). Cargile assumes that if a situation is serious enough and if a man performs an action which a good man could not bring himself to perform, then that man could not be of good character. What if I betray a friend's confidence, and that in a conflict situation, once in my life? Or steal only once in my life? And that the rest of the time I perform as a good man should? Cargile's estimate of me will be framed around the single case and not on the run of cases, which, of course, given the seriousness of the single case, amounts to the view that a man's 'true character' emerges when 'his back is to the wall'; if I betray a friend's confidence in a serious situation, my character on this view will be tainted forever, on the strength of this single case. While a harsh view, the surprising thing is that it appears to be held simultaneously by Cargile with a dispositional analysis of character; he says, for example, that 'there are actions which someone possessing the dispositions characteristic of a good man could not possibly perform in any circumstances'.[13] If in

serious circumstances, but a conflict situation, I once in my life betray a friend's confidence, it cannot be said that I lack the dis_ positions characteristic of a good man; for a single performance could not alone show whether I lacked, e.g., disposition X. I range of cases in which I have opportunity to X and in which a good man would not X but in which I do X are required to establish this; but in the bulk of cases I do not X; there is only one case in which I do. So a single instance will not show that I lack this disposition. Of course, an Aist may reply that if I could bring myself to perform such an act once, I can bring myself to do so again. However, the burden of proof here is upon the Aist, for the very good reason that this reply ignores (once again) the circumstances or situation in which I was able to bring myself to betray a confidence, viz., in a situat_ ion of conflict between betraying a confidence and perjuring myself. Then, too, I have to date, in normal circumstances, performed as a good man should, and it is only in extreme circumstances that I have been able to bring myself to betray a confidence: why should one go on to suppose that in subsequent cases in normal circumstances I am now more likely to break my established, good habits?

(2). What sorts of acts may we cite as those which a good man cannot bring himself to perform? Cargile instances murder, torture and sadism, and to be sure these are some of the kinds of acts supporters of A1 have said may never be done. Of the other kinds,

e.g., idolatry, blasphemy, perjury, lying, masturbation, etc., it is not clear whether Cargile thinks a good man could bring himself to do one of these. This raises the question of how the good man finds out what he can and cannot bring himself to do. If he has, in advance, a list of acts and consults this list to find out what he cannot do, then the arguments of the previous section are applicable. Whereas, if he says that he cannot tell in advance of an actual situation just what he can and cannot bring himself to do and that the way he finds out is by examining the alternatives in each case, then we are entitled to complain that we have no idea how this examination is to be conducted. The difficulty finds an analogue in Ross' solution to the problem of conflicts among his prima facie duties: we simply examine each of the conflicting duties, weigh them up, and decide which, in the circumstances, is the more stringent. However, to make this decision, we need principles or a measure by which to determine stringency. Likewise we need principles by which to conduct the examination of alternatives, to see whether a good man could bring himself to perform one of them. For example, just as in the case of Ross, it may be suggested that it is by consulting our feelings that we determine which actions a good man could or could not bring himself to perform - assuming that we are good men.[14] However, the problem with this suggestion, if we set aside any question of whose feelings we consult (our own, no matter who we are?), or

which feelings we consult, or what we do when conflicts arise amongst people about how they 'feel' over this or that case, or what I do when I feel one way and then another about the same case _ the problem is how we consult our feelings in the required manner. Is an action a good man cannot perform one that engenders an intense feeling of repugnance? An intense feeling of repugnance over a certain duration? Some sort of measure or calculus seems required for consulting our feelings.

(3). By breaking down, the Aist shows that he is a good man and that he will have no truck with intrinsically wrong acts or, in any event, acts of a particular kind. However, the result of his having no truck with them, whether or not accompanied by a break_down, is that the Aist allows great harm to befall others for the sake of his own goodness. This may be seen by the following familiar example: let us suppose that an Aist obstetrician is confronted with a situation _ not altogether uncommon in the case of ectopic pregnancies _ in which, if the foetus is not removed, both mother and foetus will probably die, whereas if the foetus is removed, the mother will very probably live. An Aist is likely to see this situation as one in which an intentional killing of an innocent human being stands against allowing two human beings to die: since the former constit_ utes for him an intrinsically wrong act or an act he as a good man could not bring himself to perform, it is ruled out, with the result

that mother and foetus are allowed to die. A non-Aist, who certainly need not be a Cist, is apt to see the situation as one in which, for the sake of one man's goodness (the obstetrician's), great harm is allowed to befall others, and remarks such as 'it is preferable by far that a million mothers and foetuses perish than that a physician stain his soul with murder'[15] are apt to confirm him in this opinion. In a word, the avoidance of intrinsically wrong acts or acts a good man could not bring himself to perform and the avoidance of harm to human beings do not always coincide; and when they do not, a non-Aist may be excused for thinking that the Aist will often increase the amount of harm to human beings.

(4). Another, and much more important, point about the ectopic pregnancy case is that the Aist obstetrician must be prepared, if he is to decide consistently on the basis of A, to permit harm to befall others which would not have befallen them if he had acted differently. The harm permitted will include the ultimate harm, loss of life: our peasant, instead of breaking down, could remain silent and not betray the hamlet chief, but then he must be prepared to see many others die. By deciding to remain silent, he preserves his own goodness at the expense of the lives of others; to hand over the hamlet chief, he sullies his own goodness in order to preserve the lives of others. (This sort of decision is not of the kind that confronted Jewish doctors in the camps, who alleged that they had to choose between

performing experimental operations and being shot; for refusal to perform such operations prima facie brings harm only upon oneself.) Now most importantly, the Aist obstetrician has control over the situation he is in, in the sense that the harm that befalls mother and foetus comes about because he decided one way rather than another. We have not hereby said that he is to be blamed for what happened; but we are saying that there is a case to be answered; the fact that he had control over the situation makes him answerable for the out_come. (I return to this theme in my discussion of the doctrine of double effect.) What a man is prepared to allow or see happen is as much a part of his moral make_up as what he does, and in any assess_ment of character, is equally entitled to be included. A good man, I want to say, has not only to ask himself what he can and cannot bring himself to do, but also to ask himself what he can and cannot bring himself to allow, permit or see happen.

(5). Finally, breaking down is not a position that everyone will be content to have characterize the good man: I think someone might well argue that, instead of breaking down in the face of a conflict between, e.g., betraying a confidence and perjurying oneself, the good man is precisely the individual who should not break down in this situation but who should be equipped to think out moral problems like this one, the more so since the well_being of others is at stake.

## IV.4 A verbal ploy

Rather than tell a lie, betray a confidence, etc., the Aist may elect to remain silent. Of course, the Aist's opponent will now seek out cases in which remaining silent constitutes performing what the Aist regards as an intrinsically wrong act or one a good man could not bring himself to perform, and will press these cases against the Aist. But I think the strongest reply to the Aist on this score is the one given in the case of the peasant and the one I have been emphasizing, viz., that in electing to remain silent, the peasant is prepared to allow the other peasants to be murdered. If one is asked a question in the witness box to which one can only give a truthful answer only by betraying a confidence, and one is an important, material witness to the innocence of the accused, then in electing to remain silent one must be prepared to allow a miscarriage of justice. To all but the most hardened of Aists, what is permitted to befall others through one's remaining silent renders that course in the present cases unacceptable; for one to go on, then, and elect to remain silent is as much a sign of character as betraying the confidence.

Of course, the Aist may elect not to remain silent but to perform a kind of verbal conjuring trick. For example, in his note on 'Lying and Equivocation', J. H. Newman argues that, while we may be under a strict duty not to lie, we are under no such strict duty, under certain circumstances, to tell the truth; and he envisages in certain hard

cases the use of equivocation and other verbal devices in order to avoid performing what he considers to be intrinsically wrong acts.[16] Our peasant may resort to equivocation and hope for the best, though this device more likely than not will only delay the moment of reckoning. But in the case of conflict between betraying a friend's confidence and perjurying himself, where he is the only witness to the innocence of the accused, I think equivocation is ruled out for the Aist. The Aist's taking of the oath and solemnly swearing to tell the truth places him under a strict duty to tell the truth and so to forego equivocation and similar devices. Clearly, in such cases, the Aist would do better, if he could foresee the dilemmas into which he might be placed, not to take the oath; but refusal to take the oath in a case in which he is the only material witness to the innocence of the accused may for a variety of reasons lead to a miscarriage of justice. In essence, then, the case before us is one in which the Aist cannot betray the confidence, cannot perjure himself, cannot resort to equivocation and other such devices, and can only remain silent if he is prepared to allow a miscarriage of justice. It makes no difference for the purposes of assessing character whether justice here actually miscarries: whether it does or not, the Aist is prepared to see it miscarry.

If it is not clear that the good man could, for the sake of his own goodness, bring himself to allow great harm to befall many others,

neither is it clear that the good man could bring himself to equivocate. The reason for this is that the distinction between lying and equivocation can appear a fine one; that is, the difference between the two can appear to consist in the fact that lying involves intentionally trying to deceive the other party whereas equivocating involves intentionally trying to mislead the other party. Whether Cargile's good man will find this difference so significant that he can bring himself to equivocate but not to lie must be uncertain.

---

It is clear, then, that those who remain firm in their attachment to A are going to be confronted with difficult cases. However, in certain of these cases, there is another move the Aist can make to evade their thrust, a move that has been the subject of a good deal of controversy amongst philosophers and moral theologians: namely, in a case in which he would otherwise be debarred from acting, the Aist may resort to the doctrine of double effect to justify bringing about a particular effect. I turn now to a consideration of this doctrine.

# V

## DOCTRINE OF DOUBLE EFFECT

The issues involved in the Aist reliance upon the doctrine of double effect (hereafter abbreviated to DDE) are important enough to warrant separate treatment. I shall be concerned here to clarify and to comment upon certain of these issues, particularly those that may be thought to assist the Aist in escaping from several difficult cases the non-Aist (who may or may not be a Cist) brings against him; and by developing the notions of an agent's control of a situation and, therefore, of his responsibility for its outcome, I hope to show that DDE will not provide the Aist with the escape he desires.

### V.1 Conditions for its use

DDE is a principle of conduct - one, it should be added, which can appeal to both Catholics and non-Catholics alike[1] - that allows us in certain hard cases, provided certain conditions are satisfied, to bring about an effect indirectly which it would be forbidden us to bring about directly. The distinction between the directly intended and the merely foreseen consequences of an act, which we touched upon Chapter II, may be redrawn in the Benthamite terms of direct and oblique intention[2]; we may say, therefore, that DDE allows the Aist to bring about obliquely, as a foreseen consequence of his action, what would be forbidden him to intend directly. The following two examples may be said to illustrate this difference. First, a woman in labor will die unless a craniotomy is performed on the unborn foetus, while if the operation is not performed, the child can be

delivered by Caesarian section after the mother's death.[3] Second, a doctor relieves a patient's pain by administering a certain dosage of drug; since the illness is long standing, an ordinary dose will no longer suffice to deaden the pain; therefore, the doctor gives the minimum necessary to relieve the pain, but knowing as he does so that this minimum dose will in fact prove fatal or at least hasten death. To the Aist who accepts DDE these cases are fundamentally different. In the former, craniotomy is impermissible and may never be performed, and this remains true whether the mother's life alone is threatened or whether the lives of both mother and foetus are at stake.[4] To operate on the foetus in this way is a case of 'direct' killing, in that the doctor directly attacks the foetus with the intention of killing it, whereas by not operating and delivering the child after the mother's death, the doctor only foresees or obliquely intends the death of the mother. (By not operating, the mother is said by adherents to DDE to be 'permitted to die'. The same is true in the case where the lives of mother and foetus are threatened: any action which con_ stitutes a direct attack on the life of the foetus, or, for that matter, upon the life of the mother, is proscribed absolutely; and both mother and foetus are said to be 'permitted to die'. The use of 'permits' in these contexts appears standard practice amongst Catholic thinkers[5]; therefore, the reader must be careful not to confuse, because of the expression 'permitted to die', the act_

omission distinction with DDE, which applies equally to acts and omissions.) DDE is invoked, then, to show or explain why the doctor may not perform the craniotomy. In the drug case, on the other hand, DDE is invoked to explain why it is permissible for the doctor to administer a dosage of a pain-killing drug that will prove fatal or hasten death. Here, the doctor directly intends the relief of pain, not the patient's death, and only foresees as a by-product of his act that death will in fact result or be hastened. The employment of DDE to distinguish morally between these cases, i.e., between a case in which DDE does not allow the Aist to act and a case in which it does, rests in the first instance upon the fact that certain conditions must be satisfied: in the craniotomy case, he reasons that these conditions are not satisfied, whereas in the drug case he reasons that they are.

There are four conditions the satisfaction of which is necessary in order for DDE to allow the Aist to act. (1). The act to be done, considered in itself and apart from its internal and external consequences, must not be intrinsically evil, intrinsically wrong. This condition is incompatible with C; that is, because DDE (just as A or A1 itself) only makes sense against a backcloth of acts which are wrong independently of their internal and external consequences, it follows a fortiori that no Cist can accept DDE. Put differently, this first condition assumes that the line between act and consequence

is both fixed and moral in kind; a Cist, on the other hand, thinks that the act/consequence line can be drawn in different places in the causal chain, and that, wherever it is drawn, the line does not separate morally significant from morally insignificant features of a case. I should perhaps emphasize here that, since RUians are not Cists, they may be able to reconcile themselves to DDE; this possibility is denied to any and all AUians. (2). Given both a good and an evil effect of an act, the evil effect must not be intended, but simply tolerated or permitted. Here, we encompass the distinction between direct and oblique intention: the Aist aims at the good effect and merely foresees the evil effect as a by-product of his act. (3). The evil effect must not be a means of bringing about the good effect. This condition is a restatement of the traditional Aist belief that moral evil may never be done that good may come. Three points warrant mention here. First, if the good effect comes about through or as a result of the evil effect, then the evil effect as means is willed to achieve the good effect as end; in this event, the agent would directly intend the evil effect. Second, this point about means and end is sometimes put this way, that the good effect must be produced directly or immediately by the act and not by the evil effect.[6] What 'directly' and 'immediately' imply in this context is a matter to which I return below. Third, it can be misleading to express this third condition by saying that the good effect 'must come before' the

evil effect; for what is important is causal and not temporal sequence. The act and not the evil effect must be the cause of the good effect; otherwise, moral evil will be done. (4). The good to be obtained must be desirable enough to offset permitting the evil effect. Essentially, this condition amounts to weighing up what is to be gained against what is to be permitted, and the former must compensate for the latter. These, then, are the four conditions the fulfillment of which is necessary if DDE is to allow the Aist to act. All four must be satisfied: failure to satisfy one or more of them, as in the craniotomy case, where conditions two and three are violated, ensures the inapplicability of DDE to a situation or case.

I shall in what follows be concerned exclusively with conditions two and three. I have in earlier chapters already dealt with condition one, whereas condition four lacks both the importance and interest of two and three.

## V.2 Criticisms, cautions and clarifications

The employment of DDE to distinguish morally between the craniotomy and drug cases, i.e., between a case in which DDE does not allow the Aist to act and a case in which it does, has in recent years come under attack once again. The majority of these critical efforts, including Jonathan Bennett's influential paper on 'Whatever the consequences'[7], have focused _ I am tempted to say, exclusively _ upon DDE's use of the distinction between direct intention and foresight of

consequences to draw a moral difference between (say) the craniotomy and drug cases. Rightly, they attack DDE's view of the drug case, and argue that there is no moral difference between it and the craniotomy case. For example, in his paper on 'Intention and Punishment', Hart questions DDE's use of the distinction between direct and oblique intention, not to separate the cases into one in which a doctor intentionally kills his patient and one in which he knowingly brings about or hastens the death of the patient, but rather, once the cases are thus separated, to draw a moral difference between them; according to Hart, 'neither these verbal differences nor the differences in causal structure are correlated with moral factors', and in his view, there is in the two cases 'no relevant difference between them on any theory of morality'.[8] He concludes:

> It seems that the use of the distinction between direct and oblique intention to draw the line as Catholic doctrine does between what is sin and what is not sin, in cases where the ultimate purpose is the same, can only be explained as the result of a legalistic conception of morality as if it were couched in the form of a law in rigid form prohibiting all intentional killing as distinct from knowingly causing death. 9

What Hart fails to add is that, in examining DDE, talk of sin serves only to obscure matters; specifically, it obscures the distinction between the rightness or wrongness of acts and the goodness or bad-ness of agents, and the considerations relevant to each. DDE does

not itself specify either what makes acts right or wrong or which acts are right or wrong; it is assumed, in order even to consider whether the doctrine is applicable in a particular case, that one knows whether the contemplated act is intrinsically wrong.[10] Thus, what is at issue is whether DDE and its use of the distinction in between direct and oblique intention establish a difference in the goodness or badness of the doctor in the two cases, in one of which he intentionally kills his patient, in the other of which he knowingly brings about or hastens the patient's death. Hart, rightly, thinks not.

Just as talk of sin can obscure matters, so too can labelling the doctor in the craniotomy case but not the drug case a murderer. For, even if we define murder (unless suicide by definition is murder) as 'the intentional taking of another, innocent human life', we are still left with the question of whether the doctor who knowingly brings about or hastens death, not as his aim or part of his aim, but as a foreseen and inevitable consequence of his act, is guilty of murder. Moreover, since the law may not recognize any distinction between intention and foresight, if the doctor in the one case but not the other is labelled a murderer, then we may also have to raise the question of whether the concept of murder is the same in the legal and (Aist) moral spheres.[11] Finally, labelling the doctor in the one case a murderer can obscure important questions

of responsibility: labelling him in this way ascribes guilt to him, but this guilt sense of 'responsible' is not the only one germane to assessing the goodness or badness of the doctor in the two cases, as we shall see.

Caution must also be exercised in another respect. In his book The Sanctity of Life and The Criminal Law, Glanville Williams attacks the use of the distinction between direct and oblique intention in another way:

> It is altogether too artificial to say that a doctor who gives an overdose of a narcotic having in the forefront of his mind the aim of ending his patient's existence is guilty of sin, while a doctor who gives the same overdose in the same circumstances in order to relieve pain is not guilty of sin, provided that he keeps his mind steadily off the consequence which his professional training teaches him is inevitable, namely the death of his patient. ...If [DDE] means that the necessity of making a choice of values can be avoided merely by keeping your mind off one of the consequences, it can only encourage a hypocritical attitude towards moral problems. 12

There are two questions to be kept distinct here. First, can one withhold intention from an evil effect of one's act - can one avoid willing that evil effect - if one knows that it will in fact occur? There is no easy answer to this question, and certainly some defenders of DDE have faced up to it.[13] In many cases, whether one can withhold intention from an evil effect will depend, I think, upon whether the

effect is inevitably or inseparably connected with the proposed act. If in the middle of the night I set fire to a building of mine which houses people who work for me, in order to get rid of this horrible shack that spoils the appearance of my land, then even if I claim to have been directly aiming at or intending only to rid myself of this eyesore, I must realize the people asleep in the building will very likely die as a consequence of my act. Killing these people is no part of my aim, nor do I achieve the good effect (getting rid of the shack) through or by means of the evil effect (killing the people), nor do I want or desire the deaths of these people; nevertheless, it is doubtful, in the event, whether I can withhold intending their deaths, since death seems <u>inevitably</u> linked with setting fire to a dwelling-place in the middle of the night. In other cases, withholding intention from an evil effect appears doubtful because of the effect's <u>inseparable</u> link with the act. If I drop a cannister of napalm into the midst of a village in order to deter snipers from firing at ground troops, it must be doubtful that I can withhold intending the deaths of the villagers. For napalm is an anti-personnel weapon, and killing and maiming are the very functions the weapon was perfected to fulfill; to withhold intending to kill and/or maim the villagers in this case, I should have to think both that the function of the weapon is other than I know it to be and that the napalm itself, once dropped, will be selective in alighting only upon

snipers and not upon villagers. I think the drug case is of this sort, i.e., that since the doctor knows and admits that death is inseparably linked causally with a certain dosage of a drug, then if he administers this dosage, it must be doubtful that he can withhold intending the patient's death. Nor would he be helped, even if we were to agree that there are degrees of foresight, e.g., that consequences may be foreseen as certain, probable or merely possible, and to agree also that in the case of consequences foreseen as very slightly probable or as merely possible intention can be with_ held; for the doctor foresees as a certainty that death will result from administering a certain dosage. It is tempting to conclude, then, that the degree to which we cannot withhold intending an evil effect is a function of the degree to which we foresee this effect as a certain outcome of our act; the more certain we are, the more doubtful we can withhold intending it.

A second question, to be kept quite distinct from the problem about withholding intention, is whether, as the supporter of DDE thinks, an agent is responsible only for what comes about through a direct intention of his; or whether, as Hart and Glanville Williams think, and with whom I concur, an agent is not only responsible for what he brings about intentionally but also for what he foresees as a certain and inevitable or inseparable consequence of his act.[14] In fact, in addition to the supporter of DDE, there is a further

opponent of this last view. In discussions of oblique intention, questions about what the agent wants have been thought to obtrude. In the case of Director of Public Prosecutions v. Smith[15], attention was focused on the question of whether a man intends the foreseen though unwanted consequences of his acts; in commenting on the case, Anthony Kenny noted that, on the ordinary meaning of 'intends', 'an agent intends an action only if he (a) knows he is doing it and (b) wants to it either for its own sake or in order to do some other action'.[16] Thus, if the doctor in the drug case foresaw his patient's death and desired or wanted that death, then he intended it, and is, accordingly, to be held responsible for it. One question these comments about wanting raise is whether the law should distinguish between foresight and intention, i.e., of whether the legal use of 'intends' should be construed as entailing 'foresees and wants', a development which Kenny favors. However, for our purposes, another view that may arise out of these comments on foreseeing and wanting is that an individual is not responsible for foreseen but unwanted consequences.[17] For it can appear as an upshot of the view that a man who foresees but does not want a certain consequence has not intended it, and, therefore, that he is not responsible for it.

These two views _ that one is responsible only for those conse_ quences of one's acts that one directly intends, and that one is not responsible for consequences of one's acts which one foresees but

does not want _ are but <u>different</u> sides of the same coin; both are mistaken. (a). That one is responsible for more than one directly intends seems clear enough, <u>DDE</u> notwithstanding. We may adapt a familiar example to show this. Jack and Jill are walking to work when simultaneously they see a wallet bulging with money on the path before them; both dart forward for the money. Jill gives Jack a terrific push which she intends to knock Jack over, and which she foresees as likely to cause injury to Jack. It is neither her pur_ pose nor desire to injure Jack; her only concern is to get the money. Jack falls down and sustains serious injury. Though Jill does not directly intend injury to Jack, she is nevertheless responsible for whatever injury Jack sustains. (b). The degree of responsibility an agent incurs for foreseen consequences <u>may</u> depend upon his assess_ ment of the likelihood of those consequences coming to pass. At one extreme, we have the napalm and drug cases, where the consequences are foreseen as certain, and we <u>may</u> assess degree of responsibility accordingly. At the other extreme will be cases where the likelihood of the foreseen consequences occurring is negligible; and the vast majority of cases will fall somewhere between the two extremes, with consequences foreseen as more or less probable. Assessing respon_ sibility may, therefore, more often than not depend upon the degree of probability with which the agent estimated the consequences to occur. (c). That one is responsible for consequences one foresees but does not directly intend is also clear. Jack shoots at Jill

knowing himself to be an excellent shot and intending only to frighten her; he foresees that Jill could be killed but neither intends nor desires that she should be, nor thinks that she will be, given his marksmanship; in fact, Jill is killed. Jack is responsible. (d). Responsibility for bringing about a set of consequences can be ascribed whether we consider a case such as the preceding one, involving a positive act, or a case of negligence. A man leaves his gas stove burning when he leaves his apartment; he foresees that the stove could explode, and that fire could break out in his and other apartments, but he neither intends nor wants such an explosion, nor estimates its likelihood as very high (his stove is new); nevertheless, if an explosion occurs, and the apartments burn, he is responsible. (e). In passing, it may be of interest to reflect on whether responsibility for a consequence can attach to an agent who could not have foreseen that his act was likely to have this harmful effect. For example, remoteness in the causal sequence is sometimes offered as an explanation of why a harmful effect could not have been foreseen. But we must be careful, for <u>some</u> remote consequences at any rate can be predicted. If I set out a boxwood, I can predict that in fifty years time it will completely shade my front lawn; and if I place cucumbers in brine, I can predict that I shall have pickles for my birthday. Even in the case of remote consequences that could not have been predicted at the time of acting, it is not so clear to me

that we are not in cases at least partially responsible for them. It seems to me arguable that the 1914-18 war may be held to be a partial consequence of a single individual's act of assassination, even though numerous other human agents subsequently intervened in the causal sequence; and if the war is a partial consequence of his act, even though he could not have predicted that war would result, is he not <u>partially</u> responsible for that war? (In short, I am inclined to accept the view that the consequences of our acts <u>may</u> well extend indefinitely into the future - and that the imposition of limits or boundaries to consequences is artificial - and also to accept the view that, even though we cannot at present predict what these consequences will be, we <u>may</u> be partially responsible for them.) Another interesting case in this context is this: in addition to remoteness in the causal sequence as an explanation of why a harmful effect could not have been foreseen, it may be argued that the agent <u>had no reason to think</u> now or ever in the past that this effect could be a consequence of his act. A tramp sets fire to a shack in which he has been living alone for many years; unknown to him, another tramp has crawled inside to sleep; the second tramp perishes in the fire. Here, the tramp had no reason for thinking another tramp was inside the shack - in this sense, he could not have foreseen this eventuality; moreover, no one else in that situation would have foreseen the presence of the second tramp, not even the 'reasonable man'. Is the first tramp responsible for the death of the second?

If the question is asking whether the first tramp is to be blamed, then we can be reasonably confident in the answer that he is not. If the question is asking whether the first tramp is answerable for the death of the second, the reply may be that he is, depending upon such factors as whether the first tramp was committing a crime in firing the shack (arson), whether he owned the shack, etc. If the question is asking whether the first tramp caused the death of the second, then the answer is clearly that he did. (In some Continental legal systems this answer alone is sufficient to render the first tramp guilty of manslaughter.) The death of the second tramp was a consequence of the act of the first, and he is responsible for that consequence; but this does not entail that he is to be blamed, or that we are to regard him as a wicked man.

The views that one is responsible only for those consequences of one's acts that one directly intends and that one is not responsible for consequences which one foresees but does not want are both to be rejected, and, among others, for considerations which themselves lend support to the Hart_Williams position, viz., that one not only is responsible for what one brings about intentionally but also for what one foresses as a certain and inevitable or inseparable consequence of one's act. The evidence, therefore, is against the supporter of DDE successfully using the distinction between direct and oblique in_ tention to establish a difference in the goodness of the doctor in the

craniotomy and drug cases, i.e., to establish a moral difference between the case in which he intentionally kills his patient and one in which he knowingly brings about the patient's death.

## V.3 Control and responsibility

If we turn from the drug case, and the distinction between direct and oblique intention, to the craniotomy case, I want to say, contra DDE, that the doctor is responsible for the woman's death. In this case, defenders of DDE purport to find a moral difference between operating and not operating: if the doctor performs the craniotomy, he is said to have directly intended the death of the foetus, whereas if he does not operate, the mother is said to be permitted to die. They argue that, while it is forbidden to intend the death of the foetus directly, it is not forbidden to permit the death of the mother. In sum, there is an alleged moral difference in this case between operating _ and intending a death _ and not operating _ and permitting a death, and in this and the next few sections I wish to consider just what this alleged moral difference might be supposed to consist in. (I assume that the claim that the difference consists in the fact that one death is directly intended while the other is merely foreseen has already been dealt with.)

The difference in question might be thought to consist in the fact that in operating the doctor is responsible for the outcome whereas in not operating he is not; DIE supports such a view, but it

is, I think, mistaken. The doctor is responsible, whatever the outcome of his intervention or action in the woman's case (though not, of course, if, out_of_the_blue, though no fault of his own, she dies, e.g., of cholera). Now what are we saying when we say that he is responsible whatever the outcome of his action? The concept of responsibility oscillates among (at least) three meanings. On one meaning, perhaps the most common, to be respon_ sible for X is to be the cause of X; on a second, to be responsible for X is to be guilty of having done X, i.e., guilt is laid at one's door. On a third, however, to be responsible for X is to say that there is a case to be put and so a case to be answered; while guilt is not ascribed, there is the implication that we shall go into the answer that is forthcoming. It is a variant of this third sense of 'responsible' that I am using when I suggest the doctor is responsible for the woman's death. That variant is, that while we are not labelling the doctor an evil man or a murderer, we are saying something about his control of the situat_ ion, i.e., his control over the course of events in the woman's case. This situation is one within the doctor's control, in that he may operate or not operate; whatever results in her case _ and here and elsewhere I exclude the circumstance just mentioned, where, out_of_the_blue, she dies of cholera _ cannot be said to be independent of this control.

Of course, some men lose control of a situation. It seems possible in D.P.P. v. Smith, for example, that Smith became excited and angry and lost control of the situation; events, so to speak, took charge of him rather than vice versa. This is why it is perhaps unjust to ask what an unexcited, reasonable man would foresee in this case; for the question assumes the very control of a situation that through excitement or anger or whatever the agent has lost. Nothing like this, however, exists in the case of the doctor, so that, even if excitement, etc., are defeating conditions of control, the doctor cannot make use of this fact. This is why talk of the woman's death as 'unavoidable' must not be allowed to pass[18]; her death is unavoidable only if the doctor exercises control in one way rather than another. Moreover, to say that by not operating the woman is 'permitted to die' is to say that the doctor has made his decision and decided to exercise his control in this way; the woman can be said to be 'permitted to die' only if the doctor exercises his control in one way rather than another. The woman is not simply 'permitted to die'; she is 'permitted to die' by the doctor. Of course, the woman need not have died; but then, too, the foetus need not have had its skull crushed; however, if one or the other comes about, there is no reason to refrain _ nor is anything achieved by so refraining _ from saying that the doctor 'permits' the woman to die or that the doctor crushes the skull of the foetus. It is my view here, then, that agency entails responsibility; in a real sense,

the doctor is to be held responsible whichever course he adopts, since he is the one confronted with the choice.

If to have control of a situation ensures responsibility for the eventualities that issue, it does not follow that to lose control of a situation ensures that responsibility lapses. A man who, through use of hard drugs has lost control of himself, need not be immune from responsibility for acts which he performs while under the influence of those drugs; and, clearly, he can be responsible for losing control in the first place, which he would not be if somebody drugged him without his knowledge. Slightly differently, the captain of a ship, resting below, is responsible for the disaster that befalls his ship, even though he was asleep at the time. Then we have situations in which the doing or not doing of something is 'beyond one's control'. The desire of an alcoholic for drink may be beyond his control, so that nothing short of physically stopping him will prevent his grabbing the bottle. Here, too, the alcoholic may be held responsible for getting himself into a position where situations become beyond his control. In fine, in cases of loss of control, for the purposes of allotting responsibility, we shall want to determine whether such loss is self-induced, e.g., through self-administered drugs, alcohol, etc., or whether it is induced by others, e.g., through force, violence, etc.; in cases of the latter sort, control is lost and responsibility lapses.

(The threat of unpleasant alternatives or duress is not a defeating condition of control: 'force' in this context means physical force, not the threat of physical force.)

Granted this element of control, it might be said the moral difference between operating (intending a death) and not operating (permitting a death) consists in the fact that in operating the doctor intervenes and exercises his control, whereas in not operating he does not intervene but allows events to take their course. It is a mistake, however, to think that the doctor's control of a situation can manifest itself only through acting and never through not acting. How he exercises his control is his responsibility: the woman need not have died, if the doctor had decided differently, decided to operate rather than not operate. It is true that by operating the doctor intervenes whereas by not operating he does not; but the important element here is not whether the doctor intervenes but that the decision whether to intervene is his; for the woman is permitted to die by the doctor only if he decides not to intervene and operate.

I conclude, then, that the alleged moral difference between operating (intending a death) and not operating (permitting a death) cannot consist in the fact that on the former but not the latter course the doctor is responsible for the outcome.

## V.4 Acts of omission

It may be argued that the alleged moral difference between

operating and not operating is to be found in the distinction between acts of commission and acts of omission. To an opponent of DDE, the craniotomy case may appear in this light: operating is acting, is a doing or a doing_something, whereas not operating is not acting, a not_doing, a doing_nothing; the former may be thought of as acts of commission, the latter as acts of omission.[19] This opponent will then point out that we are morally responsible for our omissions as well as our acts of commission, so that there is no need to employ the concept of control, no need to raise questions about the sense of 'responsible' in which by not operating the doctor is said to be responsible for the woman's death. For on this view the doctor is guilty of an immoral omission. In my view, this opponent of DDE is misguided.

It is said that by not operating the doctor does nothing; so let us inquire after the necessary and sufficient conditions for a not_doing or a doing_nothing to amount to an immoral omission. Suppose several men, all of them able_bodied swimmers, stand by and watch a non_swimmer drown: to say that they 'did nothing' at the very least means that they did nothing about a particular thing, for they may well have been doing something, e.g., exercising. The men took no action with respect to the particular event in question. However, failure to act is not an omission. It is necessary to add that the swimmers had the ability to do something to save the non_swimmer;

a non_swimmer in their situation, in which only by jumping in could the drowning man be saved (no boat available, etc.), does not have the ability to save him. Moreover, these able_bodied swimmers have not only the ability but also the opportunity or occasion for saving a drowning man; they will not omit to save a man if they have not the opportunity to save one. We have, then, the following conditions necessary for an individual's failure to act to amount to an omission:

(i) A does not X
(ii) A has the ability to X
(iii) A has the opportunity to X.

Besides these conditions, we need also some indication that the swimmers are expected to take action in respect of the drowning man; that is, that they took no action to save the man when they are or may be expected to. A fourth condition, therefore, is necessary:

(iv) A is or may be expected to X. 20

Now suppose the swimmers admit all this, but offer as reasons for their doing nothing that they did not want to become involved or that it was none of their affair: are they guilty of an immoral omission? I think not; their failure to act may amount to an omission but it does not amount to an immoral omission, for thus far there is no suggestion that they either ought or are required to act. Thus, a lifeguard who satisfied conditions (i) _ (iv) would, I think, be guilty of an immoral omission, since he is required to act. Therefore, to be guilty of an <u>immoral</u> omission a fifth condition is

necessary:

(v) A ought or is required to act.

One point needs amplification, namely, what it is that requires one to act. The reason we expect the lifeguard to plunge in is our knowledge that he is required to do so; he knows, moreover, that he is required, since the post of lifeguard is defined by its duties. Of course, this requirement is not in disregard circumstances: if the drowning man is in a raging sea, then the lifeguard may have cause for not going in; but it is important that he show cause for not acting. A job, a function, then, may explain why one is required to act. (In this sense, condition (v) above could be regarded as a stronger version of (iv).) Now suppose a second of the able-bodied swimmers is the son of the drowning man, and let us assume that there are no circumstances of the raging sea variety to impede rescue: if he fails to act, is the son guilty of an immoral omission? Clearly, knowledge of the family tie may be what leads others to expect the son to jump in, but it is not clear that the son is required, e.g., out of any duty to his parent, to rescue his father (though the common presumption will almost certainly be that he is). Finally, suppose the other swimmers are not lifeguards and are not relations of the drowning man, and neither know who he is nor anything about him. If they fail to act, whether or not they are guilty of an immoral omission may depend amongst other things on whether we

regard the swimmers _ not these men in particular, but anyone who might find themselves in a similar situation _ as being under a duty, e.g., to render mutual aid. Clearly, it is problematic whether we have any such duty, which may help to explain the common failure to act in this sort of case.

Well, then, what of the doctor in the craniotomy case? By not operating, is he guilty of an immoral omission? I think not. In most respects, his situation is the same as the lifeguard's: conditions (i) _ (iv) are satisfied, and he seems clearly required to save life. The reason we expect him to 'do something' is because we know his job or profession requires him to. But the important feature in the doctor's case is this: while he is under a duty to save life, that duty does not specify <u>which</u> life, mother or foetus, he is to save; <u>he</u> decides which it is to be. Of course, if he attempts to save neither mother nor foetus, then he is guilty of an immoral omission; but if he saves the child, though we may complain on behalf of the mother, he can nevertheless be said to have done what he is required to do, i.e., save life.[21] The doctor, then, is not guilty of an immoral omission. Now while I accept this conclusion, I think it can be misleading to characterize the doctor's duty as 'to save life'; for in the craniotomy case, so long as he attempts to save either mother or foetus, he will have done his duty, <u>and so</u>, it may be thought, is immune to criticism, morally. However, even

though his duty does not specify which life to save, the doctor's decision in this respect, I want to suggest, is not immune to criticism. For his decision is not made *in vacuo*, but in a particular situation, with attendant circumstances and consequences; and in order to come to a decision about what he ought to do, the *Cist* will urge the doctor to take into account all such factors as the mother's mental and physical health, the fact that she is already a functioning unit within a family, the fact that others depend on her, etc. In other words, by not operating, though the doctor will not be guilty of an immoral omission, he may have come to the wrong decision as to which life to save; and talk of his having 'done his duty' by saving either life can mask this fact. For example, in her note on 'Who is Wronged?', Professor Anscombe argues that it is not clear, in a case in which a doctor can save the many or the few, that the doctor must save the many: his duty is to save life, and this he does by saving the few; and she queries which of the many is entitled to complain of the doctor's act.[22] Now apart from taking up the question of whether the doctor's duty here is to save life or to save as many lives as he possibly can, or the question of what difference number makes, I think talk of the doctor having done his duty in saving the few may possibly mislead us into thinking that the doctor's *decision* is immune to criticism. This is not the case: the *Cist*, perplexed by such a decision, will inquire after those

considerations the doctor took into account and those he left out (a point to which I return in the next section) in reaching this decision. For even *prima facie*, it is difficult to grasp what considerations there are (he did not lack medicines, etc.) that could have led him to such a decision[23]; and even to one who thinks the doctor does his duty whether he saves the few *or* the many, an explanation may be required for why he chose the course he did, for why he did his duty so narrowly and not more broadly and so more generously. (Nor is it difficult to conceive of answers to these questions that point to defects in the goodness of the doctor.) The doctor does not escape responsibility for what happens to the many, for the decision as to which lives to save is his, and the *Cist* will hold him answerable for what he decides. He has control over the situation, and the many would not have perished if he had decided differently.

## V.5 *Acting and refraining*

It may be argued that the (alleged) moral difference between operating and not operating in the craniotomy case is to be found in the distinction between acting and refraining. That is, it may be said that a not-doing or a doing-nothing in the doctor's case amounts to refraining or 'letting the mother die'; and since refraining or 'letting die' implies wrong-doing, the doctor does wrong in not operating.[24] I do not accept this argument.

The necessary and sufficient conditions for a not_doing to amount to a refraining are, I think, as follow:

(i) A does not do X
(ii) A has the ability to X
(iii) A has the opportunity to X
(iv) A considers Xing but rejects it.[25]

The important difference between refraining and omitting[26] is that to refrain from Xing one must consider or deliberate whether to X. If I do not turn out the lights or feed the baby, I may in certain circumstances omit to do so; but there is no suggestion that I con_ sidered whether or not to turn out the lights or feed the baby. In the craniotomy case, however, the doctor considers whether to save the mother, and the case seems a clear instance of refraining. However, it does not follow from the fact that the doctor refrains from saving the mother that he has ipso facto done wrong. If I debate the merits of each but decide to stay inside to practice the piano and refrain from going walking, I have not done something wrong or something I ought not to have done. Why should the same not be true in the doctor's case? The only reason I can see why one might want to treat the craniotomy case any differently is because one goes on to construe refraining in the doctor's case as failing in his duty; and if he fails in his duty, he has done wrong. How_ ever, the doctor's duty does not indicate which life he is to save; he has to decide this, since he is in control. (Of course, a case might be devised in which the doctor's duty did indicate which life

he was to save, e.g., a case in which he has the choice between saving his patient or another doctor's but cannot save both.) Moreover, we must be careful not to confuse the view that 'the doctor ought to save the mother' with the view that 'the doctor's duty is to save the mother'; which life he is to save is not specified by his duty, but his decision as to which life he saves can, I am suggesting, nevertheless be open to criticism; and at this level one can say that the doctor made the wrong decision and 'ought to have saved the mother'.

It is not difficult to conceive of situations in which such criticism may be levelled against the doctor; he may, for example, fail to take into account relevant circumstances and consequences in the woman's case, such as the effect of her health, on her family, etc. More seriously, suppose the doctor knows that the foetus will almost certainly die anyway (or, for that matter, will almost certainly be born with a serious malady, such as spina bifida): to a Cist, should the doctor not take such knowledge into account in reaching his decision, then he could leave himself open, not only to the charge that he came to the wrong decision and ought to have saved the mother, but also to the ascription of blame and guilt.[27] This is a point of the utmost importance. In virtue of his control in the craniotomy case, the doctor is responsible for the outcome, though herein we do not ascribe guilt or blame to him; let us call this

control responsibility. To a Cist, who wants the circumstances and consequences of decisions, acts, policies, etc., taken into account and who finds in a particular case an important circum_ stance or consequence left out of consideration, such as know_ ledge in the craniotomy case that the foetus will die anyway, control responsibility can be (and in the case envisaged doubtless will be) a ground on which he goes on to ascribe moral blame and guilt. That is, what the doctor takes into account and what he leaves out, what he includes as morally relevant and what he ex_ cludes, in reaching his decision, may make him liable to blame, at least in the eyes of the Cist; and the doctor's answer to why he saved the foetus, when he knew it was going to die anyway, will have to be a good one to prevent the Cist from ascribing blame. In certain situations, then, control responsibility can be the basis upon which one holds the agent responsible in the 'blame' or 'guilt' sense, and this explains why the Cist desires to establish firmly that the decision as to which life to save is the doctor's.

## V.6 Positive and negative duties

In her paper on 'Abortion and Double Effect', Mrs Foot con_ strues the difference between operating and not operating to be the difference between positive and negative duties, a distinction

she broadly characterizes as that of bringing aid and avoiding injury.[28] That is, in the craniotomy case, by operating the doctor brings aid, whereas by not operating he avoids injury; and Mrs. Foot is in no doubt that the duty to avoid injury is more stringent than the duty to bring aid.[29] Moreover, she thinks that, in the usual cases in which DDE is employed, 'the distinction between avoiding injury and bringing aid is very important indeed'.[30] I doubt, however, that this distinction will carry the weight Mrs. Foot would have it carry.

First, and most importantly, where have we got if we agree that by operating the doctor brings aid and by not operating he avoids injury? If we are unsure of what the moral difference is between operating (intending a death) and not operating (permitting a death), will we not be equally unsure of what the moral difference is between bringing aid and avoiding injury, since they are in the case at hand but other words for describing operating and not operating? I think we shall have failed to advance the argument by this reformulation, unless one or other of the two conflicting duties always is overriding. For if this is not the case, in each instance the doctor will have to deliberate whether the duty to bring aid is more or less stringent than the duty to avoid injury; how is the doctor any better off weighing these two things than in weighing the merits of operating and not operating? On the other

hand, if one of the duties, e.g., avoiding injury, is always overriding, then once again circumstances are ignored. Is the duty to avoid injury overriding in a case in which the doctor knows the foetus will almost certainly die anyway? To think it overriding even in this case evinces the same disregard for 'saying all that can be said about a case morally' that we have noticed already. Second, Mrs. Foot is mistaken in thinking the bringing aid_avoiding injury distinction is applicable to the usual range of cases where DDE is employed. For example, a student in order to get his degree in literature must read an erotic text; he foresees that he will be aroused but neither directly intends nor wants to be aroused; and DDE may be invoked to show that reading the text is permissible. I do not see how to characterize this case in terms of bringing aid and of avoiding injury, since to speak of reading as bringing aid and of not reading as avoiding injury does not make sense; and, in any event, it is not clear what we are to make of 'bringing aid to oneself'. Or suppose a woman of great beauty but modest reputation comes to our small village: she foresees that when she goes about the streets she will evoke feelings of lust in us villagers, but she neither directly intends nor desires that we should have those feelings. DDE will support the view that her walking through the streets is permissible; but is her walking through the streets

bringing aid to us villagers? In a sense, it is, viz., it may satisfy our lusts, but this is not the sort of thing I think Mrs. Foot had in mind. Lastly, and rhetorically, is there not a point beyond which any duty to avoid injury is extinguished? The answer to this question is by no means clear, which is one of the reasons why proponents of euthanasia often seize upon the cases of 'human vegetables' to build their plea. Any claim that the duty to avoid injury never lapses will have to be defended in the face of such cases.

V.7 Place in the causal series

Finally, and briefly, the defender of DDE purports to find a moral difference between operating and not operating in the cran_ iotomy case: just where, then, does he think the alleged moral difference lies? The key to answering this question is not to be found in the distinctions between direct and oblique intention, act and omission, acting and refraining, bringing aid and avoiding injury; rather, it is in my opinion to be found _ where it has been in front of us all the while _ in a supposition that underlies the third condition for the legitimate employment of DDE. This condition states that if an act has two effects, one evil, the other good, then the evil effect must not be a means of bringing about the good effect. The point is sometimes put this way, that the good effect must be produced 'directly' or 'immediately' by the act and

not by the evil effect; otherwise, moral evil will be done that good may come, and this is proscribed. Thus, in the craniotomy case, the good effect _ the saving of the mother _ must be pro_duced directly by the operation and not by the evil effect _ the killing of the foetus; otherwise, this third condition for the employment of DDE will be violated. I believe it is just here that the moral difference between operating and not operating is supposed to be found, for underlying this view that the good effect must be produced directly by the act is the supposition that it is morally significant in the evaluation of consequences or of acts where a consequence falls in the causal series. On this supposition, it is place in the causal sequence (and not, it must be stressed, in the temporal order) that morally distinguishes saving the mother and causing the death of the foetus; and the intention or foresight of consequences only becomes an important issue in so far as it enables us to deter_mine the position of a consequence in the causal series. Though I cannot prove this supposition mistaken, I nevertheless reject it.

It is difficult to understand how situation in the causal series of act and consequences can have any bearing on the goodness or badness of an effect. Saving the mother is not good if it comes before the foetus' death but bad if it comes after it in the causal

series: the goodness of saving the mother is not contingent upon position in the causal order, for saving life is saving life, and all human life is to be valued. It is a good thing in virtually all situations that life is saved, and position in the causal order will not negate this goodness. However, it might be argued that although position in the causal series has no bearing upon the goodness of the effects, it does have a bearing upon the morality of the act.[31] Thus, act A causes good consequence B and bad consequence C: on the view in question, it makes all the difference for the morality of A that B occur before C in the causal order (the temporal order makes no difference); change B's place in the causal series and the morality of A changes. But is this the case?

Suppose I am imprisoned unjustly and during the course of my show trial am asked to name my fellow conspirators in the liberation army. I decide to perjure myself, but I do not make up fictitious names; rather, I name members of the present government, including the chief of police, whom I hope to discredit. The chief, discredited, commits suicide and so ceases to persecute my friends. The good effect, that my friends are no longer persecuted, comes through the bad effect, that the chief commits suicide, and the chief's suicide is a bad effect of my act of committing perjury. I think that to any non-Aist, to anyone, that is, who accepts that there are no acts absolutely proscribed (including perjury), whether

the cessation of the persecution of my friends comes before or after the chief's suicide in the causal order is of no significance to the rightness or wrongness of my act of perjury. For if my friends had ceased to be persecuted before the chief's suicide, my act of perjury would not suddenly have altered in its moral evaluation. In a word, that my act has consequence X is important to Cists, be that consequence near or remote, predictable or unpredictable, in the causal chain, since it is relevant to the act's rightness. But that consequence X comes before or after consequence Y in the causal series is of no importance to assessing rightness. (In another context, I return to place in the causal series in Chapter X.)

---

From these and the other considerations of this chapter, I conclude that, even if we leave aside the fact that DDE only makes sense against a backcloth of acts wrong apart from their internal and external consequences, we still have grounds for rejecting it as a principle of conduct. Hence, by recourse to DDE, the Aist will not extricate himself from the dilemma in which he is placed through conflicts between absolute proscriptions. Moreover, if only we recall that what a person is prepared to permit, allow or see happen is as much a part of his moral make-up as what he

does, and that he is responsible for both, we shall be far less likely to go off at a tangent in discussions of DDE. For we shall have before us a simple reminder that 'permitting to die' or 'permitting nature to take its course' are not per se defences against control responsibility, which in turn may be a ground for the ascription of blame.

## V.8 Legal doctrine of necessity

It is high time somebody wrote a good book on the relations and differences between the doctrine of necessity in the common law and DDE, but I can hardly take up this subject in detail in this essay. Still, I do wish to say something on the matter, since the doctrine of necessity covers a good deal of the same ground that DDE is intended to cover.[32]

In civil cases, necessity has long been recognized as a permissible defence: if to prevent flood waters from spreading I take sandbags from your lot, it is no tort.[33] In criminal cases, however, the matter is not nearly so clear, for the limits within which necessity operates as a permissible defence have never been fully defined; as a result, what conditions must be satisfied in order for a plea of necessity to be entered successfully have never been made explicit.[34] The best-known case in which necessity was entered - unsuccessfully - as a defence to a criminal charge was in

R. v. Dudley and Stephens[35], in which two shipwrecked men with almost no food for twenty days killed and ate the cabin boy who was with them. However, to give an example where necessity does serve as a justification for action, we can cite more ordinary cases. Usually, in an operation, the surgeon has the permission of the patient (or perhaps next of kin) to cut (it being otherwise against the law to do so), but situations arise, obviously, in which this permission cannot be obtained; nevertheless, the doctor proceeds to operate and necessity is his justification.

A defence of necessity is entered not to show that the accused could not have done otherwise, for 'necessity' in this sense renders what was done involuntary, and the accused is said not to have acted at all in any full sense of 'act'; rather, this defence is entered to show that there were strong and compelling reasons for acting as the accused did. While it is not clear what counts as a strong and compelling reason in these cases (e.g., in R. v. Dudley and Stephens it appears to be the finding of the Justices that the saving of one's own life even in extreme circumstances is no defence to murder), it is nevertheless possible to discern two conditions that must be satisfied if a plea of necessity is to be made out: (i) the evil averted must be greater than the evil done, and (ii) no more evil may be done than is necessary to avert the greater evil. The most crucial decision, of course, is which evil

is the greater and which the lesser. In many cases, in which we would ordinarily describe ourselves as having to choose 'the lesser of two evils', we may be quite sure which is the lesser evil: thus, in the eyes of the law it is an evil for a doctor or anyone to lay open an individual, but this is a lesser evil than the death of that individual. Obviously, however, cases in which doubt exists as to which is the lesser evil are the hard cases. Is it a greater evil to murder a cabin boy to save two lives than for three people to starve to death on an open raft? To all but an Aist, there is no easy answer to this question.

Now it has sometimes been proposed that the doctrine of necessity be extended to cover certain 'hard' cases of abortion and euthanasia; that is, that legislation be enacted which incorporates necessity as a defence and/or justification for doctors aborting and killing in certain cases. I am inclined to look favorably upon this proposal, even though in some of these cases it will be difficult to decide which is the lesser of two evils. In the craniotomy case, the deaths of the mother and of the foetus are evils; to decide which is the lesser a doctor will not only want to consider the interests of the mother but also not to lose sight of the interests of the foetus.[36] And a host of other considerations, some of which were set out earlier, may enter as well. But the decision remains a difficult one, and doubtless will not

always be made in the same way.

In the face of such legislation, two important issues we have discussed would come more directly into focus. First, in cases where necessity operated as a defence and/or justification, the doctor would be left in control; for he is still the one in the position both of having to decide which is the lesser of the two evils and which course of action is then to be adopted. The crucial decisions will still be his, as they are now, in cases where necessity justifies a surgeon in operating; and what occurs in these cases will not be independent of how he exercises his control. In a word, necessity will not vitiate control responsibility, which in turn may be a ground for the ascription of blame. Second, the distinction between direct and oblique intention will not, contra DDE, be relevant to a judgment of the goodness or good character of the doctor. In the craniotomy case, one that will be covered by the legislation, it is true that the doctor must decide whether to operate or not, and that how he makes this decision and what decision he makes may be germane to an assessment of his character; but in a case in which he must act out of necessity, whether he directly intends the death of the mother or the death of the foetus will not be revelative of character. Nor, when the doctor must act out of necessity, will directly intending the death of the foetus make the doctor more wicked than one who, in the same

situation, directly intends the death of the mother. In those cases where there is an overlap between DDE and the doctrine of necessity, as promulgated in the new legislation, then, it will be tempting for the Cist to stand up and boldly proclaim that it is at last clear that the distinction between direct and oblique intention is not the index of character the defenders of DDE take it to be, a view he has been arguing for all along.

# VI

## ACT_UTILITARIANISM AND THE HARD CASES ARGUMENT

We come now to the second group of issues that surround C, to those that stem from the inclusion of C within AU theories. While there are many such issues, we shall confine ourselves here to the question of whether the AUian can support the useful social rules and institutions of our society; or, conversely, whether his position commits him to undermining them. Our aims are to show that a new species of AU _ tempered act_utilitarianism _ can accommodate such rules and institutions, and so, in this limited sense, may be put forward as a satisfactory utilitarianism; and to show that the traditional species of AU is indeed committed to undermining, e.g., the useful social institution of promising.

In this chapter, I shall only briefly touch upon the best_known argument against AU theories. My reason for mentioning it at all _ the ins_and_outs of this argument are by now exceedingly familiar _ is that it sets the stage for the introduction of tempered act_utilitarianism.

### VI.1 The 'plain man's' fear

In its general characterization, AU is the view (i) that an act is right if, and only if, its consequences are at least as good as those of any alternative[1], and, as a corollary, (ii) that an ought to be performed if, and only if, its consequences are

better than those of any alternative. What we ought to do as well as what is right is linked to optimificity, to the maximization of utility, where 'utility' is a blanket term, to be filled in by whatever standard of intrinsic goodness we adopt. Now this linkage with optimificity has been alleged to lead in particular cases to conflicts with our 'ordinary moral convictions', or the views or opinions of the 'plain man'.[2] Such conflicts form the basis of what I call the hard cases argument, which consists in setting our certain (usually 'hard') cases, comparing what the AUian says with what the 'plain man' thinks about them, and using any discrepancy between the two as an index of the inadequacies of AU. Most criticisms of AU have taken this form; and while I think there is an underlying strand to this argument that should worry AUians, I do not think the argument alone can provide a conclusive reason for rejecting AU. This is why I shall not consider it, though it is also true that AUian replies can be developed to many of its facets, and that, for various reasons, not the least of which is the tendency to leave out in the delineation of cases features which the AUian would take into account, it can fail to persuade. Most importantly, a form of AU can be set out which avoids a good many of these well-documented conflicts with the 'plain man'.

A typical example around which the hard cases argument is

developed is this: Smith who is dying elicits from Jones, the only other person on their island, a promise to give all of his (Smith's) fortune to a jockey club, while Jones, off the island and back in society, now realizes that best consequences will obtain if the money is given to medical research. The anti_AUians argue that the 'plain man' would say that it would be wrong for Jones to break the promise, whereas the AUian is forced to say it would be right.[3] Exotic cases are not required to make the point: if Baker promises to repay Thomas, then the 'plain man', it is said, would think it wrong for Baker to give the money to a charity, even if utility were thereby maximized. Now one of the most important things such conflicts are said to bring out is a difference between the 'plain man' and the AUian as to the nature of at least some moral rules. The 'plain man', it is said, sometimes thinks it right to adhere to a rule, e.g., that promises are to be kept, even though utility would be maximized through breaking it; and this in turn is taken to show that he does not regard that rule as a rule of thumb, an empirical generalization or guide. What more the rule is, whether, for example, it is taken to have a hand in determining rightness, and how much of a hand it has, are matters on which the critic of AU may not wish to reveal his hand; but that the 'plain man' thinks not all moral rules are rules of thumb is certainly a commonplace in the anti_AUian literature. (It is

also the point of departure for the development of the different versions of RU.) Why is the 'plain man' wary of rules of thumb? The answer is clear: once moral rules are regarded as rules of thumb and once optimificity justifies the AUian in departing from them, the 'plain man' fears that far too many departures, and in cases where he would not approve of them, will be justified by the AUian. Clearly, this underlying fear of the 'plain man's' is not without foundation, since, traditionally, AUians have thought both that moral rules are rules of thumb and that optimificity justifies one in departing from them; and this fear persists, however the hard cases argument fares, precisely because these two elements in AU theories which give rise to it are never deleted by AUians in their replies to the hard cases argument. Nor can the first of these elements be dropped by an AUian; but the second can, and this provides the opening for a new form of AU, one which accepts that moral rules are, in a sense to be given the term, 'rules of thumb', as it must, if it is to remain a species of AU, but which does not incorporate optimificity as a justification for departing from them. This said, I must at once warn the reader that the expression 'rule of thumb' may be a highly misleading term for the kind of rules these are. In the following chapters, I shall reveal, gradually, certain characteristics of the kind of rules to be found in tempered act-utilitarianism; and, as we shall

see, these characteristics are such that, to describe these rules as _just_ or _merely_ 'rules of thumb' (or as _mere_ labor-saving devices) is seriously to misjudge their character. The reason I nevertheless employ the expression emerges below.

AUians of _all_ stripes _have_ to regard moral rules as 'rules of thumb'; the inclusion of _C_ in their theories sees to this. Nor should it be surprising that they offer no explicit argument in support of the view that moral rules _are_ 'rules of thumb': the adoption of a strong consequence component, _C_, renders _pointless_ any further argument on this score, since _C_ is incompatible with any view of moral rules that, most especially, allows them to have a hand in determining rightness. _This_ is why I call these rules 'rules of thumb': they do not make right or wrong the acts falling under them. If something in addition to consequences is allowed to determine rightness, the theory which includes this weak consequence component will not be a form of AU: it is the inclusion of _C_ that marks off AU theories as a type of utilitarianism and distinguishes them, e.g., from RU theories. Nor can the AUian, to accommodate the 'plain man', abandon the consequence component itself, since, as we saw in Chapter I, it is a fundamental part of the utilitarian theme. Thus, the AUian is in this position: it is the inclusion of _C_ in his theory that forces him to regard moral

rules as 'rules of thumb', but he can neither drop nor weaken C, and remain an AUian, nor abandon the consequence component itself, and remain a utilitarian. I conclude, therefore, that AUians of all stripes can accommodate moral rules in their theories only as 'rules of thumb'. However, as we shall see, it does not ensue upon this conclusion that moral rules have only a subsidary role to play in moral thinking or that they enjoy the status only of mere labor_saving devices; nor does it follow that AUians of all stripes have to accept optimificity as a justification _ let alone the justification _ for breaking them.

## VI.2 Impasse

The 'plain man' fears that the AUian will justify a number of cases of promise_breaking that he (the 'plain man') would not, and what critics of AU have sometimes gone on to assume is that this number will be of a magnitude to undermine the useful social inst_ itution of promising or the useful social rule that promises are to be kept. An argument is needed to support this assumption, and the hard cases argument is not equal to the task; it will not show that the number of acts justified is of the required magnitude, unless it is in turn assumed that any particular instance of AU promise_breaking will be staggeringly influential in getting others to break promises. For neither a single instance of AU promise_breaking nor a single individual breaking promises on many occasions will under_

mine the institution; and all the hard cases argument shows is that our AUian calls some acts of promise_breaking right or justified which the 'plain man' does not. What is needed, and what I will supply, against one form of AU (though the traditional form), is an argument that shows that the continuous and consistent application of AU will undermine the institution of promising and other useful social institutions. Such an argument is a powerful one, since the form of AU in question exhorts the maximization of utility, and yet its continuous and consistent application results in disutility. Equally important, the argument envisaged will not turn upon the success or failure of the hard cases argument.

Now the AUian can defend himself against the causal influence of his example. He can, of course, resort to certain ploys: he may only engage in AU promise_breaking in secret, or in those cases where the possibility of discovery is minimal; since single acts of secret AU promise_breaking are less detectable than a series, he will be careful to space out these acts; he may keep his AU theory to himself, using it privately to assess rightness or justify promise_breaking, while espousing a non_AUian theory in public; and he may find it has good consequences to praise publicly acts which privately he regards as wrong. But the most important defences the AUian can deploy against the causal influence of his example are those which emphasize the implausibility of assuming

enormous influence. After all, in a non_AUian society such as our own, it is extremely unlikely, because of training and the way people are brought up, that most people will suddenly desert their old habits and begin, with an example to guide them, to engage in AU promise_breaking; and, of course, even if we assume causal in_ fluence, it is unlikely that a case of AU promise_breaking will prove so catastrophic in its influence as to threaten the institution.

Such devices as these do protect the AUian against the causal influence of his example, but they neither still the 'plain man's' fear nor go any way towards alleviating it. Indeed, such devices are intended to preserve AU intact, so that the two elements which, together, give rise to this fear, are retained. The result is, as the literature so abundantly manifests, a kind of perpetual impasse or stand_off between the 'plain man' and the AUian; this so far as I can see is the only legacy of the hard cases argument.

## VI.3 The way forward

Now there is a way of breaking this impasse, a course by which an AUian can both still the fear of the 'plain man' and yet remain an AUian, though not, however, of the old variety. The attraction of such a course is obvious; it not only appears to arbitrate nicely the clash between 'plain man' and AUian but also to draw the alleged sting of the hard cases argument. I shall explore this way forward in the next chapter.

Then I shall return to the attack against the old form of AU, with an argument that <u>is</u> conclusive against this form (the new form escapes the argument completely), in the sense that it establishes that the <u>practice</u> of this form does indeed undermine the institution of promising, i.e., that it is productive of disutility. The ultimate answer, in short, to an AUian who fails to see the merit of any concession to the 'plain man's' fear is that his position collapses entirely in the face of our conclusive argument, and that the only hope of retaining a <u>practically viable</u> AU theory is to take the steps we are to take.

# VII

## TEMPERED ACT-UTILITARIANISM

While it is the inclusion of C within his theory that forces him to treat moral rules as 'rules of thumb', an AUian, as we have seen, can neither drop nor weaken C, and remain an AUian, nor abandon the consequence component itself, and remain a utilitarian. He can, however, supplement C with certain items his theory permits him to include and which have as their intended upshot the removal or diminution of conflicts with the 'plain man'. The point to such a course is that it enables an AUian both to make concessions to the 'plain man' and his worries and yet to retain a strong consequence component, a *sine qua non* of AU theories. I shall argue in this chapter that such a strategy of supplementation is indeed successful in soothing the fears of the 'plain man'; but the important aspect to the strategy for us is the identity of those considerations with which C is supplemented, since those given here mark off a new form of AU.

### VII.1 Philosophy and practice

Earlier, we distinguished between the use of imaginative cases when we are doing philosophy and when we are trying to find the best way, the best guides for doing the right thing.[1] When it comes to selecting those guides by which to act, then if we employ imaginative cases at all, they will be most helpful if they resemble the actual cases our guides will have to deal with. One is looking for the

best guides to doing the right thing, in those sorts of situations one finds or is likely to find oneself in, and the rules (or whatever) that will be of practical assistance in these cases are hardly likely to be those framed out of consideration of unrealistic cases. Now if an AUian, therefore, were concerned with the practical side of his theory, with its practical implementation and usefulness, and if he were to agree, for reasons to be given shortly, that we are most likely to do the right, i.e., optimific thing if we follow the commonplace rules of the 'plain man' as our guide, then he may ignore as not to the point cases that are unlikely to occur.[2] On the level of philosophy, his theory will have to cope with them, but on the level of practice, since they are unlikely to occur, he can safely stick to those cases that will occur and with which he will have to do. An example about a man in search of moral advice should help to make this distinction clear. Suppose a police officer has apprehended a drug_pusher whom he knows to be connected with central figures in the drug trade in this country and from whom, if only the man would talk, the names of these men may be had. The officer has begun to consider the possibility of torturing the man in order to obtain these names, and he comes to us for advice about what to do. (1). The officer does not want philosophy from us. It will be of limited _ if any _ assistance for us to help him

through the ins_and_outs of AU theories, or to point out the attractions of RU theories, or to discuss desert_island cases. Nor will it help him to be told that, on an AUian view of rules, every rule, including one proscribing torture, has exceptions; for what the officer wants to know is whether <u>this</u> is an exceptional case.[3] (2). If we do not know what <u>is</u> right in this situation, i.e., if, on AUian lines, we do not know what will be optimific in <u>this</u> situation, and I will shortly give reasons why we may not know this, then our advice to the man will be in the way of in_ dicating the best path for him to follow, not to <u>do</u> the right thing, for this we do not know, but to have the best <u>chance</u> of doing the right, i.e., optimific thing. We do not thereby tell the officer what <u>is</u> right; rather, we indicate to him, if we can, that course of action which we think will give him the best chance of doing the right thing. (3). If we raise with our officer this question, 'What is the most likely path for arriving at the right decision in this case?', and if we bear in mind that we may not know what <u>is</u> right, i.e., optimific, then is a departure from the commonplace rules of the 'plain man' to be advised in this case? If we are ignorant of what is optimific, then we may conclude that the officer will have the best <u>chance</u> of doing the right, i.e., optimific thing if he sticks to the rules. For the rules of the 'plain man' have long been recognized by AUians as reliable guides

to doing the right thing, and only to be broken for the sake of optimificity; so when an AUian is unclear about what course will be optimific, he has a good reason to advise adherence to these rules. Accordingly, the officer should forego torture to have the best chance of doing the right thing, and this is the advice we should give him, even on AUian grounds. Rules here are operating on the level of practice, and on this level there are AUian reasons for adhering to them.

VII.2 Ignorance, weakness and optimificity

What reasons has an AUian for thinking he is ignorant as to what will be optimific in this situation? In Chapter Five of Principia Ethica, Moore gives several reasons why an AUian will always be unclear as to what is optimific in the circumstances.[4] These are (i) the difficulty in assessing the total consequences, including those in the remote future, of any action, and (ii) the difficulty in calculating carefully, without resort to special pleading, the consequences of any action.[5] Special pleading, however, is not the only form of human 'weakness' an AUian has to contend with. Suppose a child is not yet able to recognize or an adult does not recognize that animals suffer pain when beaten: having a rule, reinforced with sanctions, proscribing cruelty to animals may serve to keep child and adult up to the mark. Here, a

rule aids those individuals who cannot or do not recognize the reasons for doing or not doing an act. Again, suppose we had reason to think that, in the absence of a rule proscribing cruelty to animals, people would indulge their aggressive impulses: here the reason for having and observing rules is to protect people against themselves and to try to get them to behave in ways they otherwise would not.[6] Or again, if I know from experience that I am tempted not to return books or tools I borrow, then having a rule that things borrowed are to be returned may, as an antidote to temptation, serve to keep me up to the mark. Having had such a rule ingrained in me, I will be less likely to be tempted; and if I frequently do not return books I borrow, then acting in accordance with this rule is probably a better guide to doing the right, i.e., optimific thing in a particular case than is breaking the rule. In fine, lack of knowledge and foresight of all the consequences of our acts, the danger of special pleading and these other forms of human 'weakness' render our situation murky vis-a-vis optimificity; accordingly, an AUian may conclude that we shall have the best chance of doing the right, i.e., optimific thing if we follow the rules of the 'plain man' as our guide.

These sorts of considerations make a strong case for having and observing the 'plain man's' rules and for inculcating them into our characters. Nor is there anything smacking of rule-worship

here: what our practical considerations make plain is that the demand for having and following these rules issues from the predicament in which we find ourselves, limited in knowledge, deprived of a God's eye view of things and plagued by various weaknesses. These reasons for having and observing also indicate why I am emphatic about not regarding the rules of the 'plain man' as merely 'rules of thumb: they are far too important, practically, to be set aside very easily.

In institutional cases, two additional reasons may be cited to strengthen the case for adhering to an institution's prescript_ ions. These are (1) that as a participating member of the instit_ ution one gives others to believe and to expect, strongly enough for them to be willing to make plans and to act upon them, that one will do one's share in maintaining the institution and its benefits; and (2) that as a participating member of the institution, one who enjoys its benefits, one gives others to believe and to expect, strongly enough for them to make plans and to act upon them, that one will consider the effect upon them if one pro_ poses an exception in this case to the institution's prescriptions. These provide the AUian with additional reasons for sticking to the institution; otherwise, one runs the risk of harming both the institution and those who participate in it, and this is hardly likely to be the best path to doing the right, i.e., optimific

thing. If one had to advise a man whether it would be right, i.e., optimific to break his promise in this particular case, then in addition to factors about the murkiness of our position vis-a-vis optimificity and factors about human 'weakness', there are further considerations - possible harm to the institution and its participants - which will also lead the AUian to advise on AUian grounds sticking to what the institution prescribes, viz., to keeping his promise.[7]

In sum, provided one realizes it is practical considerations with which C is being supplemented or tempered, an AUian can make a success of the strategy of supplementation. He retains his strong consequence component, but tempers it with such practical considerations that go a long way towards accommodating the views and worries of the 'plain man'. Such an AUian I shall call a tempered-AUian (hereafter abbreviated to TAUian), to distinguish him from the usual untempered-AUian (hereafter abbreviated to UAUian).

## VII.3 Rightness and justification

The traditional AUian position on justification has been that exceptions to rules and/or institutions are justified, and only justified, on grounds of optimificity. Thus, their position with respect to justifying cases of promise-breaking may be broken down as follows. Either a rule that promises are to be kept has

exceptions or it does. If one says it does not, then one is at odds with the 'plain man' who thinks that it does; if one says it does have justified exceptions, then one requires some means by which to distinguish the justified from the unjustified ones. The only procedure to adopt to determine whether a case is an exceptional one is, on this view, to consider each case 'on its own merits', i.e., in the light of its circumstances and consequences. This commitment to what I shall call a case by case procedure is simply a restatement of the AUian view of rightness: one is justified in breaking this promise if it is right to break it, where rightness is determined by optimificity; and if it is right to break this promise, it is ipso facto a case of a justified exception to a promise_keeping rule. Put succintly, then, the position is that, if breaking this promise is optimific, it is right to break it and one is ipso facto justified in breaking it. Now it is to this thesis of justifi_ cation, a feature of all untempered forms of AU, that many of the usual counterexamples to AU are directed; they are intended to show that the 'plain man' does not think optimificity a justification _ let alone the justification _ for promise_breaking. Thus, it may be optimific for Smith not to return money to Jones, but if he has promised to return it, then the 'plain man', it is said, does not think optimificity justifies doing something else with it. Moreover,

a UAUian will be in perpetual conflict with the 'plain man' on this score, for the employment of a case by case procedure, i.e., examining each case on its own merits and deciding to keep or break each promise solely on the ground of optimificity, is precisely what the 'plain man' desires to exclude.[8]

At this point, the TAUian may draw an important distinction between what is right and what we are justified in doing in the circumstances. Suppose an AUian were to retain the link between rightness and optimificity but to sever the link between justification of exceptions and optimificity. That is, he retains a strong consequence component but accepts the criticism that optimificity is no justification for breaking promises or taking an exception to the institution. A case of promise-breaking will still be right if optimific, but one will no longer be _ipso facto_ justified in breaking this promise; for our situation is murky _vis-a-vis_ what _is_ right, i.e., optimific, and we stand the best chance of doing the optimific thing if we stick to the rules of the 'plain man'. On this view, an act can be justified but wrong: we may be justified in sticking to the rules for those practical reasons given in the previous section, even though in the final analysis or from a God's eye point of view adherence to the rules in this case is not optimific. A TAUian, then, as in the example of our police officer, thinks he is justified in adhering to the rules of

the 'plain man' and advising others to do so, since his and our positions vis a vis optimificity are unclear, even though such a course turns out in the final analysis to be wrong, i.e., not optimific.

Since our positions are murky vis a vis optimificity, we look for the best path to follow, not to do the right thing, but to have the best chance of doing the right, i.e., optimific thing. AUians have recognized the rules of the 'plain man' as highly reliable guides to doing the right thing, and even to UAUians, they are only to be broken for the sake of optimificity; when we are unclear what course will be optimific, there is good reason for the AUian to adhere to these rules. TAU, then, is tempered as a result of taking into account those practical considerations that, together, weigh heavily in favor of adherence to the rules and institutions of the 'plain man'. Because TAU retains a strong consequence component, as it must, to remain a species of AU, these rules of the 'plain man' will be still regarded as 'rules of thumb', since the acts falling under them will not be made wrong by the rules; however, on the strength of a distinction between what is right and what is justified in the circumstances, it does not follow from the fact that such rules are 'rules of thumb' that optimificity is a general justification for breaking them. The link between justification and optimificity is rejected by the TAUian; and he has made substantial moves towards

calming the 'plain man's' fears. For if the reason the 'plain man' is wary of treating moral rules as 'rules of thumb' is that he does not think such rules are to be set aside whenever best consequences obtain through doing so, then TAU caters to this view; for though such rules are still regarded as 'rules of thumb', for the reason just given, optimificity is no longer held to be a general justification for breaking them _ indeed, the TAUian can even advocate <u>on AUian grounds</u> positive reasons for adhering to them in the face of uncertain optimificity, as we have seen. I think this form of AU, in short, can come to terms with the 'plain man', and, moreover, is the only form of AU that can; for it is the only form of AU that rejects the case by case procedure, so noticeably a part of other AU theories, and with it optimificity as a thesis of justification. As we shall see, when we turn to our utilitarian argument against UAU, there is a very potent reason for rejecting these items, apart from any desire to come to terms with the 'plain man'.

Instead of saying that TAU is tempered by certain practical considerations which support sticking to the rules and institutions of the 'plain man', we could have said, for example, that TAU is tempered through the introduction of rules. Two points here are absolutely vital. First, as is clear, in TAU rules are operating

not on the level of theory but on the level of practice: they do not make right acts right, and they are introduced on AUian grounds, as the best guides to follow to have the best chance of doing the right or optimific thing. For this reason they are immensely important, practically. Second, on this level of practice, there is no transition from AU to RU, for TAU retains a strong consequence component and so the link between rightness and optimificity. If it be objected that any introduction of rules into AU, even on the level of practice, is a transition to RU, the conclusion must be that some forms of AU and RU are consistent. For both will prescribe the same course of action _ sticking to the rules _ in the face of uncertain optimificity, and both may agree on whether optimificity justifies exceptions to the 'plain man's' rules. So if those forms of AU which incorporate rules on the practical level are held to be forms of RU, then these forms of AU and RU seem indistinguishable in practice or 'extensionally equivalent'.[9]

## VII.4 Essential rules

In Principia Ethica, Moore claims that some rules are so essential to the preservation of society that they ought always to be adhered to, even if on some occasions harmful results obtain.[10] Moore's reasons for saying an AUian is never justified in breaking these essential rules are, as we have seen, lack of knowledge of the consequences of acts and the danger of special pleading. I do not

agree with Moore that we are never justified in breaking essential rules: to hold that there are no justified exceptions to the rule that promises are to be kept is as objectionable to the 'plain man' (and to a Cist, and so an AUian, for reasons I gave in earlier chapters) as is the UAUian reliance upon optimificity to justify promise-breaking. However, if the TAUian rejects Moore's counter-intuitive view that no exceptions are justified and rejects as well the UAUian's reliance upon optimificity to justify exceptions, then he stands in need of some means of distinguishing justified from unjustified claims to an exception. To this problem, we shall return in a later chapter.

Moore argues that we ought always to adhere to essential rules, not that we ought always to adhere to all rules or all moral rules; but the essential/non-essential distinction is not a perspicuous one, for reasons which I shall not go into here, but which obtrude in the Hart-Devlin controversy and Basil Mitchell's discussion of it, in his Law, Morality and Religion.[11] I think the concept of an 'essential' rule is a flabby one, and I shall proceed by dropping it altogether. Nothing is lost to our discussion by so doing. We shall still be concerned with what justifies a person in breaking a promise; only we shall cease to speculate whether a rule proscribing such acts is essential to the preservation of society. Then, too, TAU is unaffected by dropping the concept, for the practical

considerations in favor of sticking to commonplace moral rules (and institutions) do not depend for their force upon whether the particular rule in question is an essential one or not. Nor is any important problem brushed aside by dropping the concept: given that we stick to the rules because doing so gives us the best chance of doing the right or optimific thing, yet given also that the rules have exceptions, we shall still be confronted with the problem of trying to tell the justified from the unjustified claims to an exception.

VII.5 Moral rules: some doubts

Since the appearance of Hart's The Concept of Law, with its comparisons of and contrasts between legal and moral rules[12], it is increasingly common for philosophers to point to characteristics which legal rules or park regulations or library rules exhibit which moral rules do not, and to conclude either at worst that there are no moral rules or at best that, if there are, then they are rules only in some derivative sense of that term - the central sense being defined by those characteristics.[13] The following is a representative sample of these alleged characteristics of the central cases of rules.

(1). There is no book of moral rules as there is a book of rules for games, parks, libraries, universities, etc. Library rules are written down and may be checked if we are in doubt whether X is

a rule of this library; moral rules have not been written down, and cannot be checked if we are in doubt whether X is a moral rule of this society. (2). In these central cases, rules amount to regulations: the rules of tennis regulate that activity[14], just as the rules for checking out books regulate that activity. (3). With parks, games and universities, there are people whose business is to know the rules currently operative and to enforce them; with moral rules, even if we agree that it is the business of everyone to know the rules and enforce them, there is still no one person or body whose pronouncements are more authoritative than another's or whose pronouncements can tell us whether X is a moral rule of this society. (4). Library rules are promulgated and from time to time altered or rescinded; this is done by an individual or body in authority empowered under the rules to do so. Nothing similar exists in the case of moral rules. (5). The rules of libraries, games, parks, zoos, etc., are made whereas moral rules are not, and no individual or body is empowered to make _ or to unmake _ them. In fine, on the strength of characteristics such as these, one might be led to claim that moral rules are rules only in a derivative sense of that term.

From the point of view of a discussion of utilitarianisms, this claim calls for three comments. First, and least important, we can quibble with some of these characteristics, especially with two of

them, viz., that rules are promulgated and capable of being re-scinded, and that rules are made and capable of being unmade. (a). Rules of etiquette have not been promulgated, nor is there a body who can make or unmake them. They change without any formal procedure for their being changed, e.g., they do not include a rule specifying how all the other rules may be changed. On the view in question, either the rules of etiquette are not really rules or else the term 'rule' is being used derivatively. What is gained by saying either of these things? (b). The rules of ordinary English grammar are not made, nor have any of them been promulgated or re-scinded; there is no body which by its pronouncements can change them. Are these rules not really rules? (c). A number of people treat the teachings of Confucius as rules by which to live, which, in a sense, have been promulgated or, in any event, made, but which cannot be rescinded or unmade. What is gained, therefore, by insisting that these followers live not by rules but by something else, e.g., maxims? Second, these questions about whether we should call this or that a rule or something else are largely irrelevant in a discussion of utilitarianisms. For whatever 'Promises are to be kept' is, whether a rule, principle or maxim, a UAUian will argue that it has justified exceptions and that these are justified solely on the ground of optimificity; a TAUian will argue that it has exceptions but agree with the 'plain man' that optimificity is

no general justification for breaking it; and an RUian will turn to formulating his theory in terms of, e.g., maxims instead of rules. Of course, on some theory or other important differences may be drawn between rules, principles and maxims; nor do I deny that it would be useful to establish clear distinctions among them, and to observe such distinctions carefully. All I am suggesting is that, for our purposes, it is unimportant whether 'Promises are to be kept' be called a rule or something else, since nothing of substance turns upon the outcome. Similarly, philosophers speak of the informal rules governing the uses of words; yet clearly, these rules are not promulgated and rescinded, made and unmade. To one who says that whatever it is that governs the uses of words, it cannot be rules, we shall want to object that precisely what is of interest is <u>what does</u> govern the uses of word, and whether these are called rules or something else is of little importance. Lastly, it is frequently overlooked that the controversies surrounding AU and its strong consequence component <u>need not</u> be conducted in terms of <u>moral</u> rules at all. Let us agree on an unambiguous 'central' case of a rule, e.g., the rule of the zoo that 'animals are not to be fed': whether optimificity justifies breaking this rule, whether it <u>makes</u> wrong acts falling under it, etc., may all be discussed in terms of this 'central' case. The same will apply in the case of legal rules, as Wasserstrom, in <u>The Judicial Decision</u>, has recently

shown[15]; and there is in principle no reason to think that any of the 'central' cases of rules may not be used to discuss the issues that concern us here.

---

I am inclined to think that the proponents of any move that minimizes the importance of what I have been calling 'rules of thumb' fail to appreciate the lack of knowledge and weaknesses that characterize all of us. Were we different from what we are, perhaps such rules could be dispensed with; but in our present condition, they are necessary. Whether one calls them 'rules' or even 'rules of thumb' or something else is unimportant so long as the reasons for having and observing them remain to the fore and so long as their characteristics are marked.

## VIII

## THE DISUTILITY OF UNTEMPERED ACT UTILITARIANISM

The TAUian, by rejecting the UAUian link between optimificity and justification of exceptions and by exploiting the fundamental distinction between what is right and what is justified in the circumstances, can on AUian grounds advocate adherence to the useful social rules and institutions of the 'plain man'. Since we do not know what is right or optimific, we have to decide what action will give us the best chance of doing the right or optimific thing; this estimate of probability is made by the TAUian in favor of the rules and institutions of the 'plain man', because he judges that adhering to them gives him the best chance of doing the optimific thing; and his judgment is justified even though it turns out to have been wrong or not optimific. The TAUian and the 'plain man', then, can on this score, I think, bridge their differences. But UAUians, on the other hand, because they employ a case by case procedure and decide to keep or break each promise solely on the ground of optimificity, will be in perpetual conflict with the 'plain man'.

Now I think a stronger argument can be brought against UAU, a utilitarian argument, to the effect that acting upon UAU will undermine useful social institutions. It will be my aim in this chapter to show that the justification of exceptions to useful social institutions on the ground of optimificity will undermine them, and

so be productive of gross disutility. First, however, I shall distinguish my utilitarian argument against UAU from the utilitarian argument brought against all forms of AU by D. H. Hodgson, in his Consequences of Utilitarianism.[1] Very roughly, the difference between the two arguments is this, that whereas Hodgson argues that in theory on AUian assumptions social institutions such as promising and truth-telling could not get started, I shall argue that in practice justifying exceptions to these institutions on the ground of optimificity will erode them. Since TAUians do not accept optimificity as a thesis of justification, my argument will be conclusive against UAU only.

## VIII.1 Hodgson's argument[2]

Hodgson's argument is intended to show that, even if on each occasion the AUian succeeded in doing that act which maximized utility, he would in the long run nevertheless bring about worse consequences than if he acted on some other, non-AUian (or non-utilitarian) theory; the upshot of which is that AU is self-defeating.

The essence of Hodgson's argument is as follows: certain good consequences depend upon the existence of expectations of actions, and, given certain conditions of knowledge and rationality, acting upon the 'act-utilitarian principle'[3] could preclude such

expectations and thereby such good consequences. This is the case with promising. A promised act would not have better consequences, better than it would have if it had not been promised but, e.g., only mentioned as intended, unless it were expected more, more than it would be if it had only been mentioned as intended. However, there would be no good reason for such increased expectations, on the assumption that the promisor is an AUian, unless the promised act had greater utility than any alternative; but the promised act would have greater utility only if it could be expected to occur. Thus, a spiral or regress is set in motion: a promised act will have best consequences only if expected more; it will be expected more only if it has best consequences; but it will have best consequences only if it is expected more; and so on. Once in this spiral or regress, vicious in kind, there is no escape for the AUian; hence, since the whole point of the 'act_utilitarian principle' is that the best possible consequences are to be promoted, but since the adoption of that principle precludes the promotion of certain good consequences, AU is self_defeating.

A successful refutation of this argument has been given by Peter Singer, in his paper 'Is Act_Utilitarianism Self_Defeating?'[4], to which I have very little to add. But what I do have, though it supports Singer, also supports TAU in the face of Hodgson's thrust.

Singer exposes the fact that all that is required to cut this vicious spiral and to set off a virtuous one is some initial presumption that the AUian will not make false promises; for then the promisee will to that extent expect the promise to be kept, which yields some utility to be gained in keeping it, which yields an AUian ground for keeping it. Singer goes on to set out and argue for an initial presumption of this sort. Now the way I would establish this presumption is by going through the arguments of the preceding chapter; for if there are AUian reasons for adhering to the rules and institutions of the 'plain man', then the promisee has strong grounds on which to expect the promise to be kept, and the virtuous spiral Singer wants to set in motion gets underway quite easily. Moreover, established in this way, the initial presumption that a TAUian will keep his promise is strong (or high), since he rejects optimificity as a thesis of justification and is cautious about taking or advising exceptions to useful social institutions and rules. To be explicit, the initial presumption that a TAUian will keep his promise is of the same strength as the initial presumption that a 'plain man' will keep his, since the rules and institutions adhered to by the TAUian are the 'plain man's'. In short, the TAUian can meet Hodgson's argument.

One other aspect of Hodgson's argument wants comment. For it

to get off the ground, certain conditions must be granted, which, in addition to certain factors of knowledge and rationality, include assumptions that form part of what Hodgson calls the 'act-utilitarian's ideal'.[5] An important part of this 'ideal', according to Hodgson, are the assumptions that 'there are no conventional moral rules, and everyone's only personal rule is the act-utilitarian principle'.[6] These assumptions are certainly not 'ideal' from a TAUian point of view, nor are they assumptions he will grant, for reasons already given; nor does Hodgson provide an argument to show that all AUians must treat them as either desirable or 'ideal'. The reason he fails to provide such an argument is that he overlooks the distinction between UAU and TAU, an oversight the explanation for which is to be found in his belief that acknowledgment of, reliance upon and acting in accordance with conventional moral rules is not a real possibility for an AUian. Here, Hodgson fails to take into account the level on which a TAUian relies upon conventional moral rules; this is the level of practice, on which there appears to be, as we have seen, no reason why on AUian grounds the TAUian cannot consistently advocate action in accordance with these rules. And, for reasons already gone into, the TAUian's decision to adhere to these rules is justified even if it turns out to have been wrong. Such distinct-ions are vital, and failure to make them ensures that all forms of

AU will be lumped together, so that what may be 'ideal' for one form is construed as 'ideal' for all.

Unlike Hodgson's, my utilitarian argument against UAU only is an argument from practice, to the effect that acting upon UAU will undermine useful social institutions. If it is correct, UAU is in deep trouble: it exhorts us to maximize utility, but its consistent application produces disutility.

## VIII.2 A preliminary to the argument

One feature of UAU that critics will want to look into is this: traditionally, UAUians have been willing to see all of us become like themselves; but if enough of us, e.g., some large majority of the population, begin to engage in UAU promise-breaking, i.e., in breaking promises on the ground of optimificity, then this useful social institution will collapse, with consequent disutility. Now the escapes one might expect UAUians to adopt at this point in fact, at least in print, have few adherents. For example, a UAUian might argue that he can conceive an institution of promising of even greater utility than the 'plain man's' and that, therefore, for the sake of this other institution, he is prepared to see the 'plain man's' undermined; or he might argue that only he is entitled to an exception to the useful social institution of promising on the ground of optimificity (a view which would seem to involve a breach of universalizability). Neither of these lines, however, find many

adherents in the literature; instead, by far the most common UAUian tactic has been to argue that, e.g., because of their good habits and moral upbringing, a UAUian can be practically certain that others <u>will not</u> engage in UAU promise_breaking. What seems to have been overlooked about this reply is, first, that it indicates that a UAUian is unwilling to undermine the social institution of promising, but second, and more importantly, that it does not really assist the UAUian on all fronts. For it assumes a certain situation to be at hand, viz., that there are a few UAUians living in a non_UAUian society and that, because of the other people's good habits and upbringing, a UAUian can be practically certain that these others will not engage in UAU promise_breaking. However, the fact that UAUians are (or have been) willing to see us all become UAUians suggests at least one other initial situation to be considered, viz., one in which more and more people are becoming UAUians and so beginning to engage in UAU promise_breaking. In <u>this</u> situation, reliance upon other people's good habits is mis_placed, since at least more and more of them are breaking whatever good habits they have in respect of promise_keeping. Thus, if we suppose that a UAUian both allows and advocates that all of us become UAUians, then he is allowing and advocating the collapse of the institution of promising; for the more people succumb to his advocacy the greater the risk the institution will be harmed

and confidence in it undermined, as UAU promise_breaking spreads. Paradoxically, then, though the UAUian advocates the adoption and practice of his theory by fellow members of his society, he must hope to be relatively unsuccessful in his advocacy, at least to the extent that the vast majority remain unconverted. Of course, he may not advocate UAU or try to persuade others to become UAUians, and this we note; nevertheless, traditionally, UAUians have (more minimally) at least been in favor of us becoming like themselves, and this licenses us to consider a situation in which more and more people are engaging in UAU promise_breaking.

Confronted by a situation in which a number of people are breaking promises on UAUian grounds, though as yet not in numbers sufficient to undermine the institution, the UAUian may argue that it is optimific to engage in UAU promise_breaking only so long as some number sufficient to undermine the institution does not do so, just as it may be optimific for a few to walk on the grass provided the vast majority refrain. Thus, it ceases to be optimific, and therefore right, to engage in UAU promise_breaking when this number sufficient to undermine the institution is approached; at this point, the UAUian, on the ground of optimificity, refrains and, we may suppose, urges that others refrain. Though perhaps a strange argu_ment to have from one who is in favor of all of us becoming UAUians, it is for all that consistent with UAU. Nevertheless, it is exposed

to a weighty objection.

If its collapse is to be avoided, the UAUian must know when the institution is threatened; that is, he must know the point at which it ceases to be optimific to engage in UAU promise_breaking; and to know this he needs to know when the number of people engaging in and the number of acts of UAU promise_breaking are sufficient to _ or are approaching being sufficient to _ undermine the institution. Put summarily, he must know when the critical point is or is about to be reached. There are two problems here for the UAUian, both concerned with the state of his knowledge. First, to know when the critical point is or is about to be reached and so when it ceases to be optimific for him and others to engage in UAU promise_breaking the UAUian must command a clear view of what is optimific in any situation in which he contemplates a UAUian exception to promise_keeping, which in turn means that he must know reasonably accurately how many and how often others are engaging in UAU promise_breaking. Now I have argued that our situation vis_à_vis optimificity is murky; but even if one is inclined to think our situation not as murky as I have made out, one may nevertheless be unwilling to allow that it is so clear that the UAUian can, in each instance when he is contemplating UAU promise_breaking, ascertain whether the critical point is being approached. At the very least, it will not be easy to find out how many and how often others are behaving

similarly; and for UAUians to engage in UAU promise_breaking without this information places the institution at risk, so that it is mere good fortune that it escapes, if it does, unscathed. Second, in speaking of a critical point (I take up this notion later), we must not we deceived into thinking that in all instances the critical point is a point, i.e., is such that a single act at the critical point can be sufficient to bring on harmful consequences. A case where a single act at the critical point suffices to bring on harmful consequences is not difficult to imagine; e.g., if it requires two individuals dumping garbage into a stream for the level of contamination to reach proportions dangerous for human consumption, then the dumping at the critical point by the second man is sufficient to attain this dangerous level. On the other hand, in the case, e.g., of promising, there is no reason to believe or to suppose that a single act of UAU promise_breaking at the critical point is sufficient toundermine confidence in the institution; a series of such acts at the critical point may be required, so that instead of a critical point there is a critical region.[7] However, the fact that a series of acts of UAU promise_ breaking in the critical region are (or may be) required to bring on harmful consequences adds to the UAUian's burdens. For in order to determine whether it is optimific to take a UAU exception to the institution and its prescriptions, he must ascertain whether the

critical region is or is about to be reached; in order to ascertain this he must not only know how many and how often others are engaging in UAU promise_breaking; but he must also know whether the series of acts required at the critical region to bring on harmful conse_ quences has commenced and has progressed to its final stages. Since the series of acts will not be performed by one but rather by many agents, it seems unlikely the UAUian can have the knowledge in question and can perform the fine estimate demanded of him. Once more, to proceed without this knowledge places the institution at risk, so that it is mere good fortune if it escapes unscathed; but, more importantly, without this knowledge of the critical region and without the knowledge of how many and how often others are engaging in UAU promise_breaking, the UAUian will not <u>know</u> whether it is optimific to engage in UAU promise_breaking <u>in this case</u> (or whether it is optimific to disengage from UAU promise_breaking altogether). Of course, I do not say that it is impossible for the UAUian to know these things, only that it is implausible _ on anything less than a God's eye view _ to assume he knows them on any single occasion, let alone on all occasions, when he contemplates UAU promise_breaking. So, without these pieces of knowledge, he will not know in a particular case whether it is optimific to engage in or optimific to cease engaging in UAU promise_breaking .

## VIII.3 A utilitarian argument

In the light of the foregoing preliminaries, we can construct a utilitarian argument against UAU, which consists in showing that the general practice of UAU will undermine the useful social institution of promising and, by implication, other useful social institutions. I begin by setting out certain ingredients to the argument, and conclude by presenting the argument.

(1). It has long been acknowledged, even by UAUians, that the 'plain man's' institution of promising is of a high degree of utility. This acknowledgment is completely independent of any question of whether one can conceive another institution that is even more useful than the 'plain man's'; whether or not one can, the 'plain man's' institution possesses a utility usually considered immense, for reasons of fostering trust, facilitating interpersonal cooperation, predicting behavior, coordinating plans, etc. Moreover, the acknowledgment of utility is independent of any question of whether firm statements of intention may not suffice to bring about, e.g., the same degree of coordination of plans as a promise; even if they may (which, of course, may be doubted), it remains true that the coordination of plans is an important factor in the utility of the 'plain man's' institution of promising. Given, then, such utility, it follows that harm to the institution will involve a certain disutility, which may be and very likely will be considerable if the

institution is undermined.

(2). A single case of UAU promise-breaking may have negligible effect on the institution, though, as we have seen, a UAUian is probably not in a position to know that it will; but he can be certain that the practice of UAU promise-breaking by a sufficiently large number of individuals will have a damaging effect on the institution, i.e., will impair, if not undermine, its utility.[8]

(3). In his Generalization in Ethics, Marcus Singer distinguishes between individual and collective consequences.[9] According to Singer, the collective consequences of everyone's acting in a certain way is for certain kinds of acts not the sum - or, more exactly, is greater than the sum of - the individual consequences of each individual act. The point is put more straightforwardly by Sir Roy Harrod:

> There are certain acts when performed on n similar occasions have consequences more than n times as great as those resulting from one performance. ...For example, it may well happen that the loss due to a million lies uttered within certain limits of time and space is much more than a million times as great as the loss due to any one in particular. 10

Another way of putting Harrod's point is that the harm done by a number of acts of a certain kind is different from - or, more exactly, is greater than - the sum of the harms done by each act singly; and still another way is that there may be consequences as the result of

the general performance of a certain kind of act which are not the result of individual performances of it. Two much_discussed examples in this connection are voting and lawn_crossing. If one man abstains from voting for a popular candidate, it is unlikely that his vote will be decisive in causing that candidate's defeat; but if a large number abstain, then that candidate's defeat will be ensured. Likewise, if one man nips across the grass, he is unlikely to damage the lawn; but if some large number walks on the grass, the lawn will be damaged. So far as I can see, there is no reason why promising should not be included among these examples. Thus, negatively, if everyone engages in UAU promise_breaking, the institution will be undermined, whereas if only one man takes UAU exceptions, his act will not have this consequence. Or, positively, the valuable consequences associated with the institution are the result not of only one many keeping his promises but of the vast majority keeping their's, so that a great many acts of UAU promise_breaking will prevent these valuable consequences arising. This makes it appear, then, that the general practice[11] of UAU promise_breaking has consequences which do not occur through the individual performance of that kind of act. Since the collective consequences here are not the sum of the individual consequences of each act, the effect of an aggregate of acts of UAU promise_breaking <u>is not</u> a linear function of the effect of a single such act. This

view David Lyons calls the thesis of non-linearity between the consequences of individual acts of a certain kind and the consequences of the general practice of acts of that kind.[12]

(4). For collective consequences to occur, universal conformity or non-conformity is not necessary; general conformity or non-conformity will suffice. Thus, for the harmful collective consequences of damage to the grass and thereby to people's enjoyment of a beautiful lawn to occur, it is not necessary that walking on the grass be universal; they obtain if there is general non-conformity with the prescription to keep off the grass.

(5). How general non-conformity must be for harmful consequences to ensue will depend upon the kind of act under consideration. The institution of promising or the appearance of the lawn will be threatened only when promise-breaking or lawn-crossing reaches a certain critical point or region, which, to recall, is that point or region beyond which, because others are behaving similarly, harmful collective consequences ensue. Thus, if a great many people begin to walk across the grass, the lawn will be damaged, and the critical region will be that range of acts upon the performance of which harmful collective consequences ensue, i.e., the lawn is damaged. The concept of a critical point is what Lyons has called a threshold and the harmful or beneficial collective consequences that ensue with the threshold is crossed are by him called threshold

effects:

> For many kinds of act there are thresholds which must be passed before effects of certain kinds or of certain magnitudes can be produced. For many kinds of acts, in other words, the curves describing the relation between total effect and the number of performances are ir_regular or at least not straight lines. ...In the lawn_crossing case, for example, the threshold may be indicated roughly as that point at which, if crossing becomes more frequen , the lawn would be damaged. This damage (which is no longer negligible) may be called a threshold effect. 13

(6). Applied to promising, we may say that the collapse of confidence in the 'plain man's' institution is a threshold effect, the result of many people on a number of occasions engaging in UAU promise_breaking. Accordingly, if a series of acts of UAU promise_breaking or its general practice may have consequences which would not accrue through any single act not a part of the series or general practice, the threshold will be that point in the series beyond which, through additional acts of AU promise_breaking, confidence in the institution collapses.

(7). A UAUian employs a case by case procedure and retains optimificity as his thesis of justification; that is, he examines each case on its own merits, assessing its consequences and deciding, e.g., whether to keep or break each promise solely on the ground of

optimificity.

With these ingredients, we may construct our argument. Given the utility of the 'plain man's' institution of promising, one act of UAU promise-breaking may not be disastrous, but a large number of such acts would carry us over the threshold, and harmful threshold effects, i.e., the collapse of confidence in the institution, would ensue. On the other hand, so long as the threshold remains uncrossed, acts of UAU promise-breaking will have individual consequences at least as good as those of any alternative and will not have the harmful collective consequences of damaging the institution. Therefore, if the UAUian employs his case by case procedure, deciding whether to keep or break each promise solely on the ground of optimificity, then of no instance of UAU promise-breaking considered singly will it be true to say that it is the cause of damage to the institution; hence, of no particular instance will it be true to say that it has the harmful collective consequences of damaging the institution. It follows, therefore, that in any particular case of UAU promise-breaking the individual consequences will be the determining factor for the UAUian in deciding to keep or break the promise; and employing his case by case procedure he will determine to break the promise, since, with harmful collective consequences not being ascribed to his act, it will be optimific to do so. However, if the practice of UAU

became general, so that the above chain of reasoning became general, it would follow that the number of individual acts of UAU promise-breaking justified would be sufficient to cross the threshold and produce harmful threshold effects, viz., undermining (confidence in) the institution. Thus, the general practice of UAU will undermine the useful social institution of promising, i.e., the general practice of UAU produces disutility, which amounts on the practical level to a utilitarian argument against UAU. Moreover, unless some difference can be established between promising and other social institutions, so that our argument is inapplicable in respect of these latter, the general practice of UAU will be productive of disutility in these cases too.

Lastly, if the UAUian advocates the general adoption and practice of UAU by fellow members of our society, he will be seen in the light of this argument to be unwittingly advocating the production of disutility. More than ever, then, he must hope to be unsuccessful in his advocacy, to the extent at least that a sufficiently large number of people - sufficient to preserve the institution from impaired utility - forego the adoption and/or practice of UAU.

# IX

## A DEFENCE TO THE UTILITARIAN ARGUMENT

There is a defence to our utilitarian argument that, gleaned from David Lyons' Forms and Limits of Utilitarianism[1] and put in the mouth of the UAUian, may be thought _ mistakenly, I hold _ to challenge our argument fundamentally. I propose to consider this single defence to the exclusion of all others, partly because of its importance, which stems from the fact that several of its steps figure in Lyons' well_known argument for extensional equi_ valence between forms of GU and/or RU and AU; and partly because other defences appear not to pose a serious or fundamental challenge. For example, to the objection that there are not thresholds for all kinds of acts, we may retort that our interest is in those for which there are, viz., institutional cases such as promising.[2] To the objection that it is difficult to determine the threshold of a kind of act, we may reply, first, that whether difficult to determine or not, we know that the general practice of walking on the grass damages the lawn, and it is the general practice of UAU that conerns us; and second, that ignorance of the threshold may equally well serve _ and, one may think, should serve _ as a reason for refraining from walking on the grass as a reason for walking on it _ indeed, this is the point of TAU. It will not be fruitful to consider such defences as these.

I shall instead focus upon Lyons' argument for extensional

equivalence and therein upon the thesis of non_linearity. For a UAUian might argue as follows: the utilitarian argument of the previous chapter only works on the assumption that non_linearity holds, whereas Lyons has demonstrated that non_linearity does not hold between the consequences of a single act and the collective consequences of a series of such acts; so the utilitarian argument does not succeed.

## IX.1 UAU and Lyons' argument

Lyons' argument in support of extensional equivalence is, despite appearances, really quite simple. (1). The appearance of non_linearity between the consequences of a single act and the collective consequences of a series or general practice of such acts is due to the fact that the description under which the act is being referred is not completely specified. In particular, it fails to include a significant causal property of the act, viz., the extent to which others are performing acts similar to the one in question.[3] Thus, Lyons argues that in the voting case what effect my voting or abstaining has depends upon what others are doing; and because the acts or behavior of other people determine what effect mine will have, this behavior is a causal property of my act and, consequently, should be included in its description. (2). Non_linearity is thought to obtain, therefore, because of failure to include in the

description of the act how other people are behaving; once this is done, the appearance of non-linearity vanishes. Thus, Lyons thinks the only way a supposed difference between the effects of one act of lawn-crossing and the effects of many such acts may be non-linear is to leave out of the description of the single act the extent to which others are behaving in the same or a similar fashion.[4] (3). The inclusion in the description of a single act of lawn-crossing of how others are behaving eliminates the appearance of non-linearity through taking into account the threshold phenomenon.[5] This means that in the description of a single act of lawn-crossing will be included some indication of whether or not harmful threshold effects ensue as a consequence of it; once threshold effects are included in the description of the act, it becomes clear that the act 'contributes' to the damage to the lawn, if damage should result from a number of such acts.[6] Thus, when an act is described completely, to include its social context, i.e., threshold effects and thereby the extent to which others are behaving similarly, it will include within its description the fact that it does or does not contribute to the damage to the lawn. In this way, the appearance of non-linearity is eliminated; we are left simply with a description of a single act of lawn-crossing which includes within it an indication of its contribution to threshold effects. Consequently, for Lyons, the harm done by a number of acts of

lawn_crossing <u>is</u> a linear function of a single such act: 'If actions are viewed as completely relevantly described, the generalized simple utility remains linear'.[7] (4). Equivalence is shown by establishing that, once acts are completely specified in respect of their causal properties, linearity obtains between the harmful or beneficial effects of a single act and the harmful or beneficial effects of a series or general practice of such acts. Hence, when a single act of lawn_crossing has been completely specified in respect of its causal properties, including the extent to which others are behaving similarly, so that possible threshold effects may be included, assessment of its consequences will yield a result on all fours with the assessment of the consequences of its general or universal practice, i.e., of what would happen if everyone were to behave in similar fashion. (5). Lyons thinks he is entitled to include in the description of a single act of lawn_crossing the extent to which others are behaving similarly because he thinks that in a utilitarianism we are concerned with all the consequential aspects of acts, i.e., with all those properties of an act that cause consequences. Thus, because the causal efficacy of my act of lawn_crossing is determined by how others act, Lyons thinks he is entitled to include the behavior of others as a causal property of my act and, thereby, to make this property of importance for utilitarians. As

Lyons claims, '<u>general utilitarian properties</u> - those properties that are relevant for the application of utilitarian generalization - are causal properties in virtue of which the universal performance of acts of that kind would produce some utility or disutility'.[8] For the purposes of utilitarian generalization, then, an act's causal properties assume importance; but the inclusion of all its causal properties in one's description of the act will include the contribution of the act to threshold effects; and when this is done the appearance of non_linearity between the effects of a single act and the effects of a series of such acts, which, after all, is the basis of utilitarian generalization, is destroyed. It is, therefore, upon establishing linearity and uprooting non_linearity that Lyons rests his case for the extensional equivalence of most forms of GU and AU: since the individual act of lawn_crossing, completely specified in its causal properties, will include within its des_ cription its contribution to threshold effects, to ask what would happen if everyone were to engage in lawn_crossing reveals nothing more than is revealed through looking at the effects of a single act of lawn_crossing. (6). Lastly, it should perhaps be mentioned that Lyons does not argue that non_linearity never obtains. It obtains, however, only for kinds of acts where some are performed <u>within</u> and some <u>without</u> a series or general practice of such acts. If the acts considered are all performed <u>within</u> a series or general

practice of such acts, then Lyons holds that non-linearity does not obtain.[9]

The UAUian may use portions of this argument for extensional equivalence to secure a defence against our utilitarian argument. We have argued that, so long as the threshold remains uncrossed, acts of UAU promise-breaking will have individual consequences at least as good as those of any alternative and will not have the harmful collective consequences of undermining people's confidence in the institution. Therefore, of no instance of UAU promise-breaking will it be true to say that it is the cause of damage to the institution; hence, of no instance will it be true to say that it has the harmful collective consequences involved in damaging the institution. It follows that in any particular case of UAU promise-breaking the individual consequences will be the determining factor for the UAUian in deciding to keep or break his promise, and employing his case by case procedure he will determine to break the promise, since, with harmful collective consequences not being ascribed to the act, it will be optimific to do so. Thus, if the practice of UAU became general, it would follow that the number of individual acts of UAU promise-breaking justified would be sufficient to cross the threshold and produce harmful threshold effects, viz., collapse of confidence in the institution. Now the UAUian may suggest that unstated, but at work in this argument, is the view that in assessing

the consequences of an individual act of UAU promise_breaking or lawn_crossing, the UAUian cannot take into account the fact that his act figures in a series or general practice of such acts, since his case by case procedure limits him to examining the consequences of each act, act by act. If harmful collective consequences are not to be ascribed to any one of the individual acts of UAU promise_breaking or lawn_crossing that comprise the series, it follows that they do not figure in the individual consequences of any of the individual acts, and so may not be assessed by the UAUian in deciding to keep or break his promise or to nip across the grass. Arguing on the strength of Lyons, the UAUian may claim that the way to avoid this utilitarian argument against UAU is to show that a UAUian <u>can</u> include in his assessment the fact that his act figures in a series or general practice of such acts; and he can show this by arguing; and he can show this by arguing that the harmful collective consequences of the practice or series <u>are</u> to be distributed in some proportion (to be deliberately vague) to their constituent acts; the upshot of which is that linearity will hold between the harm done by the series and the harm done by a single such act. Since a proportion of threshold effects will then be distributed to each constituent act, the UAUian will be able to take into account the contribution of his act to the production of harmful threshold effects, which are, of

course, the result of the series or general practice. In which case, then, the utilitarian argument against UAU will not succeed.

Put baldly, this defence to our argument would have it (i) that the main issue between ourselves and the UAUian, and his mentor Lyons, is whether threshold effects of a general practice of acts of UAU promise_breaking are to be allotted in some proportion to constituent acts; (ii) that, therefore, the success or failure of our argument rests entirely upon whether non_linearity holds; and (iii) that, since non_linearity does not hold, our argument fails.

## IX.2 Reply to the defence

The way to breach this defence is to show that our utilitarian argument against UAU may be maintained even if linearity holds. It will follow that the main issue between ourselves and the UAUian, and his mentor Lyons, is not whether linearity holds.

If we assume for the sake of argument that linearity holds, so that threshold effects of the general practice or series of acts of UAU promise_breaking are to be allotted in some proportion to constituent acts, we can still argue (i) that the UAUian does not *know* when the threshold is about to be crossed, and (ii) that *if* the threshold is in danger of being crossed, *anybody*, including himself, ought to refrain from UAU promise_breaking or lawn_crossing[10], for he is in the position of *not knowing* but what his proposed act of promise_breaking or lawn_crossing is one he ought

to abstain from. Now in a matter of important consequences, where there was not simply a perceptible difference in utility between performing the act and abstaining, this consideration would actually justify the UAUian in abstaining from that act at the threshold. Significantly, however, the UAUian does not distinguish between justification and rightness, and he is obliged by his position to say that in this case _ because of the proposed act's individual consequences, and because harmful threshold effects have not yet occurred, since the threshold remains uncrossed _ it would be right, and therefore justified, to do the act. Once again, the UAUian will find himself justifying a number of acts of promise_ breaking and lawn_crossing, some at the threshold, and it requires only a certain number of such acts at the threshold to cross it, which number will be achieved if the practice of UAU becomes general. I conclude, therefore, even if linearity holds, that we may still maintain our utilitarian argument against UAU, and that a fortiori the UAUian defence to it fails in its aim.

It is a mistake to suppose that a utilitarian argument of the kind presented here can only be generated against UAU on the assumption of non_linearity. Even if linearity holds, so long as the UAUian fails to distinguish between justification and rightness and continues to employ his case by case procedure and

adheres to optimificity as a thesis of justification, he will find himself justifying more and more cases of lawn_crossing and promise_breaking, just because these acts on their individual consequences would be right. Moreover, because of lack of knowledge in previously specified respects, the UAUian does not know when it ceases to be optimific to cross the lawn or to break promises, and does not know when the threshold is being crossed, and, thus, even if linearity holds, at what point to ascribe harmful threshold effects to constituent acts in the series. In short, the UAUian is no better off if linearity does hold: he is still plagued by lack of knowledge in crucial respects; and he still fails to make important distinctions which would circumvent _ albeit at a cost of transforming his theory into a species of TAU _ those elements in his theory which give rise to our utilitarian argument. These are the important issues for our purposes, not whether linearity holds.

## IX.3 The late ascription of threshold effects

The crucial step in the UAUian's defence is that he distributes to constituent acts of the series or general practice some portion of the harmful collective consequences, which enables him to assess along with an act's individual consequences its contribution to threshold effects. What this step overlooks in our argument, however,

is that the argument begins from the fact that, so long as the threshold remains uncrossed, acts of UAU promise-breaking will have individual consequences at least as good as those of any alternative and cannot have the harmful collective consequences involved in undermining people's confidence in the institution; they cannot have these collective consequences since the threshold remains uncrossed; and one cannot ascribe threshold effects when there as yet are no such effects. Put succintly, so long as the threshold remains uncrossed, there are no threshold effects, and it will be optimific for the UAUian to engage in UAU promise-breaking; but if the practice of UAU becomes general, the threshold is then crossed and harmful threshold effects ensue. The UAUian, in terms of his theory, is both powerless to prevent this approach to the threshold and only gains the advantage of distributing harmful threshold effects once the threshold is crossed, i.e., in restrospect, after the damage is done. Clearly, this ascription is too late to be of use practically, in deciding what to do, and such distribution of threshold effects will be only for the purpose of assessing responsibility. (I return to these matters in the next chapter.) Of course, it will be said that each UAUian will consider, before he engages in an act of UAU promise-breaking, whether his act forms part of a series or general practice of such acts, and, therefore, whether his act is likely to contribute to bringing on

threshold effects: if it is likely to contribute, he will refrain. Once again, however, for reasons given earlier, it is doubtful that the UAUian can possess the requisite knowledge about when the threshold is or is about to be crossed, about how many and how often others are behaving similarly, etc.; he is thrown back, therefore, upon the proposed act's individual consequences, and here he finds it optimific to break the promise. So, the UAUian is in a dilemma: either threshold effects cannot be ascribed, because the threshold has not yet been crossed and there are no such effects, or else they can be ascribed but only after the threshold has been crossed and the damage done. Since the UAUian is in this dilemma whether or not, or even if, linearity holds, we have firm ground for concluding that our utilitarian argument remains compelling.

# X

## UTILITARIAN CONSEQUENCES

The defence of the UAUian to our utilitarian argument turns upon the assumption that Lyons has successfully uprooted (the appearance of) non_linearity between the effects of a single act of UAU promise_breaking and the effects of a series or general practice of such acts. So, another way of coming at this defence is to take a closer look at Lyons' discussion of linearity, which I shall do in this chapter. It has been overlooked by critics, I think, that Lyons offers only a single indirect argument in support of his claim that linearity obtains; and an examination of this argument reveals that he is operating with a different conception of causal consequences from that of his non_linearian opponents. I think it can be shown that the dispute between the two camps is nothing more than a dispute over the relative merits of these two conceptions of consequences in respect of certain tasks; and here, too, we shall find grounds for thinking TAU attractive.

### X.1 Lyons and linearity

As we have seen, the cogency of Lyons' argument for exten_ sional equivalence depends in part upon the establishment of linearity, i.e., upon the view that the effects of an aggregate of acts of a particular kind is always the sum of the effects of each individual act that constitutes the aggregate.[1] For if

linearity is not established, the non-linearian (and thereby utilitarian generalization[2]) gains a foothold: he will be free to argue (i) that there is a difference between the simple and generalized utilities of some acts, so that (ii), though an act of lawn-crossing may have a balance of utility over disutility in its individual consequences, when generalized there is a balance of disutility over utility.

So far as I can determine, Lyons offers no direct arguments by which to establish linearity; rather, he hopes to achieve this goal indirectly, by showing that the supposition that <u>non-linearity</u> obtains is paradoxical.[3] His argument in support of this charge is that, under a condition of causal non-linearity, either the sum of collective consequences of a series of acts is <u>less than</u> the sums of the consequences of each individual act or it is <u>greater</u>; in the former case consequences apparently 'somehow disappear', whereas in the latter 'it would seem that some effects arise uncaused or appear mysteriously'.[4] To illustrate what Lyons has in mind, we may recall Harrod's remark that '<u>there are certain acts which when performed on n similar occasions have consequences more than n times as great as those resulting from one performance</u>'.[5] Another example is to be found in this passage by Jonathan Harrison:

> On this point Lyons is mistaken. Putting a straw on the camel's back never does the camel any harm, however heavily laden the camel already is, but a billion straws very likely will break its back. Hence the harm done to the camel by putting a billion straws on its back is more _ much more _ than the harm done to the camel by each individual straw.... 6

(This is obvious; but Harrison has not said what he meant to say, and I correct his oversight in what follows.) Here, I think, Lyons would find Harrison's conclusion paradoxical: if the harm done the camel by a billion straws is more than the sum of the harms done by each individual straw, then we appear to be getting something out of nothing, i.e., consequences which are uncaused, which are not the consequences of any particular act of putting a straw on the camel's back. I want to suggest that the paradoxical appearance of this result can be explained away.

X.2 Causal and contributive consequences

There is a car stalled at the bottom of a hill; five men are required to push the car up the hill; five men are to hand; all push; the car goes up the hill. If we ask of whose act of pushing the motion of the car up the hill is a consequence, the proper reply is that it is a consequence of all five acts of pushing, considered together; considered separately, no single act of the five has the motion of the car up the hill as a consequence. One may wish to say of one of these acts that it contributes to the

motion of the car, but it would be a mistake to say of it that it caused the motion of the car: all five acts, considered together, cause the motion of the car up the hill. This distinction between contributes and causes is the key to unlocking the mystery Lyons alludes to over non-linearity.

As we have seen Lyons thinks the appearance of non-linearity vanishes when we describe, e.g., an act of car-pushing completely, i.e., when we take into account all of what he calls the 'causal properties' of the act, including, we must recall, its social context or the extent to which others are behaving similarly. The expression 'causal property' in this context, however, is far from clear. Ordinarily, a causal property of an act or thing is a property of bringing about something: a causal property of Exlax is that it brings about bowel movements, of Ajax that it removes grease and stains. Now the sense in which the similar acts of others are a property of my act is that their acts and mine jointly determine or jointly contribute to the production of a set of consequences - my act alone does not bring about this set. To call this property a causal property risks confusion; a more happy expression for Lyons to have used would have been contributive property, since my act does contribute to the production of a set of consequences over and above those (individual) consequences of my act alone. In the car-pushing case, the motion of the car is a

causal property of all five acts of pushing, considered together, whereas it is a contributive property of each act of pushing considered singly.

The supposition that non-linearity holds appears paradoxical because Harrod, Harrison and other non-linearians restrict themselves to a consideration of the causal properties and causal consequences of an act, whereas Lyons, simply referring to the lot under the rubric 'causal properties', takes into account both the causal and contributive properties and the causal and contributive consequences of each act. The justification Lyons offers for adopting his course is that a utilitarianism - any utilitarianism - is concerned with all those properties of an act that cause consequences; because, therefore, the causal efficacy of my act of pushing vis-a-vis the motion of the car is determined by how the others act, Lyons counts the behavior of the others as a causal property of my act. The justification the non-linearians could offer for adopting their course is that utilitarianisms are concerned with causal consequences, in the sense in which the act itself brings about something or other; that, accordingly, they mean by 'causal consequence' what is ordinarily meant by that expression; and that the theory of causation Lyons is operating with, though never explicitly stated, must be bizarre, if the behavior of others can be called a causal property of my act. The mystery over disappearing

consequences and consequences arising uncaused, then, is no mystery at all; it is simply a reflection of a difference of opinion over whether contributive properties and consequences are to be considered along with causal properties and consequences for a utilitarianism. Lyons is operating with a different and broader notion of causation than are the non_linearians, which explains why he finds, on the supposition that non_linearity holds, consequences disappearing or arising uncaused.

X.3 Individual and collective acts

To the non_linearian, consequences to which his act contributes are not, in the relevant sense, caused by his act. Because he distinguishes between causal and contributive consequences, therefore, it is important that he be able to tell what is and what is not a causal consequence of his act, which in turn necessitates some account of what a causal consequence is. The account given may be an ordinary one, such as that to be found in G. E. Moore's 'A Reply To My Critics':

> One natural way, and perhaps the most natural way, of understanding the ex_ pression 'the total consequences of the action, A' is one in which among the consequences of A nothing is included but what is the case subsequently to the occurrence of A, so that the 'total consequences of A' means everything which is the case subsequently to A's occurrence, which is also such that it would not have been the case if A had not occurred. 8

Moore's view is that for c to be a consequence of act A, the occurrence of A must be a necessary condition for the occurrence of c. Such a view provides the non-linearian with a means of telling what is a consequence of his act: he must find out of everything which occurs subsequently to his act to which of these things his act stands in the relation of necessary condition.

Three nations announce that they are each to explode a nuclear bomb in the atmosphere. The bombs will be of the same megatonnage and will each release the same amount of strontium 90 into the atmosphere, viz., 2 units each on an internationally agreed scale. Studies carried out by an impartial body, the findings of which are accepted by the three nations, show that a level of 4 units of atmospheric pollution due to strontium 90 is a serious health risk, and is known to have detrimental and lasting effect upon human and animal genes. The study further reveals that this danger only arises if the 4 units of strontium 90 are released into the atmosphere within the same month. Each nation announces separately from the others both that it is to explode a bomb in the atmosphere this month and that the fall-out from its bomb will be 2 units; moreover, the three nations are enemies, and so do not exchange information on their nuclear-testing programs. Each fears that either of the remaining two may achieve a breakthrough

of enormous importance in weapons development, which will be turned into a military threat against it; accordingly, each nation decides that it cannot wait to explode its bomb. Without any communication, nations A and B explode their bombs, roughly simultaneously: 4 units of strontium 90 are released into the atmosphere, and the threshold of danger due to atmospheric pollution is crossed. Did A cause the danger, and are threshold effects to be ascribed to its act? Given the account of consequences in terms of necessary conditions, A may argue that, though its bomb releases 2 units of strontium 90 into the atmosphere, it does not cause atmospheric pollution. Its act is not a necessary condition of pollution: its bomb releases 2 units only, and, moreover, if nations B and C explode their bombs, pollution occurs whether or not A explodes its bomb. Likewise, B will argue that its bomb is not a necessary condition of pollution, since even if it refrains, C could explode its bomb; so B's act did not cause pollution and the present danger. We have, then, a situation in which atmospheric pollution occurs, but those who perform the acts which together have this harmful threshold effect argue that the effect is not ascribable to either act in the series; it is, in other words, to be ascribed to the series itself. There is no mystery here, no consequences arising uncaused. On the one hand, neither act <u>considered singly</u>

is a necessary condition for the occurrence of atmospheric pollution (and therefore cannot have pollution among its causal consequences), since the other two nations could explode their bombs and still bring on pollution. (It is important to the case, therefore, that there are more than enough parties to surpass the threshold and bring on threshold effects: were there only two nations with bombs to explode, and the threshold remained at 4 units, then each act <u>considered singly</u> would be a necessary condition for the occurrence of pollution and pollution would therefore figure in each act's causal consequences.) On the other hand, it is true that both acts <u>considered together</u> (i.e., the series of acts) have the consequence of pollution. Let us refer to both acts considered together as the <u>collective act</u>: what the non_linearian is arguing, given his treatment of consequences in terms of necessary conditions, is that we cannot ascribe as threshold effects of an individual act what can only be a threshold effect of a collective act. In other words, a collective act is the necessary condition for the occurrence of collective conse_ quences or threshold effects, i.e., pollution; since neither A's nor B's act is itself a collective act, it follows that pollution is not a consequence of their individual acts considered singly. Thus, there are individual acts and individual consequences and collective acts and collective consequences; to the non_linearian,

only confusion will result in the bomb case from treating collective consequences as if they were the consequences of individual acts. Likewise, in the car-pushing case, to ascribe the motion of the car up the hill to my act of pushing is to treat a collective consequence as if it were the individual consequence of my act; it is, on the contrary, a consequence of a collective act, i.e., all five acts of pushing considered together.

Lyons, on the other hand, will argue that each bomb contributes to atmospheric pollution, for each bomb releases 2 units of strontium 90; and he will want this contribution to figure in the consequences of each act. He operates on the view that *c* is a consequence of act *A* if the occurrence of act *A* contributes to the occurrence of *c*. Nor does he care how many other acts contribute to *c*: so long as act *A* does, *c* (or some portion of *c*, if it is divisible) can figure as a consequence of act *A*. Clearly, then, the non-linearian has a narrower view of what a causal consequence is than does Lyons, and he relies, as Lyons does not, on distinctions between individual and collective acts, causal and contributive properties and consequences, and singly determined and jointly determined consequences.

X.4 The non-equalizer

It is perhaps also worth mentioning that Lyons will have to

contend with the non_equalizer. Unlike the non_linearian, the non_equalizer accepts Lyons' view that threshold effects are to be ascribed to constituent acts of a series; however, he denies, should Lyons think otherwise, as I think he does, that each act in the series contributes equally.

A twist to our bomb example gives the non_equalizer his opening. A and B explode their bombs under the same conditions as before, but, instead of simultaneous explosions, we have sequential explosions; and A's bomb is exploded first. B's position is now altered; for although its bomb is not a necessary condition of pollution (since C may still explode its bomb), never_theless, in this sequential case, <u>because</u> A has already exploded its bomb, B's act is a sufficient condition for pollution.[9] (Thus, A's getting in first alters the situation of those who subsequently act similarly: (i) A gets the benefits of performing the act without the disadvantage of bringing on harmful threshold effects; (ii) A brings us nearer the threshold, without crossing it; and (iii) A increases the likelihood that a subsequent, similar act by another agent will carry us over the threshold.) So, asks the non_equalizer, given that B's explosion suffices to bring on harmful threshold effects, why should not total threshold effects be ascribed to B's act? Or, if this is too extreme, why should we not say that B's act does disproportionate harm? For it is B's act that

pushes us over the threshold. The non_equalizer, moreover, may present the same argument in the other cases as well; thus, if the camel's back breaks after but not before the billionth straw is placed on its back, why should we not say that the billionth straw does disproportionate damage, or even that it is the straw that breaks the camel's back?

It may even be thought that one act in a series can make a negligible or no contribution to harmful threshold effects. Suppose nations A, B and C explode their bombs sequentially, releasing 6 units of strontium 90 _in toto_, the threshold remaining at 4 units. Just as A's getting in first alters B's situation, so, argues the non_equalizer, the fact that A and B explode their bombs before C alters C's position. Specifically, it is now impossible for C's act to be either a necessary or sufficient condition for pollution, since pollution has already occurred by the time C acts. In this case, pollution is not a causal consequence of C's act, nor is C's act even part of the collective act necessary to produce pollution. Therefore, pollution is neither a causal not a contri_ butive consequence of C's act. So, in spite of the fact that C's act increases the level of strontium 90 in the atmosphere, for which it bears responsibility, the _non_linearian_ will argue that C's act neither contributes to crossing the threshold nor brings us nearer the threshold; whereas the _non_equalizer_ will argue that, while all

three explosions contribute to the production of harmful threshold effects, it is by no means clear that C's act makes the same contribution as B's or even A's. B's act suffices to push us over the threshold, and A's act brings the threshold a little nearer and increases the likelihood of its being crossed; neither of which are true of C's act. So why should we not say, since the damage has already been done by A and B, that C's act makes a negligible or no contribution to the production of threshold effects?

The most obvious move against the non_equalizer fails. In the case of sequential explosions by A and B, the non_equalizer claims that, though A and B both contribute to pollution, they do not contribute equally; B contributes more. Here, one may be tempted to argue against the non_equalizer that A's act should not be described as 'exploding a bomb' but as 'exploding a bomb in the knowledge that further bombs are to be exploded'. One hopes in this way to bring out that A had a hand in bringing on pollution or threshold effects. Obviously, however, this move at best only shows that A and B <u>make a contribution</u> to pollution; it does <u>not</u> show they <u>contribute equally</u>. And, in any event, it <u>may be</u> quite false to describe A's act as 'exploding a bomb <u>in the knowledge</u> that further bombs are to be exploded', since, among other things, B and C <u>may not</u> announce publicly their intentions to test in the atmosphere.

Another way of trying to show that each act does equal harm is

this: in the camel case, one can argue that each straw does equal harm, on the assumption that the others are or will be there too, in that, if it were omitted, the others will not break the camel's back. Whichever straw we select, if it is omitted, the others do not break the camel's back; therefore, each straw makes an equal contribution. This line appears plausible, I think, only because a billion straws are required to break the camel's back and there are just enough straws to do the job. Suppose, however, there are more than enough acts of the kind in question (as was the case when nations A, B and C all exploded their bombs) to produce harmful threshold effects; e.g., while a billion straws are required, one and a half billion are placed on the camel's back. Do all one and a half billion contribute equally to breaking the camel's back? The present argument will not establish the point, since whichever straw we remove, the camel's back will still be broken. So even on the assumption that all the others are or will be there, it has not been shown that the last straw makes a contri_ bution equal to that, e.g., of the billionth straw. Moreover, by the time the last straws of the additional half_billion are added, the camel's back is long since broken, and it seems misguided to say even that these straws contribute to producing an effect that has already been produced. Whatever it is that the additional half_

billion straws contribute to _ the body of the camel and all that straw may create an obstruction _ it is not to breaking the camel's back. In situations, then, in which there are more than enough acts of the kind in question to bring on threshold effects, and in which the acts are performed sequentially, the non_equalizer has a strong case.

Finally, the only argument, so far as I can determine, that Lyons offers to show that threshold effects are to be distributed equally to constituent acts of the series is itself inconclusive.[10] Put summarily, Lyons argues that acts exactly similar in their causal descriptions fall into 'causally homogeneous classes'[11], and acts are exactly similar in their causal descriptions only if they are equally efficacious; therefore, causally homogeneous classes of acts contain only members of equal causal efficacy. Accordingly, if two acts are exactly similar in the causal descriptions, they are equally efficacious in their consequences, which I take in part to be saying that each act may be regarded as making the same contribution to threshold effects. But only in part, for acts that form part of a causally homogeneous class must also have similar ordinary, non_threshold_related effects.[12] In essence, then, Lyons argues that two acts contribute equally to threshold effects, if there is no relevant difference between them in causal efficacy.

Here, however, Lyons construes 'causal efficacy' in the broad sense, to include both singly determined and jointly determined consequences as consequences of my act; this implies that linearity obtains, which implies that linearity has been shown to obtain; and this means that Lyons' argument to establish that contributive acts contribute equally to threshold effects is inextricably bound up in the controversy over linearity. In other words, we now know that Lyons will overcome neither the non_linearian nor the non_equalizer unless he can, by arguments other than the ones he uses and we have considered, establish conclusively that linearity obtains. Moreover, to emphasize the obvious, neither the non_linearian nor the non_equalizer need (or would) agree with Lyons as to which causally homogeneous class certain acts belong to. For example, no non_linearian need (or would) agree that two acts can have exactly similar causal descriptions, if one is framed in terms of his narrower conception of consequences while the other is framed in terms of Lyons' broader conception; and no non_equalizer need (or would) agree that two acts can have the same causal efficacy if one suffices to push us over the threshold and the other does not, though both figure in the same series of acts. Thus, the placing of certain acts into causally homogeneous classes may (or will) itself be a bone of contention among the parties.

X.5 In prospect and in retrospect

Finally, to these complaints by the non-linearian and the non-equalizer we must add the complaint of our TAUian, viz., that Lyons' contributive view of consequences is of little use for assessing consequences in prospect. If we are trying to decide what to do, and we try to envisage the consequences of our proposed acts, the contributive view would have us assess not only our acts' individual or singly determined consequences but also their collective or threshold-related or jointly determined consequences. The view the TAUian takes of this suggestion is the one set out in the preceding chapters: for reasons given there, it is implausible to assume that agents in prospect know what the collective consequences of their acts will be; and I have argued throughout that it is implausible to assume agents know all the individual consequences of their acts. Our lack of knowledge in prospect both of the individual and collective consequences of our acts is the really important point vis-a-vis Lyons; and, too, it is precisely because I do not know either sort of consequence in prospect that I must not act on UAU.

On the other hand, for assessing consequences in retrospect, the contributive view works nicely. This broad view of causal consequences enables us to take into account all the consequential properties, both causal and contributive, of our acts; since both

sorts of properties issue in consequences, and since in retrospect a good many of these consequences become known to us, we may assess many of the act's singly and jointly determined consequences. If a ship sinks, and a board of inquiry is formed to find the causes of the disaster, it will try to uncover, not merely the necessary conditions for the ship's sinking, but anything else that may have contributed to the disaster. In this case, in assessing consequences in retrospect, the narrow view of causal consequences is too narrow, just because we have before us in retrospect, which we do not have in prospect, a great many _ though almost certainly not all _ of an act's singly and jointly determined consequences. Then, too, since our aim in employing the broad view of causal consequences in retrospect is assessing responsibility and/or allocating blame, the fact that we are apt to feel impatient with a defence of the narrow view is probably a reflection of our desire to hold respon_ sible all members of a group which collectively performs some act, e.g., pollutes the atmosphere with strontium 90. I conclude, therefore, that it is in retrospect, especially because it enables us to include all the consequential aspects of acts, that the broad view appears most useful; and it is in retrospect, therefore, and primarily for the purpose of ascribing responsibility, that the assumption of linearity is most plausible.

In prospect, the broad view, and the UAUian's reliance upon it

to generate a defence to our utilitarian argument against UAU, runs up against a major hurdle, viz., that we do not know at the time of acting to what our act will contribute and therefore what the collective, threshold-related or jointly determined consequences of our acts will be. Accordingly, we shall be unable to include them in our assessment of consequences for the purpose of deciding what it would be right to do here and now, in this case. As a result, we are thrown back on the individual consequences of acts, such as they are in prospect, and here the UAUian will find himself in a spiral of justifying acts, e.g., of UAU promise-breaking, because he thinks they are right or optimific. Moreover, if actual consequences make acts right or wrong, then it is in retrospect that we shall know whether our act was right or not, or whether the course we adopted was in fact the optimific one; thus, the individual trying to decide what to do is 'reduced' to asking what he would be justified in doing, what course would give him the best chance of doing the right, i.e., optimific thing. Once again, the TAUian's distinction between rightness and justification looms important. Lastly, since we also experience difficulty in calculating carefully the singly determined consequences of our acts, we shall stand the best chance (for reasons given earlier) of doing the right or optimific thing, then, because our situations are murky vis-à-vis the singly and jointly determined consequences of our acts, if we

adhere to the rules and institutions of the 'plain man'. Here, too, TAU can assist us.

In short, then, I do not think Lyons has established that linearity obtains (or, for that matter, that the non-equalizer has an untenable position); and the assumption of linearity appears most plausible in retrospect, for the purposes of allocating responsibility and blame. The UAUian, relying, as he does, upon Lyons, will find that it is only in retrospect that he can take into account the fact that his act figures in a series of such acts and, as a result, the contribution of his act to threshold effects; by then, through the general practice of UAU, harmful threshold effects will have been produced and the institution of promising undermined.

One final point. There is a type of case where the employment in prospect of the narrow view of causal consequences appears important for utilitarianisms. This is the case where similar acts are performed sequentially and an agent knows his place in the sequence. Nation A explodes its bomb, and B is trying to decide what to do, i.e., whether or not to explode its bomb. B knows that A has let off its bomb, that A's bomb has released 2 units of strontium 90 into the atmosphere, that 4 units exceeds the threshold, and that its own bomb will release 2 units. In

this case, the narrow view is of some benefit. For B can know, with C still to decide, that while the explosion of its bomb is not a necessary condition for the occurrence of pollution, since C may still explode its bomb, it will be a sufficient condition for pollution however C decides. Thus, B can know in advance and take into account in reaching its decision that the explosion of its bomb will suffice to pollute the atmosphere. Now if A and B could explode their bombs underground, if negligible amounts of strontium 90 would thereby be released into the atmosphere, and if A explodes its bomb in the atmosphere, then B can know in advance that exploding its bomb underground avoids pollution. Therefore, B is in this position: if it explodes its bomb under_ ground, under no circumstances can its act have pollution either as a causal or contributive consequence, whereas if it explodes in the atmosphere there are circumstances, i.e., the present case, in which its act can have pollution as a contributive consequence. Such information in advance, while it does not tell B that ex_ ploding its bomb underground is in fact optimific, will neverthe_ less justify B in exploding underground, in this case of important consequences; such information, then, will be important to B in reaching its decision about what to do. Furthermore, this infor_ mation will be of importance to us as well, for purposes of ascribing responsibility and allocating blame, if we find that B

decided to explode its bomb in the atmosphere and so intentionally pushed us over the threshold of pollution.[13]

# XI

## JUSTIFICATION AND EXCEPTIONS

I want in this chapter to consider further the topic of justifying exceptions to the useful social institution of, e.g., promising. My aim is modest, and is certainly not to attempt to resolve anything; rather, it is to make one or two points that seem to me important for discussions of institutions and utilitarianisms but which appear to go unremarked in much recent work in this area.

### XI.1 Justification and consistency

If we see the 'plain man' as demanding some account of how his social institution of promising is compatible with a species of AU, then it is clear that UAU fails to achieve compatibility, for, in a word, the 'plain man' does not think optimificity a general justification _ let alone the justification _ for promise-breaking. Worse yet, the general practice of UAU will undermine this institution of promising, and, by implication, other useful social institutions as well. By contrast, because it severs the link between justification of exceptions and optimificity, TAU fares better; and by further exploiting the distinction between what is right and what is justified in the circumstances, TAUians can on AUian grounds advocate adherence to the social institutions of the 'plain man' and thereby demonstrate the compatibility in question. Since he does not accept optimificity as a general

justification for exceptions to these institutions, the TAUian escapes our utilitarian argument against UAU: either singly or in conjunction with a great many others, the TAUian may act on his theory, and may encourage and even try to persuade others to act on his theory, without fear of undermining, e.g., the institution of promising.

However, if the TAUian rejects the counterintuitive view that no exceptions to the institution of promising are justified and rejects as well the UAUian's reliance upon optimificity to justify exceptions, then he both thinks that there are justified exceptions and requires some way of telling, some means by which to distinguish, the justified from the unjustified claims to exceptions. How is the TAUian to meet this requirement? The answer, I think, is reasonably clear from one of the lessons to be drawn from the misfortunes of the UAUian; put succintly, the grounds the TAUian will accept as justificatory for exceptions to the institution will have to be such that their general employment or practice - unlike the case of optimificity - will not undermine the institution.

For a host of reasons, nearly all of which figure prominently in the literature and a good many of which have been emphasized in past chapters, and a rehearsal of which may therefore be dispensed with here[1], the 'plain man's' institution of promising is

immensely useful; and if, in addition, it really is part of the cement of human relationships and helps to foster communal cooperation, then its immense utility is even more earily estab_ lished. I should say, in other words, that utility itself appears to demand that we support this institution, where 'support' in this context means to preserve (or to act so as to preserve) the institution's utility. Of course, the obvious way to support the institution is through keeping promises, especially since at least general promise_keeping is required in order to preserve the institution and confidence in it unimpaired. Now it has very often been overlooked that accepting certain grounds as justifi_ catory for exceptions to the institution is logically inconsistent with support of that institution, which utility demands. Thus, accepting optimificity as such a ground is logically inconsistent with one's support of the institution, since the general practice of optimificity undermines the institution's utility. The same applies to the ground of self_interest: the general practice of breaking promises on the ground of self_interest will impair confidence in the 'plain man's' institution. Because, then, utility demands that we support this institution, where this amounts to acting so as to preserve its utility, we can only accept as justificatory grounds for exceptions those the general practice of which will not undermine the institution's utility;

or, negatively, to claim exceptions on grounds the general practice of which will undermine the institution's utility is inconsistent with our support of it.[2] Nor should we be side_tracked by some claim that one can be practically certain others will not justify promise_breaking on the ground(s) one cites: unless one is prepared to argue _ as utilitarians are not _ that this ground is justificatory for oneself alone, then to accept ground X as justificatory for promise_breaking is implicitly to permit acts justified on this basis, whether performed by oneself or another party. On the view in question, then, an exception will be justified if the general practice of the ground on which it is claimed will not impair the institution's utility; if the general practice of this ground will impair the institution's utility, we cannot consistently _ consistent with our support of the institution, which utility demands _ accept it as justifi_catory for an exception. We have here, then, a means of separating a justified from an unjustified claim to an exception, which is what the TAUian seeks.

Just as utility makes certain demands on us, so too does logical consistency: once we understand in this matter what utility demands of us, i.e., support of the 'plain man's' institution, then we shall find that consistency places us under certain

restrictions as to what we can and cannot accept as justificatory grounds for exceptions. (I shall not here go over those factors which emerged in previous chapters, e.g., lack of knowledge of the threshold, etc., that must make us even more tentative about the grounds for our claims to exceptions.) However, I want to suggest that there is yet a further restriction we are under in this matter, one of some importance, and so worth drawing attention to.

## XI. 2 Discharge restrictions and roles

An opponent may try to avoid these restrictions consistency places us under by trying to limit the applicability of an alleged justificatory ground for an exception both *to this case only* and *solely to himself*.[3] In essence, such an opponent claims that *only he* is entitled to an exception on *this* ground, and that he is claiming a justified exception on this ground *for this case only*. It does no good for him merely to argue for a justified exception on *this* ground for this particular case alone; for, once again, to accept ground X as justificatory for promise-breaking is implicitly to permit acts justified on that basis, whether by himself or someone else; so he must also add that ground X is justificatory for himself alone. This individual is claiming a privilege; when he objects to repaying the £5 he promised to repay because to do so would be to deprive himself of

a meal at a fashionable restaurant, it is because he is deprived that is the important feature of the case for him. Now, obviously, adherents to some universalizability principle are going to find this individual in breach of it; and I see no reason to dissent from this verdict. However, there is another opponent, not to be confused with this first one, who can pose a problem for us, and it is upon this individual that I wish to focus.

Two sorts of cases need to be distinguished: first, those, as in the case above, in which the individual, the promisor, cites features about himself as justificatory for exceptions, e.g., that he (because he is who he is) is deprived of something; and second, those cases in which the individual, the promisor, cites features about the promisee as justificatory for exceptions, e.g., that he has black skin. Even if one agrees that cases of the former sort do involve breaches of universalizability, it does not follow that cases of the latter sort do. Moreover, a promisor of this second sort may go on to admit that, in accepting the black skin of the promisee as a justificatory ground for exceptions, he is implicitly permitting the acts of others (i.e., other non_blacks) to be justified on this ground; but he points out that, even if all non_blacks did not keep promises to blacks, we need have no fear of undermining the utility of the institution, since there are vastly more non_blacks than blacks in our society, enough, in

fact, to sustain the institution's utility.[4] This individual would appear to satisfy the TAUian's requirements: he understands what utility demands of him in respect of the institution of promising; he recognizes that, in view of his support of the institution, consistency places him under certain restrictions in accepting justificatory grounds for exceptions; and he does not accept as a justificatory ground one the general practice of which adversely affects the utility of the institution. So how, then, are we to deal with this opponent, whom, for convenience, I shall call White?

White cites Black's skin color as a justificatory ground for breaking his promise; but there is a clear sense in which he could equally well have cited the promisee's hooked nose or country of birth or religion; for he is the one doing the deciding as to which characteristics of the promisee justify breaking promises to him. This is the crux of the matter, and, I think, the point on which we may develop an argument against White. It is the institution of promising that is under discussion, and institutions are in part defined by roles: when Smith promises to repay Jones, Smith steps into or adopts the role of promisor and Jones, if he accepts the promise, into the role of promisee. White promises Black that he will repay him the money, and Black accepts the promise, but when it comes time to repay Black, White cites Black's

skin color (or religion, etc.) as a justificatory ground for breaking his promise: given that these are the facts of the case, what is to prevent us from describing _ and correctly describing _ White's action (citing skin color) as acting in accordance with the prin_ ciple that promisors may set their own restrictions to the dis_ charge of their promises? By promisors setting their own restrict_ ions, I mean in this matter deciding upon those characteristics of the promisee that justify them in not keeping or discharging their promises: let us for reasons of economy refer to this as setting their own discharge restrictions. Why, then, may we not say of White that he is acting in accordance with the principle that promisors may set their own discharge restrictions? Of course, White might argue that he is not really acting on the principle that promisors may set their own discharge restrictions since he really does believe that promises need not be kept to blacks. The state of his belief, however, is immaterial to whether his action in citing skin color as a justificatory ground for an exception bears the interpretation in question. What is material is whether he occupies the role of promisor, and in the present case he does. Now I think White is in a quandary. His action is in accord with the principle, his action indicates that he accepts the principle that promisors may set their own discharge restrictions; but since anyone stepping into the role of promisor

may then set his own discharge restrictions and since we all are or are potentially promisors, White's acceptance of this principle, his action in accordance with it, appears inconsistent with his support of the institution; for the general practice of such a principle appears certain to impair confidence in the institution and so to affect adversely its utility. If White accepts that only some promisors (i.e., whites) may set their own discharge restrictions and, further, may set such restrictions only in respect of skin color (i.e., black), then his position is radically arbitrary. For, to mention only a few qualities, skin color, hair color, eye color, hooked nose, height, weight, etc., appear of the same logical type; and it can only be arbitrarily that one but not the others are permitted to justify exceptions.[5] If White accepts that some promisors, i.e., whites, may set their own discharge restrictions, but that these promisors are free to set any such restrictions they wish, then his position appears arbitrary in another respect, viz., in picking those promisors who enjoy the privilege - that no other occupants of that role may enjoy - of setting their own discharge restrictions. (And, of course, if all whites in our society may set their own discharge restrictions, the utility of the institution of promising will be impaired; there are more than enough of them to see to this. And what would the consequences be if more and more whites began to cite white skin color

as a discharge restriction?) Lastly, if White argues that only he as promisor may set his own discharge restrictions (and he cites black skin), then he appears to be ignorant of what it is to occupy a role such as that of promisor. To step into or adopt a role is in part to give up the possibility of appeal to some such view as, 'it is because I am affected or because I dislike or because I am who I am that entitles me to this exception': the personal characteristics we bring with us do in no way characterize the role we have adopted; furthermore, to appeal to them as if they could confer entitlements upon us in that role is simply a failure to understand what stepping into a role means. Put differently, we can demand from White some explanation of how it is that he has an entitlement as promisor (to appeal to characteristics about himself as justificatory of an exception) which no other occupant of that role has or can have. Doubtless, the fact is that White has no such entitlement in the first place.

A crucial point against White is that, by citing skin color, he is acting in accordance with a principle - viz., that promisors may set their own discharge restrictions - that is inconsistent with his support of the institution of promising, which utility demands; for the general practice of this principle appears certain to affect adversely confidence in the institution.[6] We must be careful, however, not to go too far; for there are cases in which

a promisor is justified in setting his own discharge restrictions. For example, suppose Smith has promised to return a gun he borrowed from Jones, but that when it comes time to return it, Smith discovers that Jones is out of his senses, so that Jones might harm himself and others if the gun is returned and the promise kept; thus, Smith decides not to keep the promise. Smith is here setting his own discharge restrictions, and I suppose that we all think him justified in doing so. The reason for this, and the reason this and similar cases need not threaten the position taken here on promisors setting their own discharge restrictions, is that we think that promises should operate for the benefit of the promisee; and a kept promise in the gun case is almost certainly not going to operate for the benefit of the promisee. On the other hand, if we look at the grounds White cites or may cite (or that are cited for him in the literature) as justificatory of exceptions, e.g., skin color, religion, place of origin, etc., we shall find in virtually every case that the discharge restrictions he sets operate to the detriment of the promisee and, it should be added, to his personal benefit.

---

In fine, in order to avoid going seriously wrong in discussing claims to exceptions to the prescriptions of institutions, I think the TAUian is right to insist that we always bear in mind these

points about the demands of utility, the restrictions logical consistency places us under, and the outcome of promisors setting their own discharge restrictions to the detriment of the promisee. He is also correct in insisting on these points as material to utilitarianisms, at least to those versions, such as his own, that are concerned with practice and with being practically implemented, implemented here and now, with the world as we find it and with ourselves as we are, and implemented generally or universally, and not by just one or a few persons. The general practice of TAU will not undermine confidence in the 'plain man's' institution of promising, and the findings of the present chapter, vis-à-vis the justification of exceptions to this institution, support this view.

# XII

## GENERALITY IN RULES

In this concluding chapter, I want to say something further on the rules to be found in TAU; specifically, I wish to set out another important characteristic of these rules, to supplement what has been said on this score already.

### XII.1 RU and TAU

Cumulatively, our various arguments, together with the failure of the UAUian's defences against them, especially the failure of his Lyons-supported defence against our utilitarian argument, have provided us with conclusive reasons for rejecting UAU. Conversely, the success of TAU, formed in awareness of these arguments, in escaping them altogether, without recourse to RU, provides us with a powerful reason for adopting it. Essentially, what is being claimed for TAU is, first, that it both escapes arguments that prove awkward for UAU and remains a viable position vis-a-vis the views of the 'plain man'; and second, that it achieves this result without recourse to RU, i.e., without turning itself into a species of RU. The fact that TAU not only escapes the arguments that prove the undoing of the UAUian but also remains a viable position in respect of the 'plain man' is a ground upon which we can proclaim TAU's superiority to UAU as a satisfactory species of utilitarianism; for TAU can get round the objections to UAU that we have chosen as our obstacle course, i.e., TAUians can demonstrate that a species of AU

is compatible with and finds a place in it for the common social rules and institutions of our society. On the other hand, the fact that it achieves this result without recourse to RU is a ground upon which we can proclaim TAU to have rendered unnecessary the various forms of RU. Specifically, virtually all RU theories have had their genesis in _ and have been designed to overcome _ alleged defects in AU theories in respect of accommodating the views of the 'plain man' on particular cases or kinds of cases, especially, as we saw in connection with the hard cases argument in Chapter VI, institutional cases, such as promising; TAU, however, does not suffer from these defects; therefore, the adoption of some RU theory to remove such defects is unnecessary. Moreover, I in_ cline to the view that TAU renders the adoption of RU theories un_ necessary in another way, viz., that it brings us the benefits the latter are supposed to bring (without the abandonment of its essential AU features); but I here advance this view only tentatively, for this is not the occasion to undertake an examination of all the major and important benefits the different versions of RU have been thought to confer.

(It should be clear that this explanation for why so little has been said throughout this essay of RU theories is completely at odds with Narveson's reason for ignoring RU theories. In Morality and

Utility, Narveson claims that the development of RU theories has been unnecessary and, too, misguided, because the objections to classical AU theories, i.e., species of UAU, they have been designed to meet are largely incorrect.[1] I think just the opposite, viz., that the objections to UAU canvassed here force us to give it up as a satisfactory utilitarianism; however, I part company with the RUians over the question of whether resort to a species of RU is necessary in order to overcome these objections.)

XII.2 Cooperation

It is sometimes said that there is one enormously important benefit of RU theories which UAU does not share, viz., that action in accordance with them enables us to foster cooperation amongst individuals. Failure to achieve cooperation is then, quite rightly, portrayed as striking at the very glue of our society and, therefore, as a serious objection to UAU.[2] I have said enough already of the AUian grounds on which the TAUian can advocate adherence to the rules and institutions of the 'plain man' and of the exceptionally strong presumption that the TAUian will keep promises, tell the truth, etc., to indicate that the application of this serious objection to TAU would be inapposite. But there is another point here that must not be overlooked.

A certain interdependence of people and actions appears necessary for individuals living in communities, and for a number of reasons,

not the least of which is that a degree of cooperation in our everyday lives saves time and energy for other essential tasks and so may serve to improve our lots. Centrally, however, communities have goals the attainment of which requires cooperation on the part of its members. The pursuit of such goals often demands the coordination of actions, plans, decisions, policies, etc., and whether the required degree of coordination is forthcoming is often a function of whether individuals in the community do their share. If we must be able to count on individual members doing their share when the time comes, then it is probably important that they on occasion be bound to the performance or non-performance of certain acts: if they are bound, we can count on them and proceed accordingly. (It is for some such reason that we may also on occasion demand more from an individual than a statement of intention or even of firm intention.) However, we shall not be able to count on the UAUian, because even if he acknowledges that he is bound, he will only do what he is bound to do if what he is bound to do is optimific at the time when he comes to do it; and this consideration introduces uncertainty into the situation. On the other hand, it is suggested that we <u>can count</u> on a RUian. What it is important to stress is that, even if true, the reason we can count on a RUian has nothing to do with the role or place of rules in his theory as the arbiters

of rightness or with any particular analysis or treatment of rules his theory may incorporate; rather, it has to do with the supposition that he rejects optimificity as a thesis of justification, i.e., as justifying him in departing from his rules (maxims, principles, etc.) and the prescriptions of institutions. Put differently, the general application of UAU undermines the useful social institution of promising, and an RUian will avoid this fate not through any link between rightness and rules but through rejecting the crucial feature of UAU that brings it about and which generates uncertainties over the UAUian's dependability, i.e., through rejecting optimificity as justifying him in departing from these rules and institutions. Thus, if we can count on a RUian to keep his promise and do what he is bound to do, it is because we suppose he rejects the UAUian's thesis on justification. There is nothing inherent to this picture which prevents an AUian, once he takes the steps the TAUian has taken, including the rejection of optimificity as a thesis of justification, from being counted on at least as firmly as the RUian.[3] So, if cooperation within a community requires our being able to count on the individuals within it, then the TAUian can be counted on as strongly as the next man; so there is no reason why the TAUian cannot foster this benefit of cooperation as well as the RUian. Put negatively, and more plainly, the general practice of TAU will not prevent this important benefit of cooperation from arising, and this result obtains without

TAU having to turn itself into a species of RU and so giving up its strong consequence component.

XII.3 Cooperation and rules

Now I think the TAUian can foster this necessary element of cooperation even better than most forms of RU, and for a particular reason, the importance of which for our purposes extends beyond the present discussion.

I think rules have a useful role to play in securing this interdependence of people and actions, simply because following them and encouraging others to follow them appears the simplest as well as the easiest method for achieving it. ( I emphasize the simplest and easiest method, not the only method: the view that cooperation or coordination can be achieved without rules finds solid support in Thomas C. Schelling's theory and analysis of coordination games.[4]) In this, perhaps, is to be found yet another reason for inculcating rules into our characters and fostering a habit of rule-following in ourselves and others. However, this method will remain the easiest only so long as the rules to be inculcated themselves remain general (in a sense to be explained); for such general rules as 'Promises are to be kept', 'Animals are not to be fed', etc., are easily learned and taught, are capable of being widely understood, may be applied by even the less intelligent or developed members of

a community, and, most importantly of all, have applicability to a very broad range of cases. This last feature is central to my use of the term 'general'. The rule 'Tigers are not to be fed' is less general than the rule 'Mammals are not to be fed', which is less general than the rule 'No creature in the zoo is to be fed': there are degrees of generality in rules, and I take this to mean that the more general a rule becomes the broader the range of cases to which it applies, whereas the more specific a rule becomes (e.g., 'White tigers in this part of the zoo are not to be fed cookies') the narrower the range of cases to which it applies.

Following rules, and encouraging others to follow them, will remain the easiest method for achieving the necessary interdependence of people and actions only so long as the rules to be followed, and, perhaps, to be inculcated into our characters, themselves remain general, and so remain easily taught, learned, understood, and broad in the range of cases to which they apply. In other words, their very generality accounts in large measure for their practical usefulness, and it is a noteworthy feature of TAU that the rules it advocates adherence to, those of the 'plain man', are of this general sort. Now while it would be an over-simplification to say that all RUians have preferred specific, if not highly specific rules, (and an error to suppose that a RUian cannot develop his theory around such general rules as those of the 'plain man'),

this has certainly been the trend, as Lyons makes clear.[5] The general explanation for this trend is, I think, to hand: because exceptions to rules have been thought to pose a problem for RUians, since direct appeals to optimificity can crop up as an attractive way of determining whether this is in fact an exceptional case, the temptation has been for RUians to treat exceptions as only apparent, as but further specifications of the rules, to be 'built into' or included into them and their statements or formulations.[6] Clearly, the more exceptions 'built into' the rules, the more specific they become, and in turn the more difficult, because complicated, to formulate. Degree of specificity in rules, then, can be _ I do not say will be _ a point of division between TAU and RU theories; certainly, the point is important enough to warrant our asking RUians for specimens of the rules in their theories. For, as is frequently commented upon, very specific rules may not be capable of being learned, taught or understood, or of being applied to even a moderately extensive range of cases _ hence, perhaps, the un_ welcome necessity for yet additional rules. And, too, some philo_ sophers have diagnosed _ correctly, I think _ the collapse of RU into some form of AU, if the rules are allowed to be of <u>any</u> degree of specificity.[7]

From the TAUian point of view, however, the really crucial ob_

jection to very specific rules _ whether found in RU theories or any other _ is that their practical usefulness, i.e., their utility value from the point of view of practice, is diminished the more specific they become. If we are looking for those guides to give us the best chance of doing the right or optimific thing, we shall find rules of a very specific character not very helpful, precisely because such rules will not be able to take account of the varied and complex circumstances surrounding cases, so that their range of applicability cannot be very great. What we want are reliable guides the following of which will, in the broad range of cases we are to expect to meet with, will see us through. I must emphasize that the TAUian, concerned as he is with practice, will be especially impressed by the utility of such general rules, of which the 'plain man's' are instances, and will use this fact in his general argument in favor of adhering to these rules. The TAUian calls these rules 'rules of thumb', since they do not make right or wrong the acts falling under them; but this rubric should not mislead us into thinking the TAUian will let these rules be dispensed with easily in practice. And any move which minimizes the importance to practice of these rules the TAUian will oppose and, let us be clear, will do so on grounds of utility. In fine, if the TAUian is to equip himself with sound guides for practice, sound guides by which to have the best chance of doing the right or optimific thing, he will not

advocate adherence to rules which, by their very specificity, suffer diminished utility from the point of view of practice; and in this he remains true to his utilitarianism.

XII.4 Generality in rules and exceptions

One final point. How does the TAUian both allow his rules to have exceptions and yet preserve their generality? Preserve their generality he must, since in this consists their usefulness; but if exceptions are incorporated into them, they will soon cease to be general. There are two solutions to this problem. First, obviously, the TAUian can allow exceptions to his rules but not go on to build them into the rules, or into their statements or formulations. Therefore, 'Promises are to be kept' will be formulated thus even after a justified exception has been allowed. Second, the TAUian can allow his rules to be overridden without being modified.[8] Allowing them to be overridden does not require a change in their formulation, and it preserves their generality and thereby their usefulness; whereas to modify rules amounts to incorporating exceptions into them. In my opinion, nothing hangs(for our purposes) on whether we speak of a rule having justified exceptions or of a rule being overridden justifiably; for cases in which a rule may be overridden justifiably will be cases in which a rule may be broken or set aside justifiably.[9] However, I do not insist on the point, and am content to accept that the rules of the TAUian do not have

exceptions, though they may be overridden justifiably. For so far as I have been able to determine, nothing of importance to TAU is adversely affected by so doing.

## FOOTNOTES

### Chapter I

1. D. Lyons, Forms and Limits of Utilitarianism (Oxford, Clarendon Press, 1965).

2. cf. H. L. A. Hart, A. M. Honore, Causation in the Law (Oxford, Clarendon Press, 1959), pp. 64_78.

3. cf. J. J. Smart, An Outline of a System of Utilitarian Ethics (Melbourne, Melbourne University Press, 1961), pp. 22_23.

4. Ibid., p. 23.

5. cf. H. B. Acton, J. W. N. Watkins, 'Negative Utilitarianism' (A Symposium), PASS, vol. XXXVII, 1963, pp. 83_114.

6. R. N. Smart, 'Negative Utilitarianism', M, vol. 67, 1958, pp. 542_543.

7. J. Bentham, An Introduction to the Principles of Morals and Legislation, ed. J. Harrison (Oxford, Blackwell, 1948), I, iii.

8. J. S. Mill, Utilitarianism (Indianapolis, The Bobbs_Merill Co., 1957), chpt. 2.

9. H. Sidgwick, The Methods of Ethics, 7th ed. (London, Macmillan, 1907), chpt. 5, also Bk. IV.

10. W. T. Stace, The Concept of Morals (New York, Macmillan, 1937), chpt. 6.

11. H. Rashdall, Theory of Good and Evil, 2 vols. (Oxford, University Press, 1924), chpt. viii.

12. G. E. Moore, Principia Ethica (Cambridge, University Press, 1959), chpt. 3_5; Ethics (London, Oxford University Press, 1961), chpts. 1, 2, 7.

13. A discussion of some of these reasons, together with an exam_ination of different sorts of hedonism, is contained in J. C. B. Gosling's Pleasure and Desire (Oxford, Clarendon Press, 1969).

14. J. Dewey, Theory of Valuation (Chicago, University Press, 1939), pp. 26_29.

15. M. Beardsley, 'Intrinsic Value', PPR, vol. 26, 1965, 1_17.

16. An expansion of this sort was set out recently by R. M. Hare, in a paper read to a seminar on utilitarianism in Oxford during Michaelmas Term, 1970.

17. cf. R. B. Brandt, 'Towards a Credible Form of Utilitarianism' in Castaneda, H_N., Nakhnikian, G., (eds.), Morality and The Language of Conduct (Detroit, Wayne State University Press, 1963), pp. 129_130.

18. D. H. Hodgson, Consequences of Utilitarianism (Oxford, Clarendon Press, 1967), pp. 26_29.

19. There are many practice versions of RU about today, but the distinction between justifying a practice and justifying acts falling under the practice is usually traced to a paper by John Rawls on 'Two Concepts of Rules', PR, vol. LXIV, 1955, pp. 3_32. The distinction is often employed in the literature, though it has its critics; see A. I. Melden, 'Utility and Moral Reasoning' in DeGeorge, R. T., (ed.), Ethics and Society (New York, Doubleday, 1966), pp. 173_196, and G. Kerner, 'The Immorality of Utilitarianism and The Escapism of Rule_Utilitarianism', PQ, vol. 21, 1971, pp. 36_50.

20. cf. R. B. Brandt, 'Some Merits of One Form of Rule_Utilitarianism' in University of Colorado Studies, Series in Philosophy, No. 3 (Boulder, Colorado, University of Colorado Press, 1967), pp. 39_65.

21. Hodgson, op. cit., pp. 63_82.

22. For an interesting attack on the 'greatest happiness of the greatest number' formulation of the principle of utility, see F. Y. Edgeworth, Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences (London, C. Kegan Paul, 1881), pp. 117_118. I am indebted for this reference to Dr. J. Griffin.

23. The whole problem of interpersonal comparisons of utility crops us in economics, in the theory of collective choice; and the possibility of a solution to the problem there, and the subsequent extension of this solution to moral philosophy, forms both an interesting and important program, which has recently been initiated: see A. K. Sen, Collective Choice and Social Welfare (Edinburgh, Oliver & Boyd, 1970), esp. pp. 131_151.

24. The consequence component first emerges as the predominant element in the utilitarian theme in Moore's Ethics, chpt. 7 ('Results the Test of Right and Wrong').

25. See Chpt. II.1.

26. J. J. C. Smart, 'Extreme and Restricted Utilitarianism', PQ, vol. VI, 1956, p. 344.

27. The case for saying that some forms of deontology can have the same view about the connection between rightness and consequences is presented in Chpt. II.2.

Chapter Two

1. W. D. Ross, The Right and The Good (Oxford, Clarendon Press, 1930).

2. H. A. Prichard, Moral Obligation (Oxford, Clarendon Press, 1949), esp. his essay 'Does Moral Philosophy Rest Upon A Mistake?', pp. 1-17.

3. See Chpt. IV, note 2.

4. cf. O. A. Johnson, Rightness and Goodness (The Hague, Martinus Nijhoff, 1959).

5. Ross, op. cit., pp. 41-47.

6. By 'absolutism', I mean to denote the same position that Jonathan Bennett ('Whatever the consequences', A, vol. 26, 1966, 83-102) calls 'conservative'. Neither term is a happy one, for reasons Bennett makes clear.

7. Immanuel Kant, 'On a Supposed Right to Lie from Altruistic Motives', in Beck, L. W., (trans. & ed.), Kant's Critique of Practical Reason and Other Writings in Moral Philosophy (Chicago, University Press, 1949).

8. G. E. M. Anscombe, 'Modern Moral Philosophy', P, vol. 43, 1958, 1-19.

9. G. E. M. Anscombe, 'Mr. Truman's Degree', privately printed and circulated, on the occasion of President H. S. Truman being considered for an honorary degree by Oxford University in 1957.

10. P. T. Geach, 'The Moral Law and the Law of God', in God and the Soul (London, Routledge & Kegan Paul, 1969), pp. 117-129.

11. R. W. Beardsmore, 'Consequences and moral worth', A, vol. 29, 1969, 177-186.

12. Anscombe, op. cit., p. 19. Italics in original.

13. G. E. M. Anscombe, 'Does Oxford Moral Philosophy Corrupt Youth?', The Listener, vol. 57, 1957, 266-267, 271). See also the ensuing correspondence on pp. 311, 349, 387, 427, 478-479, 519, 520, 564.

14. Anscombe, op. cit., p. 10. Italics in original.

15. Ibid., p. 10.

16. See Chpts. IV, V and X for more on responsibility.

17. A. B. Cody, 'Can a single action have many different descriptions?', Inquiry, vol. 10, 1967, 164_180.

18. H. L. A. Hart, 'Intention and Punishment', The Oxford Review, no. 4, Hilary Term, 1967, pp. 5_22.

19. cf. A. J. A. Kenny, 'Intention and Purpose in Law' (revised) in Summers, R., (ed.), Essays in Legal Philosophy (Oxford, Blackwell, 1968), pp. 146_163. A useful critical discussion of this paper is to be found in H. Oberdiek, 'Intention and Fore_sight in Criminal Law', M, vol. LXXXI, 1972, 389_400.

20. By 'logically distinct', I mean 'logically independent', which I construe in this context to mean 'free from implication'. By 'logically dependent', I mean 'logically implies'.

21. cf. J. Narveson, 'Utilitarianism and Formalism', AJP, vol. 43, 1965, 58_72. I am indebted to Professor Narveson for discussing this paper with me.

22. cf. M. Stocker, 'Consequentialism and its Complexities', APQ, vol. 6, 1969, 276_289. I might add that I do not believe any_thing is gained by characterizing the dispute between I and C in terms of intrinsic and consequential properties, since it does not follow from the fact that p is an intrinsic property of act A that it may not also be a consequential property of act A. Briefly, on the scheme set out in this chapter, a property p will be intrinsic to any act only under a given description; therefore, if p is intrinsic to act A relative to a given des_cription, and if different descriptions of the same act may be given, then what is an intrinsic property of A under one des_cription may be a consequential property of A under another. It is true that, if p is intrinsic to A relative to this description, then it cannot also be a consequential property of A under this description; but if different descriptions of the same act can be given, then p can be a consequential property of A under another description.

23. cf. G. E. M. Anscombe, Intention (Oxford, Blackwell, 1957), pp. 37_47.

24. A similar conclusion has been reached by Andrew Oldenqvist in his paper on 'Rules and Consequences', M, vol. 75, 1966, 180_192.

25. cf. P. H. Nowell_Smith, Ethics (Oxford, Blackwell, 1957), p. 189; R. S. Downie, E. Telfer, Respect for Persons (London, George Allen & Unwin, 1969), pp. 40_47.

26. cf. A. G. N. Flew, 'End and Means' in Edwards, P., (ed.), The Encyclopedia of Philosophy (London, Macmillan, 1967), p. 281.

27. cf. C. D. Broad, Five Types of Ethical Theory (Cambridge, University Press, 1930), p. 207.

## Chapter Three

1. J. Ladd, 'The Ethical Dimensions of the Concept of Action', JP, vol. 62, 1965, p. 637. In a paper on 'Action, Consequences and Ethical Theory' (J. Value Inquiry, vol. 1, 1967, 72-79), Miss Ruth Macklin seems to hold a position similar to Ladd's. She writes: 'if it makes no difference where we draw the line between action and consequence, then consequentialist views are surely misconceived, since the same action may be evaluated as good or bad, right or wrong, or morally neutral, depending on what is to count as consequences'(p. 73). Thus, she appears to think that there is an 'act as such', independently of any and all descriptions, to which we 'fit' possible descriptions, some of which include more and some less of 'the act's' consequences. On my view, there is no 'act' independently of any and all descriptions, only a set of events; and Macklin's use of the term 'action' already indicates to me that she is referring to the set of events under some description or other. Under that description, they cannot be evaluated both right and wrong.

2. cf. J. Atwell, 'The Accordion-Effect Thesis', PQ, vol. 19, 1969, 337-342.

3. cf. H. A. Prichard, op cit., 'Duty and Ignorance of Fact', pp. 18-34. Views similar to Prichard's are discussed in E. D'Arcy's Human Acts (Oxford, Clarendon Press, 1963), pp. 4-10.

4. One of the more interesting and important attacks on Prichardian views of the will and willing is to be found in A. I. Melden's paper on 'Willing' (PR, vol. 69, 1960, 475-484). Prichard himself is aware of difficulties in his concept of 'setting oneself' to do; see his essay 'Acting, Willing, Desiring' in Moral Obligation, pp. 187-198.

5. A very stout defence indeed would be required for what is an implication of the Prichardian position, namely, that there is at best only a contingent connection between willing and overt behavior. For it to be true that what one did was to 'set oneself' to do, on need never, on Prichard's view, exhibit any overt behavior whatever, on this or on any occasion.

6. A. C. Danto, 'Basic Actions', APQ, vol. 2, 1965, 141-148.

7. Anscombe, Intention, pp. 37-45.

8. M. Cohen, 'The Same Action', PAS, vol. LXVII, 1969, pp. 80ff.

9. cf. L. Davis, 'Individuation of Actions', JP, vol. LXVII, 1970, 520_530.

10. cf. D. Sachs, ' A Few Morals About Acts', PR, vol. 75, 1966, 91_98.

11. D'Arcy, op. cit., p. 14.

12. cf. J. Margolis, 'Danto on Basic Actions', Inquiry, vol. 13, 1970, pp. 104_107.

13. J. Griffin, 'Consequences', PAS, vol. 75, 1964_65, p. 171. For a discussion of this paper, see P. H. Nowell_Smith, 'Utilitarianism and Treating Others as Ends', Nous, vol. 1, 1967, 81_90.

## Chapter Four

1. It is remarkable that contemporary Aists have not developed the following, interesting form of A. (An exception to this stricture may be found - I find the paper difficult to follow - in Thomas Nagel's discussion of 'War and Massacre' in Philosophy and Public Affairs, vol. 1, 1972, 123-144.) Briefly, it is possible to envisage an Aist who did not disregard consequences - i.e., who did not concern himself with the question of whether there are any acts wrong apart from their internal and/or external consequences - but rather who held that to produce certain consequences (e.g., deaths of innocents) was absolutely wrong, whatever the other consequences and circumstances of the act. Such an Aist could agree with the Cist that the difference between I and C was bogus; and he would in effect be an Aist who was also, in this limited sense, a Cist. This modified form of A will still be objectionable to a genuine Cist; for he will simply not agree that, if an act produces consequences of the alleged absolutely wrong variety, we need look at no other consequences of the act in order to be able to assess its rightness or wrongness. Perhaps these other consequences will significantly offset the bad ones. And, too, the examination of a number of different descriptions of the act will enable the Cist to assess the act's internal and external consequences under different descriptions, and so perhaps better able him to make a judgment about whether certain consequences do or do not offset these bad consequences.

2. G. Kelly, S. J., Medico-Moral Problems, revised ed.(Dublin, Clonmore & Reynolds, 1960), p. 13. Masturbation and sodomy are sometimes called 'unnatural' and held to be violations of natural law, which, in place of intuition, can provide the Aist with a means of identifying which kinds of acts are intrinsically wrong. It should not be thought - what is a common mistake - that the Aist who appeals to or cites natural law appeals to or cites revelation. cf. G. E. M. Anscombe, 'Contraception and Natural Law', New Blackfriars, vol. 46, 1965, 517-521; and especially J. Finnis, 'Natural Law and Unnatural Acts', The Heythrop Journal, vol. xI, 1970, 365-387.

3. Anscombe, op. cit., p. 17.

4. Ibid.

5. On this point, see J. Bennett, op. cit., p. 90. See also K. Nielsen, 'Against Moral Conservatism', Ethics, vol. 82, 1972, 218_231.

6. It may, of course, be claimed that the philosopher is more likely to commit judicial murder in an actual case, if he has contemplated it in an imaginative way; but this claim requires evidence to substantiate it.

7. J. Bennett, op. cit., p. 84.

8. Ross' prima facie duties, on the other hand, enable him to take into account complexities in the situation to hand, since on his view, in a situation in which two or more of the prima facie have a bearing, our absolute duty is that prima facie duty which, in these circumstances, is the more stringent.

9. In reply to a criticism of mine in respect of his paper on 'Bentham's Theory of Action and Intention' (unpublished), Dr. P. M. S. Hacker raised the following question for someone with my view of C: can two acts have similar consequences but different circumstances and be evaluated differently? On the view presented here, any change in act_description (to accommodate circumstances) will change (at least) the internal consequences of the act, so unless the two acts are described in the same way, they will not have the same consequences (whether the consequences are similar or not seems unimportant); and if they are described in the same way, and so have or may be said to have the same consequences, they will not be evaluated differently.

10. An interesting account of the interrelations of circumstances, consequences and act_descriptions is to be found in Bentham, op. cit., I, chpts. 7_10. I refer especially to Bentham's theory of the circumstances of acts as set out in Chapter 8; there is a discussion of this chapter _ with which I do not wholly agree _ and of Bentham's classification of circumstances in D'Arcy, op. cit., pp. 57_92.

11. J. Cargile, 'On consequentialism', A, vol. 29, 1969, 78_88. See also G. Wallace, 'Cargile on Utilitarianism', A, vol. 32, 1972, 126_131. Wallace is confused over the distinction between consequentialism and utilitarianism.

12. Ibid., p. 82.

13. Ibid., p. 87.

14. At one point, Cargile makes explicit reference to feelings: Ibid., p. 85.

15. E. Healy, S. J., Medical Ethics (Chicago, Loyola University Press, 1956), p. 196. Cardinal Newman has a striking passage on the matter; the Catholic Church, he says, 'holds it better for sun and moon to drop from heaven, for the earth to fail, and for all the many millions on it to die of starvation in extremest agony, as far as temporal affliction goes, than that one soul, I will not say should be lost, but should commit one single venial sin, should tell one wilful untruth, or should steal one poor farthing without excuse' (from Difficulties for Anglicans, quoted by M. J. Svaglic, in his Introduction to J. H. Newman's Apologia Pro Vita Sua (Oxford, Clarendon Press, 1967, p. L). Such a passage is apt to lead one to reflect that Kingsley would have done better to bring out and attack Newman's absolutism.

16. J. H. Newman, 'Lying and Equivocation', in Apologia Pro Vita Sua (Oxford, Clarendon Press, 1967), ed. M. J. Svaglic, Note G, pp. 302-311.

## Chapter Five

1. cf. P. Foot, 'Abortion and Double Effect', Oxford Review, No. 5, Trinity Term, 1967, p. 10.

2. cf. H. L. A. Hart, op. cit. According to Hart, an oblique intention refers to the 'mere foresight of consequences', whereas in the case of a direct intention 'the consequences must have been contemplated by the accused not merely as a foreseen outcome but as an end which he set out to achieve or as a means to an end, or constituted at least part of his reason for doing what he did' (pp. 10_11).

3. J. Bennett, op. cit., p. 83.

4. cf. Holy Office Decree, May 28, 1884. This decree sets out clearly the ban on craniotomy. This ban is reiterated in the Holy Office Decree of August 19, 1889, and it is widened to include any operation which involves a direct attack upon the life of the foetus.

5. The use of 'permits' in these contexts appears standard practice amongst Catholic thinkers; see G. Kelly, S. J., op cit., pp. 73, 75; and E. Healy, S. J., op. cit., p. 196.

6. cf. A.Fagothy, S. J., Right and Reason (St Louis, C. V. Mosby Co., 1959), p. 154; S. Windass, 'Double Think and Double Effect', Blackfriars, vol. 44, 1963, p. 257; W. Conway, 'The Act of Two Effects', Irish Theological Quarterly, vol. XVIII, 1951.

7. J. Bennett, op. cit.

8. H. L. A. Hart, op. cit., p. 13.

9. Ibid., pp. 13_14.

10. On the view of A mentioned in Chapter IV, note 1, it might be argued that DDE introduces an exception to a doctrine about rightness and consequences. A Cist will deny the point.

11. cf. J. Feinberg, 'On Being "Morally Speaking" a Murderer"', JP vol. 60, 1963.

12. G. Williams, The Sanctity of Life and The Criminal Law (London, Faber & Faber, 1958), p. 286. For an astonishingly bad critique of this work from a Catholic point of view, see C. B. Daly, Morals, Law and Life (Dublin, Clonmore & Reynolds, 1962).

13. cf. J. Ford, S. J., 'The Morality of Obliteration Bombing', Theological Studies, vol. 5, 1944, 261-309.

14. The case for Williams' view is not made conclusively by pointing to a difficulty about withholding intention from consequences foreseen as certain; subsequent considerations here, however, support his case.

15. Director of Public Prosecutions v. Smith [1961] A. C. 290. Driving a car containing stolen property, Smith was told to pull over by a policeman, whereupon Smith accelerated and drove away, with the policeman clinging to the car. The car hit three others, after which the policeman fell off, directly in front of a fourth, and received fatal injuries. In essence, three things were at issue in the Smith case: (a) the maxim that a man intends the natural and probable consequences of his acts, (b) the maxim (law) that if a man intentionally inflicts grievous bodily harm, then if death results, he is guilty of murder, and (c) objective and subjective tests of intention. It was agreed by all that Smith did not directly intend to kill the policeman, but as a result of the objective test of intention, the jury (in the courst of first instance) was instructed that it could reasonably consider Smith to have had this intention - since a reasonable man would in the circumstances have foreseen the likelihood of grievous bodily harm befalling the policeman. (This anomaly brought out another peculiar aspect to the case: while everyone agreed that Smith did not directly intend to kill the policeman, they were also agreed that Smith did intend to shake the policeman from the car, and drove in such a way as to bring this about; and if the maxim that one is guilty of murder if the intentional infliction of grievous bodily harm results in death is applied, then Smith is guilty or murder - in spite of the fact that he did not intend to kill the policeman.) The Law Commission (Imputed Criminal Intent, D. P. P. v. Smith, London, Her Majesty's Stationery Office, 1967) has now come out in favor, not of an objective test of intent, i.e., one which raises the question of whether a reasonable man in the circumstances would or would not have foreseen consequence X as a consequence of his act; but of a subjective test, which concerns itself with the actual state of mind of the accused, and not with any hypothetical state of mind of any hypothetical reasonable man.

16. A. J. A. Kenny, 'Intention and Purpose', JP, vol. LXIII, 1966, p. 647.

17. Kenny does not hold this view. '...Consequences which are foreseen but which the agent is indifferent to or rejects are not intended. Of course, it may well be correct to hold the agent responsible for these consequences, but that only means that we can be held responsible for more than we intend' (Ibid., pp. 648_649).

18. G. Kelly, S. J., op. cit., p. 75.

19. cf. E. D'Arcy, op. cit., pp. 40_57.

20. cf. F. Siegler, 'Omissions', A, vol. 28, 1968, 98_106. See also D. Sachs, op. cit., pp. 93_95.

21. The same reasoning will hold if it is the mother who is saved.

22. G. E. M. Anscombe, 'Who is Wronged?', Oxford Review, Trinity Term, No. 5, 1967, pp. 16_17.

23. If a doctor treats each individual life as valuable, and if he thinks _ what he is taught _ that no individual life is beneath his craft, then in a case where he can choose between saving five or fifty persons, number will make a difference to him. For if he regards each life as valuable, then the saving of fifty valuable lives is preferable to the saving of five. To say simply that the doctor's duty is 'to save life', so that in saving the five he will have done his duty, can mask the effects of the doctor's training and of what he has been taught on his decision whether to save five of fifty. I do not wish to imply that what is his duty depends on what he has been taught; only that what he has been taught will bear on how he interprets his duty in particular cases.

24. cf. P. J. FitzGerald, 'Acting and Refraining', A, vol. 27, 1967, 133_139; J. Bennett, 'Acting and refraining', A, vol. 28, 1967, 30_31.

25. G. H. von Wright treats omissions as either identical with or a species of forbearances (Norm and Action, London, Routledge & Kegan Paul, 1963, pp. 45_54). If, and in so far as, to forbear doing X is identical with or a species of refraining from doing X, my remarks here support a difference between omitting and re_fraining and between omitting and forbearing.

26. It may not be the case that a difference between refraining and omitting is also a difference between refraining and immorally omitting.

27. Of course, all sorts of cases can be envisaged, where what the doctor takes into account horrifies us; but in these, as in all cases, we rely upon the doctor's good judgment and, in the last analysis, upon his character. The fact that the foetus will almost certainly die anyway cannot be excluded as irrelevant to the doctor's decision, unless, once again, the Aist is prepared to ignore the circumstances in which an act is contemplated and performed.

28. P. Foot, op. cit. By a positive duty, Mrs. Foot means such things as looking after children or aged parents, though she extends the notion a bit to cover acts of charity; whereas by a negative duty, she means such things as refraining from killing and robbing.

29. She notes of an example that 'one does not in general have the same duty to help people as to refrain from injurying them' (Ibid., p. 12, italics in original). She goes on: 'it is interesting that, even where the strictest duty of positive aid exists, this still does not weigh as if a negative duty were involved' (Ibid., p. 12).

30. Ibid., p. 13.

31. In so far as this claim rests upon attaching moral significance to the line between act and consequences, I reject it for reasons given in Chapter II.

32. cf. Decisions About Life and Death, Appendix 2 (Westminster, Church Information Office, 1966), pp. 49-51.

33. cf. Winfield on Tort, 7th ed., Jolowicz, J. A., Lewis, T. E., (eds.), (London, Sweet and Maxwell, 1963), pp. 49-53. It should perhaps be added that where a plea of necessity has been successfully entered in civil cases, it is not altogether clear whether or not compensation or restitution are demandable. cf. Southport Corporation v. Esso Petroleum Co. Ltd. [1956] A. C. 218, 227.

34. cf. Cross, R., Jones, P. A., An Introduction to Criminal Law, 5th ed. (London, Butterworths, 1964), pp. 80-82; see also Turner, J. W. C., (ed.), Kenny's Outlines of Criminal Law, 17th ed. (Cambridge, University Press, 1958), pp. 66-68.

35. R. v. Dudley and Stephens [1884] 14 Q. B. D. 273.

36. If one objects to speaking of the 'interests of the foetus' - on the ground that non-persons can have no interests - we may speak instead of the interests of the person the foetus will turn into.

## Chapter Six

1. What counts as an alternative available to us and how we evaluate the consequences of alternative courses of action are questions I shall not consider. cf. Lars Bergstrom, The Alternatives and Consequences of Actions (Stockholm, Almqvist & Wiksell, 1966).

2. W. D. Ross, op. cit., pp. 16ff.

3. One thing critics nearly always fail to mention in these sorts of cases is whether Smith and Jones are AUians. If so, and if each knows the other is, then the example seems to have no point, for then Smith will not elicit a promise to do something which fails to maximize utility; and if Smith at the time did not know that giving the money to the jockey club will not maximize utility, then Jones may reasonably infer, from his knowledge that Smith was an AUian, that Smith would not want his money given to the jockey club unless utility were thereby maximized. On the other hand, if Jones is an AUian but Smith is not, and Jones knows that Smith is not, will Jones give the sort of promise asked for? (He might, if he thinks it will keep Smith happy. But this raises a question as to the status of 'consoling' promises.)

## Chapter Seven

1. See Chpt. IV.2.

2. Ibid.

3. We may be reluctant, of course, to mention exceptions to rules if we know or suspect that the officer has already begun to reconcile himself to torturing the man.

4. G. E. Moore, Principia Ethica (Cambridge, University Press, 1959).

5. Ibid., p. 162.

6. For example, if one lived in a community in which many people lacked the normal human feelings of sympathy and benevolence and so would not behave kindly to animals, the point to having a rule and sanctions for breaking it would be to ensure that people in this community behaved in ways that they 'naturally' would not.

7. There is a stronger argument against this view; see Chpt. VIII.

8. There is a more serious objection to UAU here, viz., that the case by case procedure, together with reliance upon optimificity to justify exceptions, will undermine the institution of promising. See Chpt. VIII.

9. cf. D. Lyons, op. cit., chpt. 3.

10. G. E. Moore, op. cit., pp. 85-105, 162.

11. cf. B. Mitchell, Law, Morality and Religion (Oxford, Clarendon Press, 1967), chpt. 2.

12. H. L. A. Hart, The Concept of Law (Oxford, Clarendon Press, 1961), pp. 163-176.

13. For an excellent, and beautifully written, account of this position, see G. J. Warnock, The Object of Morality (London, Methuen, 1971), chpts. 4 & 5.

14. I am not here concerned with whether the rules of tennis also constitute that activity, though it seems clear to me that they do not.

15. R. A. Wasserstrom, The Judicial Decision (Stanford, University Press, 1969). See especially Chpt. 6, 'Extreme and Restricted Utilitarianism', and Chpt. 7, 'The Two_Level Procedure of Justification'. There is a general point about legal rules making acts right or wrong worth mentioning here. A rule in a zoo proscribing the feeding of animals does not make feeding them wrong; what makes feeding them wrong is, e.g., that they may suffer indigestion, or that they may be fed harmful sub_ stances or substances not good for their metabolism. Likewise, a legal rule prescribing the paying of taxes makes it illegal not to pay them; but it does not make not paying them morally wrong. Somewhat pedantically, it might be held that it is morally wrong not to pay taxes because it is illegal, if there were a rule or principle to the effect that illegal acts, sub_ ject to certain exceptions, are morally wrong.

## Chapter Eight

1. D. H. Hodgson, Consequences of Utilitarianism (Oxford, Clarendon Press, 1967), chpt. 2, pp. 38_62.

2. I am indebted to Professor Henry R. West for allowing me to read and for discussing with me his unpublished paper, 'The Consequences of Act_Utilitarianism', in which he attacks Hodgson's argument. I am also indebted to Professor J. Narve_son for allowing me to read his unpublished paper on 'D. H. Hodgson's Critique of Utilitarianism', a much shortened, revised version of which _ entitled 'Promising, Expecting and Utility' _ he read to the Philosophical Society in Oxford during Michaelmas Term, 1970.

3. Hodgson construes the 'act_utilitarian principle' as follows: 'an act is right if and only if it would have best consequences, that is, consequences at least as good as those of any alter_native act open to the agent' (op. cit., p. 1).

4. P. Singer, 'Is Act_Utilitarianism Self_Defeating?', PR, vol. LXXXI, 1972, 94_100. I am grateful to Mr. Singer for allowing me to read and discussing with me an earlier draft of this paper.

5. Hodgson, op. cit., p. 39.

6. Ibid., p. 40.

7. cf. D. Lyons, op. cit., p. 74.

8. Strictly speaking, I should say that the practice of UAU promise_breaking by a very large number of individuals over a certain interval at a certain frequency will have a damaging effect on the institution; these considerations are assumed in my argument. cf. Lyons, op. cit., pp. 69_72.

9. Marcus Singer, Generalization in Ethics (London, Eyre & Spottis_woode, 1963), p. 67.

10. R. F. Harrod, 'Utilitarianism Revised', M, vol. XLV, 1936, p.148. Italics in original.

11. Following Lyons, I shall mean by 'practice' in this context roughly 'performance'; cf. D. Lyons, op. cit., pp. 72_75.

12. Ibid., p. 65. Lyons rejects this thesis in favor of a thesis of linearity; that is, he holds the view that the effects of an aggregate of acts of a certain kind is always the sum of the effects of each individual act that comprises the aggregate. In the next chapter, I examine these two theses.

Chapter Nine

1. D. Lyons, Forms and Limits of Utilitarianism (Oxford, Clarendon Press, 1965). Especially important for our purposes is Chapter Three, 'Extensional Equivalence'.

2. Ibid., p. 73. In general, there will be a threshold, beyond which threshold effects ensue, for any cooperative enterprise in which the benefits of the enterprise accrue through general participation in it.

3. This 'causal property' of the act, the extent to which others are performing similar acts, is sometimes called by Lyons the 'social context' of the act (Ibid., pp. 63_64). This use of 'causal property' is a strange or unusual one, a point to which I return in Chapter X.

4. For a clear statement of the thesis of non_linearity, see the quotation from R. F. Harrod above, Chpt. VIII.3. To repeat, the thesis of linearity is the view that the effects of an aggregate of acts of a particular kind is always the sum of the effects of each individual act that constitutes the aggre_ gate.

5. Lyons, op. cit., p. 63.

6. Ibid., p. 72. Lyons uses the term 'contributes' (Ibid., pp. 71_ 72).

7. Ibid., p. 115.

8. Ibid., p. 57. Italics in original. For the argument in full: Ibid., pp. 56_61.

9. Ibid., p. 77. This point about linearity and the performance of acts within a practice of similar acts is never explicitly stated by Lyons, but then it is one of the bases for the argu_ ment for extensional equivalence, and as such, is discussed indirectly throughout his presentation of that argument in Chapter Three.

10. See Chpt. VIII.2.

Chapter Ten

1. Its cogency has already been attacked, in different ways, by Gertrude Ezorsky in her review of Lyons' book, JP, vol. LXV, 1968, 533_544; and by J. H. Sobel, in his paper 'Utilitarianism: Simple and General', Inquiry, vol. 13, 1970, 394_449. Sobel's paper in particular is rewarding, for he develops what he calls an 'avoidable' theory of consequences which, included in a species of utilitarian generalization, does not yield a theory that is extensionally equivalent to any simple form of AU.

2. I hasten to add that the TAUian may or may not be a non_linearian.

3. Lyons, op. cit., p. 69.

4. Ibid., p. 69.

5. Harrod, op. cit., p. 148.

6. J. Harrison, review of Lyons' book, Philosophical Books, vol. VII, 1966, p. 14. I am indebted to Professor Henry R. West for this reference.

7. I am indebted to Professor Henry R. West, who kindly allowed me to read his Ph.D. thesis 'Act_Utilitarianism and Rule_Utilitarianism' (Harvard University, 1964, unpublished) and discussed sections of it with me on a number of occasions. What in this section I call 'contributive consequences' is discussed by West under the heading of 'The Conributory Theory of Causation', but our discussions reach opposite conclusions: whereas West thinks a viable form of UAU may be grounded on a contributive theory of causation, I argue that, even granted this ground, UAU falls to our utilitarian argument.

8. G. E. Moore, 'A Reply To My Critics', in Schilpp, P. A., (ed.), The Philosophy of G. E. Moore (Evanston, Illinois, 1942), p. 559. Italics in original.

9. If B knew that A had already exploded its bomb, but nevertheless explodes its own, then B, knowing its act to be sufficient to bring on pollution, intentionally pushes us over the threshold. Judgments about the moral worth of those in B who did the de_ciding may be founded on this question of intention. I return to this case below.

10. Lyons, op. cit., pp. 76_80.

11. Ibid., p. 76.

12. Lyons writes: 'A causally homogeneous class is one the members of which are exactly similar in respect of their threshold_related effects as well as any others' (Ibid., p. 77).

13. The narrow view is helpful in prospect in this sequential case, in which the agent knows his place in the sequence; so far as I can see, however, it is <u>not</u> helpful in prospect in non_sequential cases or in sequential cases in which the agent does not know his place in the sequence and/or how previous agents have acted.

## Chapter Eleven

1. Philosophers have been especially negligent in often failing to appreciate the institution's role in helping to achieve coordination of actions, plans, policies, decisions, etc., and so in helping to achieve that degree of cooperation we find so essential. See Chpt. XII.

2. If one thinks the institution is constituted by rules, then the point here might be expressed thus: since utility demands that we support the institution, it is logically inconsistent not to accept, subscribe or commit oneself to its rules. I prefer not to speak in this way, however, because of confusions that seem rampant in giving meaning to the expressions 'accept', 'subscribe' and 'commit', particularly when some sort interior act is introduced.

3. I do not consider the case of an opponent who raises the question, 'Why bother about being consistent?' Such an opponent, if there are any, raises a question that would plague any theorist, <u>whether</u> utilitarian or not, at least if his question were construed as asking for reasons for being consistent. We might point out, however, that lack of consistency in the case of the institution of promising places the utility of the institution in peril, i.e., one accepts as justificatory for exceptions grounds the general practice of which <u>will</u> impair confidence in the institution. But I suspect this will not satisfy such an opponent.

4. Instead of allowing all non-blacks the entitlement to cite black skin as a justificatory ground, White may argue that only those like himself, e.g., other whites, are entitled to exceptions on this ground; but it would still follow, even if all whites did not keep promises to blacks, that the utility of the institution would not be impaired. The crucial question here is the one I go on to raise, viz., how is it that some occupants of a role possess entitlements that other occupants of the same role do not possess?

5. cf. R. M. Hare, <u>Freedom and Reason</u> (Oxford, Clarendon Press, 1963), chpts. 6 & 8; see also Hare's review of Singer's <u>Generalization in Ethics</u>, <u>PQ</u>, 1962, 351-355.

6. Of course, White may go on to say that he does not give a rap for the utility of the institution. But this tactic is of far less interest to a utilitarian concerned with practice than the case in which the individual supports the institution but wants <u>this</u> exception to it.

## Chapter Twelve

1. Narveson, op. cit., p. 17, chpts. 5 & 6.

2. cf. G. J. Warnock, op. cit., chpt. 3.

3. Our discussion of Hodgson's argument is relevant to this point; see Chpt. VIII.1.

4. Thomas C. Schelling, The Strategy of Conflict (Cambridge, Mass., Harvard University Press, 1960). In conversation, David Lewis has suggested to me that his discussion of conformative behavior supports the view that coordination may be achieved without rules; see D. K. Lewis, Convention (Cambridge, Mass., Harvard University Press, 1969), pp. 107_118.

5. Lyons, op. cit., chpt. IV.

6. See J. Rawls, 'Two Concepts of Rules', PR, vol. LXIV, 1955, 8ff.

7. See Lyons, op. cit., chpts. 3 & 4; Hare, Freedom and Reason, chpt. 7, and his paper on 'Rules of War and Moral Reasoning', Phil. & Public Affairs, vol. 1, 1972, 166_181.

8. The use to which this distinction might be put by a TAUian developed out of discussion of the distinction with R. M. Hare.

9. If one wishes to insist that allowing rules to be overridden amounts to not allowing them to have exceptions, which in turn amounts to regarding them as exceptionless, I can at present see no reason to alter anything in the statement of TAU if the point is accepted.

# BIBLIOGRAPHY


Acton, H. B. 'Negative Utilitarianism' (in symposium with J. W. N. Watkins), Proceedings of the Aristotelian Society, supplementary vol. 37, 1963, 83_114.

Anscombe, G. E. M. 'Does Oxford Moral Philosophy Corrupt Youth', The Listener, vol. 57, 1957, 266_267, 271. (Ensuing correspondence: pp. 311, 349, 387, 427, 478_479, 519, 520, 564.)

——— Mr. Truman's Degree. Privately printed.

——— Intention. Oxford, Blackwell, 1957.

——— 'Modern Moral Philosophy', Philosophy, vol. 43, 1958, 1_19.

——— 'Contraception and Natural', New Blackfriars, vol. 46, 1965, 517_521.

——— 'War and Murder', in Stein, Walter, (ed.), Nuclear Weapons and Christian Conscience (London, Merlin Press, 1965), pp. 45_62.

——— 'Who is Wronged?', The Oxford Review, no. 5, Trinity Term, 1967, 16_17.

Atwell, J. E. 'The Accordion_Effect Thesis', Philosophical Quarterly, vol. 19, 1969, 337_42.

Baier, K. 'Acting and producing', Journal of Philosophy, vol. LXII, 1965, 645_648.

Barnes, G. 'Utilitarianisms', Ethics, vol. 82, 1971, 56_64.

Bayles, M. D. (ed.) Contemporary Utilitarianism. New York, Doubleday, 1968.

——— 'A Rule_Utilitarian Moral Code', Journal of Value Inquiry, vol. III, 1969, 258_269.

Beardsley, M. C. 'Intrinsic Value', Philosophy and Phenomenological Research, vol. 26, 1965_66, 1_17.

Beardsmore, R. W. 'Consequences and moral worth', Analysis, vol. 29, 1969, 177_186.

——— Moral Reasoning. London, Routledge, 1969.

Bennett, J. 'Whatever the consequences', Analysis, vol. 26, 1966, 83_102.

——— 'Acting and refraining', Analysis, vol. 28, 1967, 30_31.

Bentham, J. An Introduction to the Principles of Morals and Legislation, ed. J. Harrison. Oxford, Blackwell, 1948.

Bergstrom, L. The Alternatives and Consequences of Actions. Stockholm, Almqvist & Wiksell, 1966.

Berrill, J. The Person in the Womb. New York, Dodd, Mead, & Co., 1968.

Blum, R. P. 'The True Function of the Generalization Argument', Inquiry, vol. 13, 1970, 274_288.

Bradley, F. H. 'Pleasure for Pleasure's Sake' in Ethical Studies (Oxford, Clarendon Press, 1956), pp. 85_141.

Brandt, R. B. 'Towards a Credible Form of Utilitarianism' in Castaneda, H_N, Nakhnikian, G., (eds.), Morality and the Language of Conduct (Detroit, Wayne State University Press, 1963), pp. 107_144.

——— 'Some Merits of One Form of Rule_Utilitarianism' in University of Colorado Studies, Series in Philosophy, no. 3. (Boulder, Colorado, University of Colorado Press, 1967), pp. 39_65.

——— 'Utilitarianism and the Rule of War', Philosophy and Public Affairs, vol. 1, 1972, 145_165.

Broad, C. D. 'The Doctrine of Consequences in Ethics', Ethics, vol. XXIV, 1913_14, 293_320.

——— Five Types of Ethical Theory. Cambridge, Cambridge University Press, 1930.

Brody, B. A. 'The Equivalence of Act and Rule_Utilitarianism', Philosophical Studies, vol. 18, 1967, 81_87.

——— (ed.) Moral Rules and Particular Circumstances. Englewood Cliffs, N.J., Prentice Hall, 1970.

——— 'Abortion and the Law', Journal of Philosophy, vol. LXVIII, 1971, 357_369.

Cargile, J. 'On consequentialism', Analysis, vol. 29, 1969, 78_88.

Casey, J. 'Actions and Consequences' in Casey, J., (ed.), Morality and Moral Reasoning (London, Methuen, 1971), pp. 155_205.

Chopra, Y. N. 'The Consequences of Human Actions', Proceedings of the Aristotelian Society, vol. LXV, 1964_65, 147_166.

Church of England. Decisions About Life and Death. Westminster, Church Information Office, 1965.

Cody, A. B. 'Can a single action have many different descriptions?', Inquiry, vol. 10, 1967, 164_180.

Cohen, M. 'The Same Action', Proceedings of the Aristotelian Society, vol. LXVII, 1969.

Conway, W. 'The Act of Two Effects', Irish Theological Quarterly, vol. XVIII, 1951.

Crombie, I. M. 'Social Clockwork and Utilitarian Morality' in Mackinnon, D. M., (ed.), Christian Faith and Communist Faith (London, Macmillan, 1953), pp. 98_113.

Cross, R., Jones, P. A. An Introduction to Criminal Law, 5th ed. London, Butterworths, 1964.

——— 'The Mental Element in Crime', Criminal Law Review, 1967, 215_228.

Daly, C. B. Morals, Law and Life. Dublin, Clonmore & Reynolds, 1962.

Danto, A. C. 'Basic Actions', American Philosophical Quarterly, vol. 2, 1965, 141_148.

D'Arcy, E. Human Acts. Oxford, Clarendon Press, 1963.

Davis, L. H. 'Individuation of Actions', Journal of Philosophy, vol. LXVII, 1970, 520_530.

Dewey, J. Theory of Valuation. Chicago, University of Chicago Press, 1939.

Diggs, B. J. 'Rules and Utilitarianism', American Philosophical Quarterly, vol. 1, 1964, 32_44.

Downie, R. S., Telfer, E. Respect For Persons. London, George Allen & Unwin, 1969.

Dworkin, R. W. 'The Model of Rules', University of Chicago Law Review, vol. 35, 1967_68, 14_61.

Edgeworth, F. Y. Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences. London, C. Kegan Paul, 1881.

Ezorsky, G. 'Utilitarianism and rules', Australasian Journal of Philosophy, vol. 43, 1965, 225_229.

_____ Review of David Lyons' Forms and Limits of Utilitarianism, Journal of Philosophy, vol. LXV, 1968, 533_544.

Fagothey, A., S. J. Right and Reason, 6th ed. St. Louis, C. V. Mosby Co., 1959.

Feinberg, J. 'On Being "Morally Speaking A Murderer"', Journal of Philosophy, vol. 60, 1963.

Finnis, J. 'Natural Law and Unnatural Acts', The Heythrop Journal, vol. XI, 1970, 365_387.

FitzGerald, P. J. 'Acting and refraining', Analysis, vol. 27, 1967, 133_139.

Flathman, R. E. 'Forms and Limits of Utilitarianism', Ethics, vol. LXXV, 1966, 309_317.

Flew, A. G. N. 'Ends and Means' in Edwards, P., (ed.), The Encyclopedia of Philosophy (London, Macmillan, 1967), vol. 3, p. 281.

Foot, P. R. 'Abortion and Double Effect', The Oxford Review, no. 5, Trinity Term, 1967, 5_15.

Ford, J. C., S. J. 'The Morality of Obliteration Bombing', Theological Studies, vol. 5, 1944, 261_309.

Garner, R. T. 'Some Remarks on Act_Utilitarianism', Mind, vol. LXXVIII, 1969, 124_128.

Geach, P. T. 'The Will', Newman Bulletin of the Philosophy of Science, January and July, 1958.

_____ 'The Moral Law and the Law of God' in God and the Soul (London, Routledge, 1969), pp. 117_129.

Gibbard, A. F. 'Rule utilitarianism: Merely an illusory alternative', Australasian Journal of Philosophy, vol. 43, 1965, 211_220.

Gosling, J. C. B. Pleasure and Desire. Oxford, Clarendon Press, 1969.

Griffin, J. 'Consequences', Proceedings of the Aristotelian Society, vol. LXV, 1964_65, 167_182.

Hacker, P. M. S. 'Bentham's Theory of Action and Intention', un_published.

Hare, R. M. Freedom and Reason. Oxford, Clarendon Press, 1963.

_____ Review of Marcus Singer's Generalization in Ethics, Philosophical Quarterly, vol. 12, 1963, 351_355.

_____ 'The Argument From Received Opinion' in Essays in Philosophical Method (London, Macmillan, 1971), pp. 117_135.

_____ 'Rules of War and Moral Reasoning', Philosophy and Public Affairs, vol. 1, 1972, 166_181.

Harrison, J. Review of David Lyons' Forms and Limits of Utilitar_ianism, Philosophical Books, vol. VII, 1966, 13_15.

Harrod, R. F. 'Utilitarianism Revised', Mind, vol. XLV, 1936, 137_156.

Hart, H. L. A., Honoré, A. M. Causation In The Law. Oxford, Clarendon Press, 1959.

_____ The Concept of Law. Oxford, Clarendon Press, 1961.

_____ 'Intention and Punishment', The Oxford Review, no. 4, Hilary Term, 1967, 5-22.

Healy, E., S. J. Medical Ethics. Chicago, Loyola University Press, 1956.

Hodgson, D. H. Consequences of Utilitarianism. Oxford, Clarendon Press, 1967.

Johnson, O. A. Rightness and Goodness. The Hague, Martinus Nijhoff, 1959.

Jolowicz, J. A., Lewis, T. E. (eds.) Winfield On Tort. London, Sweet & Maxwell, 1963.

Jones, D. H. 'Making and Keeping Promises', Ethics, vol. LXXV, 1966, 287-296.

Kant, I. 'On a Supposed Right to Lie from Altruistic Motives' in Beck, L. W., (trans. & ed.), Kant's Critique of Practical Reason and Other Writings in Moral Philosophy (Chicago, University of Chicago Press, 1949).

Kelly, G., S. J. Medico-Moral Problems. London, Burns Oates Ltd., 1960.

Kenny, A. J. A. 'Intention and Purpose', Journal of Philosophy, vol. LXIII, 1966, 642-651. Reprinted with revisions in Summers, R., (ed.), Essays in Legal Philosophy (Oxford, Blackwell, 1968), pp. 146-163.

Kerner, G. C. 'The Immorality of Utilitarianism and The Escapism of Rule-Utilitarianism', Philosophical Quarterly, vol. 21, 1971, 36-50.

Ladd, J. 'The Ethical Dimensions of the Concept of Action', Journal of Philosophy, vol. LXII, 1965, 633-645.

Law Commission. Imputed Criminal Intent. Director of Public Prosecutions v. Smith. London, Her Majesty's Stationery Office, 1967.

Lewis, D. K. Convention. Cambridge, Mass., Harvard University Press, 1969.

Lloyd_Thomas, D. A. 'Consequences', Analysis, vol. 28, 1969, 133_141.

Lyons, David. Forms and Limits of Utilitarianism. Oxford, Clarendon Press, 1965.

Mabbott, J. D. Proceedings of the British Academy, vol. 39, 1953, 97_118.

Macklin, R. 'Actions, Consequences and Ethical Theory', Journal of Value Inquiry, vol. 1, 1967, 72_79.

Margolis, J. 'Danto on Basic Actions', Inquiry, vol. 13, 1970, 104_107.

Melden, A. I. 'Willing', Philosophical Review, vol. 69, 1960, 475_484.

_____ 'Utility and Moral Reasoning' in DeGeorge, R. T., (ed.), Ethics and Society (New York, Doubleday, 1966), pp. 173_196.

Mill, J. S. Utilitarianism. Indianapolis, The Bobbs_Merrill Co., 1957.

Miller, L. 'Rules and Exceptions', Ethics, vol. LXVI, 1956, 262_270.

Mitchell, B. Law, Morality and Religion. Oxford, Clarendon Press, 1969.

Moore, G. E. 'A Reply To My Critics' in Schilpp, P. A., (ed.), The Philosophy of G. E. Moore (Evanston, Northwestern University Press, 1942), pp. 535_677.

_____ Principia Ethica. Cambridge, Cambridge University Press, 1959.

_____ Ethics. London, Oxford University Press, 1961.

McCloskey, H. J. Meta_Ethics and Normative Ethics. The Hague, Martinus Nijhoff, 1969.

Nagel, T. 'War and Massacre', Philosophy and Public Affairs, vol. 1, 1972, 123_144.

Narveson, J. 'Utilitarianism and Formalism', Australasian Journal of Philosophy, vol. 43, 1965, 58_72.

_____ Morality and Utility. Baltimore, The Johns Hopkins Press, 1967.

_____ 'D. H. Hodgson's Critique of Utilitarianism'. Unpublished.

Newman, J. H. Apologia Pro Vita Sua, ed. M. J. Svaglic. Oxford, Clarendon Press, 1967.

Nielsen, K. 'Against Moral Conservatism', Ethics, vol. 82, 1972, 218_231.

Nowell_Smith, P. H. Ethics. Oxford, Blackwell, 1957.

_____ 'Utilitarianism and Treating Others as Ends', Nous, vol. 1, 1967, 81_90.

Oberdiek, H. 'Intention and Foresight in Criminal Law', Mind, vol. LXXXI, 1972, 389_400.

Oldenqvist, A. 'Rules and Consequences', Mind, vol. 75, 1966, 180_192.

Olson, M. The Logic of Collective Action. Cambridge, Mass., Harvard University Press, 1965.

Prichard, H. A. Moral Obligation. Oxford, Clarendon Press, 1949.

Prior, A. 'The Consequences of Actions' (in symposium with D. D. Raphael), Proceedings of the Aristotelian Society, vol. 30, 1956, 91_99.

Rachels, J. 'On Moral Absolutism', Australasian Journal of Philo_ sophy, vol. 48, 1970, 338_353.

Rashdall, H. Theory of Good and Evil, 2 vols. Oxford, Oxford University Press, 1924.

Rawls, J. 'Two Concepts of Rules', Philosophical Review, vol. LXIV, 1955, 3_32.

Ross, W. D. The Right and The Good. Oxford, Clarendon Press, 1930.

Sachs, D. 'A Few Morals About Acts', Philosophical Review, vol. 75, 1966, 91_98.

Schelling, T. C. The Strategy of Conflict. London, Oxford University Press, 1969.

Sen, A. K. Collective Choice and Social Welfare. Edinburgh, Oliver & Boyd, 1970.

Sidgwick, H. The Methods of Ethics, 7th ed. London, Macmillan, 1907.

Siegler, F. A. 'Omissions', Analysis, vol. 28, 1968, 98_106.

Singer, M. G. Generalization in Ethics. London, Eyre & Spottiswoode, 1963.

Singer, P. 'Is Act_Utilitarianism Self_Defeating?', Philosophical Review, vol. LXXXI, 1972, 94_104.

Smart, J. J. C. 'Extreme and Restricted Utilitarianism', Philosophical Quarterly, vol. VI, 1956, 344_354.

_____ An Outline Of A System Of Utilitarian Ethics. Melbourne, Melbourne University Press, 1961. Revised edition, forthcoming.

_____ 'The Methods of Ethics and The Methods of Science', Journal of Philosophy, vol. 62, 1965, 344_349.

Smart, R. N. 'Negative Utilitarianism', Mind, vol. 67, 1958, 542_543.

Sobel, J. H. 'Rule_Utilitarianism', Australasian Journal of Philosophy, vol. XLVI, 1968, 146_165.

_____ 'Utilitarianisms: Simple and General', Inquiry, vol. 13, 1970, 394_449.

Stace, W. T. The Concept of Morals. New York, Macmillan, 1937.

Stocker, M. 'Consequentialism and its Complexities', American Philosophical Quarterly, 1969, 276_289.

Sumner, L. W. 'Cooperation, Fairness and Utility', Journal of Value Inquiry, vol. V, 1971, 105_119.

Thomson, J. J. 'A Defence of Abortion', Philosophy and Public Affairs, vol. 1, 1971, 47-66.

Turner, J. W. C. (ed.) Kenny's Outlines of Criminal Law, 17th ed. Cambridge, Cambridge University Press, 1958.

Von Wright, G. H. Norm and Action. London, Routledge, 1963.

Wallace, G. 'Cargile on utilitarianism', Analysis, vol. 32, 1972, 126-131.

Warnock, G. J. The Object of Morality. London, Methuen, 1971.

Wasserstrom, R. A. The Judicial Decision. Stanford, Stanford University Press, 1969.

West, H. R. 'The Consequences of Act-Utilitarianism'. Unpublished.

_____ 'Act-Utilitarianism and Rule-Utilitarianism', Ph. D. Thesis, Harvard University, 1964. Unpublished.

Williams, G. The Sanctity of Life and The Criminal Law. London, Faber & Faber, 1958.

Windass, S. 'Double Think and Double Effect', Blackfriars, vol. 44, 1963, 257-266.