

CATEGORICAL IMPERATIVES

M. Phil Thesis

by

Noa Latham

Bedford College, London

November 1980

ProQuest Number: 10097357

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10097357

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

A distinction between two types of person reflecting two ways in which morality impinges upon life is sought, provisionally designated as a distinction between moralist and amoralist. The distinction is not easily characterised by way of action, emotion, motive, or verbal behaviour, and is not related to any egoistic/altruistic distinctions. The distinction is finally presented as that between believer and non-believer in the existence of intrinsic reasons for actions. The problem of applying the distinction is seen as a problem concerning the vagueness of the expression 'believer in intrinsic reasons'. Three ways in which vagueness could enter this expression are isolated. Of these it is argued that only self-deception does in fact provide a source of vagueness, though it will be minimal and for the purposes of the remainder of the enquiry may be removed by stipulation.

The presence of a belief in the existence of intrinsic reasons is manifest in the use of categorical imperatives. Thus the distinction may also be presented as that between user and non-user of categorical imperatives. An analysis of the term 'categorical imperative' is then undertaken and a definition based on intrinsic reasons is shown to be equivalent to one based on the conditions under which such judgements are withdrawn. An investigation of the variety of practical judgements that can occur as categorical imperatives follows, along with an investigation of the variety of reasons for action.

The question of the extent to which there are people who use only hypothetical imperatives in making their moral judgements, and the role of empirical evidence in settling this question, is then discussed. A sketch of the difference in moral life between user and non-user of categorical imperatives is offered. Finally, some ways of justifying the use (or non-use) of categorical imperatives are considered.

C O N T E N T S

<u>CHAPTER I. Introduction</u>	<u>Page No.</u>
1.1. Personal Motivation and Aim of Thesis. Search for distinction provisionally characterised as moralist/amoralist.	1
1.2. Vagueness. Gradational and Multi-dimensional.	3
1.3. Role of Action, Emotion and Verbal behaviour in the Moralist/Amoralist Distinction. Amoralist differs from moralist in action, emotion, and verbal behaviour, but distinction cannot be made in terms of any of these.	4
1.4. Role of Belief in the Moralist/Amoralist Distinction. Distinction is one of beliefs. Distinction non-vague if criterion for possession of belief based on assent. Doubts about such a criterion.	5
1.5. An Egoistic/Altruistic Distinction Based on Motives. Amoralist/Moralist not to be characterised in terms of egoism/altruism, where egoistic act is defined as one based on motive of self-interest.	7
1.6. An Egoistic/Altruistic Distinction Based on Survival Chances.	8
1.7. Kantian Moral/Amoral Distinctions. His division of motives based on his division of judgements into categorical and hypothetical imperatives. My moralist/amoralist distinction to be equated with distinction between user and non-user of categorical imperatives.	10

<u>CHAPTER II. The Definition of Categorical Imperatives.</u>	<u>Page No.</u>
2.1. Aim of Chapter.	12
Enquiry takes the form of an analysis of 'categorical imperative'.	
2.2. The Variety of Practical Judgements.	12
Variations of tense, person, generality, and strength, positive and negative.	
2.3. The Kantian Concept of Categorical Imperative.	14
Kant uses 'categorical imperative' both to mean an imperative unconditional upon the agent's desires, and to mean the Formula of Universal Law. He may also use the term to mean principle which satisfies the formula.	
2.4. A Grammatical Distinction.	17
The categorical/hypothetical imperative distinction is independent of the distinction between sentences without conditional clauses and those with conditional clauses.	
2.5. The Withdrawal Definition of Categorical Imperative.	18
Broad and Narrow senses of desire. Withdrawal definition of categorical imperative does not depend on sense of 'desire' chosen.	
2.6. The Reasons Definition of Categorical Imperative.	21
Relation between 'X ought to ϕ ' and 'There exists reason for X's ϕ -ing'. Definition of intrinsic reason. The definition most sensitive to the sense of 'desire' chosen. Reasons definition of Categorical Imperative equivalent to withdrawal definition, though the two have slightly different applications.	
2.7. Applications of the Definitions to Morality.	25

CHAPTER II.

Page No.

- 2.8. Applications of the Definitions to Prudence. 27
- 2.9. Vagueness and Self-Deception. 30
Ways in which vagueness may enter the moralist/amoralist distinction. Justification of choice of speaker's opinion as sole determinant of whether a judgement is a categorical or hypothetical imperative.
- 2.10. An Objection to the Definition of Intrinsic Reason. 33
All desires give reasons for action. A vagueness enters the term 'intrinsic reason' due to an uncertainty as to when a desirability characterisation may serve as an intrinsic reason.
- 2.11. Modification to the Definition of Intrinsic Reason. 35
Modifications are necessary to accommodate desires which may occur essentially in an intrinsic reason.
- 2.12. Comparison of Definitions of Modern Writers. 38
The definitions of Foot, Mackie, Williams, Nagel and McDowell.

CHAPTER III. The Variety of Intrinsic Reason and their Ontological Status.

- 3.1. Aim of Chapter. 42
- 3.2. The Aristotelian Syllogism 42
Ill-suited for distinguishing intrinsic and non-intrinsic reasons.
- 3.3. Explanatory and Guiding Reasons. 44
Intrinsic explanatory reason to be identified with a belief, guiding reason with a fact that by itself constitutes reason for action. Relations between 'X ought to ϕ ', 'There exists reason for X's ϕ -ing', and 'X has reason to ϕ '.

- 3.4. Are Intrinsic Reasons Linguistically Expressible? 48
Anyone believing a fact that constitutes a reason for action must also believe that it does constitute a reason for action. McDowell's modified account to make sense of non-linguistically expressible fact. Failure to arrive at belief that there is intrinsic reason to act not a manifestation or irrationality.
- 3.5. Overall and Conclusive Reasons. 53
'Overall reason' epistemic notion, 'conclusive reason' metaphysical.
- 3.6. Conflicts between Intrinsic Reasons. 55
Weaker obligation in moral conflicts not always annulled.
- 3.7. Conflicts Between Intrinsic and Non-Intrinsic Reasons. 56
Intrinsic reason need not prevail over non-intrinsic reason.
- 3.8. Silencing Versus Outweighing. 58
A need for outweighing prior to silencing.
- 3.9. Incontinence. 60
Temperate man contrasted with the continent or incontinent. Intention derives from categorical imperative judgement for the temperate man. Categorical imperative insufficient for determining intention of continent or incontinent man.
- 3.10. Variations of Generality. 65
Moral judgements need not be regarded as universalisable.
- 3.11. Etiquette. 67
Existence of categorical imperatives that may be considered neither moral nor prudential.

<u>CHAPTER IV.</u>	<u>Do all, some or no People use Categorical Imperatives?</u>	<u>Page No.</u>
4.1.	A simple Answer. All that is needed to argue for 'some' is to produce two examples.	72
4.2.	The Prudential Case.	72
4.3.	The Views of Some Modern Philosophers. Mackie, Williams, and Foot for 'some', though Foot is sometimes interpreted as arguing for 'none'.	74
4.4.	Kohlberg's Experimental Enquiry. The possibility of finding empirical evidence to support the claim that some members of a society use only categorical imperatives, or that there are societies in which only hypothetical imperatives are used. Kohlberg's enquiry designed for different purposes.	79
<u>CHAPTER V.</u>	<u>The Differences between Use and Non-Use of Categorical Imperatives.</u>	
5.1.	Plan of Chapter. Type of desire upon which hypothetical imperative moral judgement is conditional.	85
5.2.	First Person Moral Judgements. Amoralist cannot meet 'pretheoretical' description of incontinence. Supererogation often based on Hypothetical Imperatives.	88
5.3.	Second Person Moral Judgements. Advising, providing information, raising hackles.	96
5.4.	Third Person Moral Judgements. Not only needed to influence action.	98
5.5.	Prudential Judgements.	101

CHAPTER VI. Justification of the Use of
Categorical Imperatives.

Page No.

- 6.1. Types of Justification. 104
Kantian and contractarian arguments not discussed.
- 6.2. Theological Justification. 104
Certain religious beliefs may be regarded as grounding belief in intrinsic reasons. The Euthyphro dilemma.
- 6.3. Pragmatic Justifications. 106
Discussion of need for categorical imperatives in moral education and in maintaining a legal system.
- 6.4. Metaphysical Justifications. 111
Mackie's argument from queerness. Arguments from relativity and consensus. Objectivity of evaluative predicates. Relation between evaluative judgements, partial reasons, statements and conclusive reasons statements.
- 6.5. Conclusion. 117
Psychological difficulties in changing belief.

CHAPTER I. INTRODUCTION1.1. Personal Motivation and Aim of Thesis

I have frequently heard talk of the moral life and its components - moral decisions, moral approval, moral guilt, moral perception, moral sensitivity and even moral blackmail. And it has always seemed to me as though I did not participate in this moral life, though many of my friends and associates have disputed this. In writing this thesis I am seeking to make clear the distinction that I see between those involved in the moral life and the amoralists such as myself. If I fail to find this distinction, or if it constantly eludes me, I shall be forced to concede that there is not the basic difference that seemed to be of some importance. If I find that the distinction reflects a mere difference in linguistic use - that the amoralist's way of talking can be translated into the moralist's by means of a few simple translation rules - then I shall have to accept that there was an element of delusion that was sustained by this discrepancy in language use. If I find that the distinction reflects real non-verbal behavioural differences, then I shall discuss the differences between these two types of person and the ways in which their positions can be attacked or supported. In the course of this I should discover whether my position, which I take to be unorthodox, is tenable; whether there is something necessarily vile about amoralists such as myself, as a tutor of mine once declared; and hopefully change my behaviour so as to bring myself into line if I reach any negative conclusions.

What I am seeking, therefore, is a verbal characterisation of the moralist or of the amoralist that will enable anyone to classify people into one of these two types, and that will be suitable for the purposes of assessing the two positions. But in using the term 'moralist' and 'amoralist' haven't I already found my characterisation? The reason why these terms are unsatisfactory is that they are too vague.

People have varying ideas as to what an amoral person is. Is he someone who is uninterested in the welfare of others, or is he someone who actively disrupts the welfare of others, or is he someone who never does anything for others, or doesn't believe in doing things for others, or perhaps just someone with unconventional moral views?

I do not claim that the sense in which I shall use the terms 'moralist' and 'amoralist' will match that most prevalent in the linguistic community.

Before searching for a more precise characterisation of the distinction I shall digress briefly to discuss vagueness.

1.2. Vagueness

I shall call a proposition vague if and only if there need not be universal agreement as to its truth-value, because of the way the world is constructed.

I shall call a predicate 'P' vague if and only if, given complete information, there will be some objects, x, to which it is intelligible to apply the predicate, for which there is not universal agreement as to whether or not x is P, and some objects, x, for which 'x is P' has a determinate truth-value.

I shall distinguish two types of vagueness - gradational and multi-dimensional vagueness. Gradational, or one-dimensional, vagueness can be removed if necessary by harmless stipulation, and there will be associated with the predicate a non-vague relation term. Most vague predicates will also be relative, so that a specification of the class to which the predicate is to apply will be necessary if a stipulation is to be made. Thus the term 'tall' is vague as there will not be universal agreement as to where the line is to be drawn between the tall and the not tall. Relativised to a class, e.g. the class of adult men, the vagueness can be removed by stipulation, e.g. tall men are those over 1.80 m, and associated with the vague term 'tall' there will be a non-vague relation term 'is taller than'.

Multi-dimensional vague terms, like 'left-wing' and 'chair' have no corresponding non-vague relation terms, and any attempt at stipulation will distort the meaning of the term. For example, an attempt to stipulate 'left-wing' as 'believing in the nationalisation of all industry' would inevitably leave out other factors, upon which categorisation of political views depend.

Vagueness is not necessarily an undesirable feature of predicates. Terms like 'tall' and 'left-wing' do not stand in need of further refinement in most situations in which they are used. Vagueness is also a matter of degree. The vaguer a predicate 'P' is, the greater will be the class of objects x for which there is not complete agreement, given adequate information, as to whether x is P.

The distinction that I am searching for, that has been provisionally designated as that between the moral and the amoral person, must serve in a comparison of the two types of behaviour and an assessment of the two positions. To these ends it will be advantageous to find a characterisation of the distinction with as little multi-dimensional vagueness as possible, preferably none.

1.3. Role of Action, Emotion and Verbal Behaviour in the Moralistic/Amoralistic Distinction.

Examples will be of no use in conveying the distinction unless sufficiently many can be provided to enable people to 'catch-on' and use the distinction themselves. Examples drawn from well-known public or historical figures and portraits in literature will play a role in demonstrating that there are indeed both moral and amoral characters, and in influencing evaluation of these types of character. But a characterisation of the distinction will first be necessary in order to argue for the chosen classification of the figure selected in the example.

How then is this distinction to be presented? I expect the moralist to differ from the amoralist in his actions and emotion as well as in his verbal behaviour. But dispositions to a certain sort of action (such as truth-telling) or a certain sort of emotion (such as guilt) will be difficult to measure, and only indicative of aspects of the moral/amoral distinction, i.e. never capable of providing necessary and sufficient conditions for someone's being amoral. Differences of emotion and action I shall argue to be often effects of something more fundamental underlying the distinction, rather than as constitutive of it. I shall concentrate not on the actual differences in utterances made in moral situations but consider instead whether a person would assent or not assent to propositions applying to moral situations. Any difference in assent over a particular proposition must be attributed either to

a difference of beliefs, or to a difference in knowledge of the relevant facts. Given situations in which there is shared understanding of the concepts used in a proposition, and equal knowledge of the relevant facts, a difference in assent to the proposition must be explained by a difference in beliefs. I shall take the fundamental difference between moralist and amoralist to be a systematic difference in beliefs, which I will interpret as playing a causal role in action, emotion and verbal behaviour.

1.4. Role of Belief in Moralist/Amoralist Distinction.

If it is possible to exhibit all the differences of beliefs between moralist and amoralist in terms of one central belief, and this is what I meant by a systematic difference in beliefs, then it might appear that the distinction between moralist and amoralist could turn out to be non-vague.

If the central belief is embodied in a proposition, then assent to or dissent from this proposition will depend either on a difference in the central belief or on a different understanding of the concepts involved, since background information about particular situations will no longer be of relevance. Assuming that the proposition is couched in terms that are non-vague and hence understood in the same way by all members of the speech community, any difference in assent must be attributed to a difference in beliefs. Now if those lacking an opinion that decline either to assent or dissent are classed as non-assenters, then it might be thought that the distinction between believers and non-believers would correspond to a non-vague distinction between assenters and non-assenters. However, as long as there is some uncertainty of opinion, there will be a vagueness in deciding when the proposition seems sufficiently probable to warrant assenting to. Thus a vagueness may enter this way into the distinction between moralist and amoralist.

In distinguishing those that do hold a certain belief from those that do not, there will be difficulties in ensuring that different people understand the same thing by the proposition expressing the belief, and in deciding what degree of certainty is necessary to count as belief. A further difficulty, especially relevant in cases of metaphysical beliefs such as the belief in free will or theistic beliefs, concerns self-deception. It may be disputed that sincere assent to a proposition expressing the relevant belief is necessary and sufficient for possession of that belief.

Take, for example, the belief in the existence of God. Conceptions of God are likely to vary considerably within a speech community, so a notion of God would need to be spelled out in detail before it could be assumed that people were considering the same existential proposition. Assuming this stage could be reached, and assimilating the agnostics to the atheists, a problem would still remain as to whether this non-vague distinction between assenters and dissenters corresponded to the distinction of beliefs at issue. It might be denied that someone who sincerely claims to believe in a Christian conception of God could genuinely be accredited with such a belief if he never participates in standard features of Christian life, such as prayer. Alternatively, it might be argued that someone displaying certain signs of religious behaviour should be accredited with a belief in the existence of some sort of God even if such a belief is sincerely denied. Once factors in addition to sincere assent are included amongst criteria for the possession of a belief, there will be great difficulty in determining whether someone possesses a belief, and the distinction between possession and non-possession of the belief may, though not necessarily, be thought vague. Thus in the case of the moral/amoral distinction it will be necessary first to find a minimally vague proposition that expresses the relevant belief, secondly to consider the extent to

which uncertainty (about the belief) introduces a vagueness into the distinction between assent and non-assent to the proposition, and thirdly to discuss whether considerations of self-deception should lead to the introduction of criteria for the possession of a belief in addition to sincere assent.

1.5. An Egoistic/Altruistic Distinction Based on Motives

One way I thought it might be possible to characterise the moral/amoral distinction I was searching for was in terms of egoism and altruism. 'Egoist' is a dispositional term meaning someone who is prone to choosing the egoistic act when an altruistic act is also available. The non-egoist will occasionally abjure the altruistic act in favour of an egoistic one, without it making him into an egoist. An egoistic person is thus defined in terms of an egoistic act and this is usually taken to mean an act motivated by self-interest. On this view egoistic acts will be difficult to recognise for three reasons. First, because motives are themselves difficult to pick out - people are frequently mistaken about their motives. Secondly, because there may be problems concerning the classification of motives as self-interested, or non-self-interested. Thirdly, because acts are invariably performed from mixed motives, so that an act may often be partly egoistic rather than wholly egoistic or altruistic. In addition to the variety of ways in which vagueness enters the concept of egoistic act, there is a gradational vagueness added to this in the concept of egoistic person owing to the dispositional nature of the concept.

As well as involving considerable vagueness, the distinction between egoistic act and altruistic act differs appreciably from what would commonly be regarded as the distinction between an amoral act and a moral act, and

hence also the distinction I am seeking. As a duty is generally recognised towards the self it is often considered right to forego some slight altruistic act when the cost to the self would be great. So a moral act is not necessarily an altruistic act. Nor is an amoral act necessarily an egoistic act. Cases of spontaneous benevolence such as maternal care for children I would want to class as amoral, but in terms of the egoistic/altruistic distinction, an altruistic motive-component would have to be separated from such egoistic motive components as hope of future reward from the beneficiary or other observers, fear of reprimand or lack of future assistance from others, and (contentiously) a feeling of self-righteousness. Thus an amoral act may be considered partly altruistic. It is even possible to imagine acts that would be considered immoral and altruistic, as when a mother rescues her child from the path of an erupting volcano rather than alert a whole village.

1.6. An Egoistic/Altruistic Distinction Based on Survival Chances.

It may be instructive at this stage to look at an alternative definition of the egoistic/altruistic distinction found in biology. An altruistic act is one that has a tendency to increase the survival chances of another organism or organisms whilst lessening those of the acting organism. This definition dispenses with the notion of motive which is an awkward hybrid between a cognitive and a behavioural concept. In so doing, it clearly gives a new sense to the terms 'egoistic' and 'altruistic' and should not be thought of as providing an analysis of those terms in their common use. One advantage this definition of altruism has, is that it is not vague. There is no gradational vagueness about the concept of altruistic act as the distinction between increasing and decreasing probability is sharp, i.e. there is no question of where to draw the line, as there is, for example, in the case of tall and not tall people. People

will not diverge in their classification of acts into egoistic and altruistic given adequate evidence. It will in principle be possible to verify whether or not an act is altruistic by observation of acts of the type over a long period of time. In practice, however, verification will seldom be feasible, so that there would be divergence in classifications of acts into egoistic and altruistic owing to lack of sufficient evidence.

Principle interest in the biological definition could be focussed on altruistic act, but a definition of altruistic organism could easily be given, for example as one with non-zero disposition to altruistic action. Humans will vary in their disposition to altruistic acts, but with non-human animals this disposition will be more or less constant within a species. This biological distinction between altruistic and non-altruistic acts will not be of much relevance with regard to individual human acts because moral and other human action will seldom have any appreciable effect on survival chances.

However, arguments from men's ancestry have been raised aimed at showing that members of the human species are genetically endowed so that given a suitable environment they would develop a disposition to altruistic behaviour directed beyond their immediate offspring. Such arguments, if valid, would suggest that man may sometimes automatically act altruistically (still in the biological sense), not just because of his training, but from an instinct as basic as sexual desire. This may furthermore suggest the artificiality of a distinction between egoistic and altruistic acts (reverting now to the usual sense of the terms), and suggest the extreme implausibility of psychological egoism - the view that all human action is motivated by self-interest. I say merely 'implausibility' rather than 'impossibility' because unconscious, genetically based urges need not produce conscious motives, not because owing to the need for a suitable environment, genetic endowment might never lead to such behavioural dispositions.

1.7. Kantian Moral/Amoral Distinctions

Having noted the general difficulties in presenting the moral/amoral distinction in terms of a distinction between motives, and the particular inaptness of a distinction between self-interested and non-self-interested motives, I shall consider briefly the Kantian distinction between moral and amoral motives. Kant writes of actions motivated from duty, purposes of self-interest, and inclination, though he does not treat these as mutually exclusive. Duty and self-interest are distinct, but in dividing inclinations into immediate and non-immediate, he appears to indicate that non-immediate inclinations must be either for purposes of self-interest or duty. Immediate inclinations include natural benevolence or sympathy, but it is not obvious whether Kant would have allowed such inclinations to include natural malevolence, purposes of self-interest, or duty. The question of overlap between the motive of inclination and other motives is distinct from the question of the possibility of copresence of the three motives. Kant believed that an act could be motivated by a mixture of duty, self-interest and inclination.

Though this Kantian distinction between the moral motive of duty and the amoral motives of inclination and self-interest will accord with the moral/amoral distinction I am seeking, it will fortunately be unnecessary to seek any further clarification of it as I can proceed directly to an alternative moral/amoral distinction offered by Kant, namely his distinction between categorical and hypothetical imperatives. This distinction between moral and amoral judgements is both wider-ranging and more fundamental than that between moral and amoral motives. It is wider-ranging because there will be judgements which needn't issue in actions and it is more fundamental because actions motivated by duty must be decided upon in accordance with a maxim which could be willed as universal law, i.e. actions motivated by duty must accord with a categorical imperative.

It is this distinction between moral and amoral judgements upon which I shall, in fact, base the moral/amoral distinction that I am seeking. The amoralist will be someone who does not make categorical imperative judgements. The making of categorical imperative judgements will be related to the possession of a central belief in Chapter II where I shall also discuss the sense in which I shall use the terms 'categorical imperative' and 'hypothetical imperative' and its departure from Kant. The distinction between categorical and hypothetical imperatives and the variety of forms in which these judgements appear will receive discussion in chapters II and III. In Chapter IV I shall discuss the extent to which categorical imperatives are used in moral judgements, and what sort of people dispense with them altogether. In Chapter VI I shall look at some of the qualitative differences in the moral life between those who do make use of categorical imperatives and those that don't. Finally, in Chapter VI I shall consider some ways in which the use of categorical imperatives may be justified or challenged.

3.1 The Variety of Practical Judgements

I shall begin by looking at the term 'practical judgement'. These are judgements which are principally used to direct or guide action. The term is best illustrated by presenting the variety of grammatical forms such judgements take. Many practical judgements take the form 'X ought to...' where 'X' stands for a person or group of persons and

CHAPTER II. THE DEFINITION OF CATEGORICAL IMPERATIVE

2.1. Aim of Chapter

The aim of this chapter is to provide a distinction between two types of practical judgements that I shall refer to as categorical and hypothetical imperatives. Though these are Kantian terms, I shall not be using them in the Kantian sense, but in a sense in which they have come to be used in more recent philosophical literature. The differences between the modern sense and the original Kantian sense will be sketched.

An analysis of the term 'categorical imperative' will then be undertaken so as to discuss, in Chapter IV, whether all, some, or no members of the speech community make use of such judgements; in Chapter V, any differences in behaviour that might be expected from a divergence in such language use; finally in Chapter VI, the justification of the use of categorical imperatives. To these ends, I could dispense with an analysis of the term 'categorical imperative' and simply stipulate the sense in which I intend to use it. Nevertheless, the enquiry does seem to take the form of an analysis. Although the term 'categorical imperative' is not part of the vernacular and there is no other with a similar meaning, the distinction between the two types of practical judgement does seem to be one with which all language users are conversant. No importance, however, is attached to the claim that there is a consistent modern use of the term and that I have captured it in my definition.

2.2. The Variety of Practical Judgements

I shall begin by looking at the term 'practical judgement'. These are judgements which are principally used to direct or guide action. The term is best illustrated by presenting the variety of grammatical forms such judgements take. Many practical judgements take the form 'X ought to ϕ ' where 'X' stands for a person or group of persons and

' ϕ ' is a description of an action or type of action. Other typical verbs used in such judgements are 'should', 'must', and 'have to', and there are a variety of similar forms such as 'X has an obligation to ϕ ', 'It is X's duty to ϕ ', 'It would be best for X to ϕ ', 'It would be good for X to ϕ ', 'It would be right for X to ϕ '. More contentiously action-guiding are 'It would be honourable for X to ϕ ', 'It would be courageous for X to ϕ ', 'It would be detestable for X to ϕ '.

As well as occurring in positive or negative form, in present or past tense, and in first, second or third person, there are also variations in generality and strength. Variations in generality can best be demonstrated by means of an example.

Take the following singular judgement:

- (1) NL must answer this question truthfully now.

This may be generalised in the following ways:

- (2) Everyone must answer this question truthfully now.
 (3) NL must always tell the truth.
 (4) Everyone must always tell the truth.

By variations in strength I mean variations in the extent to which the different verbs are action-guiding. The examples given cannot consistently be ranked in order of strength, but three distinct grades, or logical types, can be discerned. Initially there is the case of pure evaluation and zero action-guiding force. Whether it is a possibility in cases of moral evaluation of actions such as 'It would be honest of X to ϕ ' is a question to which I shall return in 2.6. and 6.4. Next there is the case in which some action-guiding force is present, but not so as to exclude the possibility of further action-guiding force attaching to incompatible actions. And finally there is the case in which action-guiding force cannot attach to incompatible actions. There is some matching of these

grades in ordinary language. 'Must', 'have to', and 'it would be best for' are invariably of the third type, so that 'X must ϕ ' and 'X must not ϕ ' are incompatible. 'Should' and 'ought to' are often of the second type, in which case 'X should ϕ ' is compatible with 'X should not ϕ '. Most verbs may appear in judgements of both second and third types.

It should be noted that not all judgements of the form 'X ought to ϕ ' are practical judgements. There are also epistemic uses of 'ought' which have nothing to do with action-guiding and indicate merely expectation such as 'He must arrive soon', 'He ought to arrive around midday'.

It might also be thought that ought judgements could be made concerning not actions but interests or desires, such as 'You ought to be interested in politics', 'You shouldn't be addicted to heroin'. If such judgements are to make sense other than as purely evaluative, they must be construed as shorthand for 'You ought to take steps (or ought to have taken steps) to cultivate (or eradicate) that interest'.

I shall now discuss the division of practical judgements into categorical and hypothetical imperatives, beginning with a look at Kant's use of the terms.

2.3. The Kantian Concept of Categorical Imperative.

Among the various descriptions Kant provides of these terms, unfortunately none is singled out as a definition. The following passage might be taken as most indicative:

'Hypothetical imperatives declare a possible action to be practically necessary as a means to the attainment of something else that one wills (or that one may will). A categorical imperative would be one which represented an action as objectively necessary in itself apart from its relation to a further end.

Here a categorical imperative is distinguished from a hypothetical imperative in being unconditional on the presence of some desire or purpose of the agent's. But Kant also writes:

Finally, there is an imperative which, without being based on, and conditioned by, any further purpose to be attained by a certain line of conduct, enjoins the conduct immediately. This imperative is categorical. It is concerned, not with the matter of the action and its presumed results, but with its form and with the principle from which it follows; and what is essentially good in the action consists in the mental disposition let the consequences be what they may. This imperative may be called the imperative of morality.

op. cit. p.80

From this quote it appears that a categorical imperative is not only unconditional but also subject to a formal requirement. This formal requirement is to be arrived at a priori and is embodied in the Formula of Universal Law, which Kant introduces as follows:

There is therefore only a single categorical imperative and it is this: 'Act only on that maxim through which you can at the same time will that it should become a universal law.'

op. cit. p.84

From this last passage it appears that Kant is using the term 'categorical imperative' only for the Formula of Universal Law.

How then would Kant classify the variety of practical judgements that have been exhibited that are intended to hold unconditionally upon the existence of any desire or purpose of the prospective agent's? In discussing the possibility of the existence of a categorical imperative, Kant argues

that it is never possible to demonstrate this by example. He considers the case 'Thou shalt make no false promises' and argues that it can never be certain that action in accordance with such an imperative is motivated solely by the practical law of morality. From this it may be deduced that the principle 'Thou shalt make no false promises', is not to be regarded with any certainty as a categorical imperative, as it may have served merely as a pragmatic prescription. This does not, however, entail that the principle is a hypothetical imperative. Once having presented his Formula of Universal Law as an a priori categorical imperative, it is open to question whether Kant would have accepted the principles which satisfy this formula as themselves categorical imperatives.

Certainly there is a sense in which Kant uses 'categorical imperative' to mean the Formula of Universal Law. Exploring the possibility that there is a sense in which he uses the term to apply to certain practical judgements, the obvious candidates would be subjective principles such as 'I ought never to tell a lie', or universalised subjective principles^{1.} such as 'Nobody ought ever to tell a lie' that satisfy the formula. Such principles might be regarded as categorical imperatives though there will be practical problems in applying the formula, namely in knowing how the subjective principle is to be universalised, and in knowing whether the universalised subjective principle could be willed as a universal law.

Next it would seem reasonable to regard as categorical imperatives those singular judgements such as 'I ought not to tell a lie now' that can be derived from subjective principles that satisfy the Formula of Universal Law. But unless the judgement has actually been derived from such a principle, a considerable practical difficulty will exist in isolating a unique subjective principle from which the singular judgement could be derived.

1. Kant reserved the term 'objective principle' for those valid for every rational being, i.e. those that satisfy the Formula of Universal Law.

So if there is a Kantian sense of 'categorical imperative' that applies to certain practical judgements, it will be difficult in practice to distinguish the categorical imperatives from the hypothetical imperatives. It would also appear, from the second of the passages quoted from Kant, that this concept of categorical imperative is serving as an answer, both to the question 'What is a moral judgement?' and to the question 'What moral judgements are correct?' This would leave Kant no way of classifying the sincere judgements intended as unconditional that do not meet the requirements of the Formula of Universal Law. A consequentialist, for example, may in some circumstances judge that a person ought to tell a lie, or that a person would be permitted to commit suicide. For Kant, however, such judgements could not meet the requirements of the Formula of Universal Law, and so could not constitute categorical imperatives.

If the term 'categorical imperative' is to apply to practical judgement it is perhaps best to take the unconditional requirement as distinguishing the categorical imperatives from the hypothetical imperatives that are true from those that are false. It is certainly the unconditional feature of practical judgements that is central to the concept of categorical imperative as it is used in modern philosophical writing, and it is this sense of categorical imperative that I intend to examine.

2.4. A Grammatical Distinction

First, though, it will be of interest to look at the thought that categorical imperatives may be distinguished from hypothetical imperatives by the grammatical form in which they find expression, or at least that there is a connection between the distinction and some grammatical distinction. The obvious grammatical distinction of relevance here is that between sentences of the form

'X must (or should etc.) ϕ ' and those of the form 'If C, X must ϕ '. It might be thought that all sentences of the first form express categorical imperatives whilst all those of the second form express hypothetical imperatives. Yet this is easily shown not to be the case.

Many imperatives without conditional clauses are elliptical. When it said, for example, 'You must run', thinking someone wants to catch a train, this is clearly elliptical for 'If you want to catch the train you must run', which is immediately recognisable as hypothetical. What it is that enables people to determine whether or not an imperative is elliptical is not itself a grammatical feature.

It can also be seen that not all imperatives with conditional clauses are hypothetical. For example, the imperative 'If you have borrowed money you must repay it' is normally thought of as categorical, equivalent to 'You must repay your debits.'

Hence the grammatical distinction is of no use in determining whether an imperative is categorical or hypothetical, though categorical imperatives will generally be expressed without a conditional clause.

2.5. The Withdrawal Definition of Categorical Imperative.

I shall now introduce a first formulation of the categorical/hypothetical imperative distinction as it is currently understood. I shall take as a representative judgement 'X ought to ϕ ', where ' ϕ ' is a description of a possible act, and 'X' is a name of a person, or pronoun. The definition will apply to singular judgements, classed as 2.2. as of the second kind, in first, second or third person, present or past tense.

Y's judgement that X ought to ϕ is a categorical imperative if and only if, Y would not withdraw his judgement if ϕ -ing were not conducive to the satisfaction of X's desires, and Y knew this.

Y's judgement that X ought to ϕ is a hypothetical imperative if and only if, it is not a categorical imperative.

The classification of judgements into categorical and hypothetical imperatives thus depends solely on what the person making the judgement thinks about his judgement. There is no overruling that person's own classification according to the above definition. The classification is also independent of the prospective agent's (ie. X's) belief as to whether or not the action would be conducive to the satisfying of his desires.

A vagueness may be thought to infect the distinction owing to the vagueness of the notion of desire. I shall be using the term 'desire', unless otherwise specified, to mean 'present desire'. This will not apply just to appetitive desires with a particular phenomenological character such as hunger, but also to long-term projects and interests which need not be occurrent at the time one has them, such as a desire to keep one's weight down, or a desire to deepen one's command of a foreign language. The notion thus far sketched I shall call the narrow sense of desire. This will be extended to a broad sense of desire which covers desires that are attributable solely on the basis of a person's formation of an intention to act. This extra type of desire differs slightly from that which Nagel and McDowell describe as trivially attributable from the agent's being motivated to act. I have chosen to exclude cases of desires attributable on the basis of actions that are not intentional under any description i.e. involuntary actions, and include cases where the agent formed an intention to act but failed to act in accordance with that intention, either due to change of circumstances or to incontinence.

As categorical imperative has been defined in terms of independence of present desires, this opens up the possibility that prudential judgements concerning a person's future welfare or happiness may sometimes be categorical imperatives. Kant believed that all prudential judgements were hypothetical imperatives as he believed that such judgements recommend certain actions as means to an end of personal happiness - which he took all men to have as a matter of natural necessity. Whether attribution of the end of personal happiness is the most natural way of explaining prudential judgements, or whether such an explanation is vacuous will be discussed when I come to consider the problems of distinguishing categorical imperatives from hypothetical imperatives in the domain of prudence. The role of the categorical imperative in prudence will receive attention primarily in so far as it throws light on the case of morality.

As the term 'desire' occurs essentially in the consequent of a conditional in the definition of categorical imperative, it might be expected that the definition would yield differing classifications of judgements into hypothetical and categorical imperatives according as the broad or narrow sense of 'desire' is chosen. This in fact will not be the case when the consequent is contrary to fact, rendering the conditional subjunctive. When the conditional is subjunctive, i.e. when Y does not know that ϕ -ing would not be conducive to the satisfying of X's desires, the definition can be read using the broad sense of desire, as this will be the easiest to apply.

If Y already knows that ϕ -ing is not conducive to the satisfaction of X's desires, then the conditional is material and the conditions under which the antecedent is true will be sensitive to the sense in which 'desire' is

understood. If Y knows that ϕ -ing is not conducive to the satisfaction of X's desires in the broad sense, then this will hold a fortiori also for the narrow sense. Imagine now a situation in which ϕ -ing is not conducive to the satisfying of X's desires, when 'desire' is taken in the narrow sense but not in the broad sense. This would have to be a case in which X intends to ϕ but ϕ -ing will further none of his long-term projects nor lead to the satisfying of any of his appetitive desires. It will be difficult to determine with any certainty when such situations obtain, and hence there will be few cases in which Y can be said to know such a fact. But even in these cases, it is inconceivable that Y's knowledge of such a fact would bring him to withdraw his judgement that X ought to ϕ . Hence there are no cases in which the application of the definition of categorical imperative depends on whether the narrow or broad sense of 'desire' is chosen.

The definition of categorical imperative given above in terms of withdrawal is well-suited to cases in which overall judgements taking all relevant information at hand into consideration are on offer. In such cases it is X's ^{overall} desire that is to feature in the definition. But it will be difficult to apply the definition in cases in which conflicting considerations validate contrary judgements i.e. when 'X ought to ϕ ' and 'X ought not to ϕ ' are both considered to be valid. In such cases it is any of X's desires that is to be understood in the definition. I shall therefore introduce an alternative definition that is more directly applicable and that will reveal the structure of an overall judgement.

2.6. The Reasons Definition of Categorical Imperative

To this end the putative analysis of 'X ought to ϕ ' as 'There exists reason for X's ϕ -ing' will be considered. The proposed analysis seems insufficient in view of second and third person judgements concerning supererogation. Here there is a typical moral reason for acting, yet the act would be considered beyond the bounds of the duty, and

so it would not be judged that the man ought to do the act in question. The reasons analysis may also be adjudged insufficient because of cases in which the reason statement is weaker than the ought statement. There are cases of this in which the reason is so weak as to be barely worth mentioning, such as a faint itch which provides a reason for scratching my ear. In this case there is no reason not to scratch, as the movement of my hand will not detract from anything else I have reason to do, yet it would be inappropriate to say that I ought to scratch. There are also cases in which I have a reason to do something which would conflict with an alternative course of action for which the reasons may be far better. And in these cases where the proposed action is easily overruled, 'ought' too may seem inappropriate. One might try to describe these examples in such a way as to be able to treat them as exceptions, in order to formulate a vague sufficiency statement such as the following: 'there exists reason for X's ϕ -ing' entails 'X ought to ϕ ' provided that the reason is neither very weak, very strong, nor easily overridden. Untampered with, however, it must be conceded that 'there exists reason for X's ϕ -ing' is insufficient for 'X ought to ϕ '.¹ That the reasons statement is necessary for the ought statement, on the other hand, seems unchallengeable. A notion of 'ought' in which 'You just ought' could be a satisfactory answer to 'Why?' would be useless and difficult to understand.

Assuming then that 'X ought to ϕ ' entails 'There exists reason for X's ϕ -ing', a distinction between two types of reason will produce a distinction between two types of practical judgement. I shall introduce a distinction between intrinsic and non-intrinsic reasons by means of the following definition:

An intrinsic reason for X's ϕ -ing is one which is independent of X's desires.

1. 'There exists overall reason for X's ϕ -ing' is sufficient for 'X ought to ϕ '. For a definition of overall reason, see 3.5.

or equivalently,

An intrinsic reason for X's ϕ -ing is one whose specification contains no essential reference to X's desires.

The definition of intrinsic reason may also be shown to be insensitive to the sense of 'desire' chosen. A reason involving reference to desires in the narrow sense will obviously be a reason involving reference to desires in the broad sense. The reverse is also true since a reason, R_1 , involving reference to desires in the broad sense but not in the narrow sense would have to be a reason based on a desire that is attributable solely on the basis of an intention to act. But this intention cannot have arisen from a desire in the narrow sense, or such a desire would occur essentially in the reason which ex hypothesi involves no reference to desires in the narrow sense. Hence the intention must have arisen from a reason, R_2 , that involves no reference to desires in the narrow sense. But now there seems to be nothing gained from the postulation of R_1 , as there is no means available whereby it can differ appreciably from R_2 , and no question to which it can provide an answer that would not also be answered by R_2 .¹

The definition of categorical imperative now appears as follows:

Y's judgement that X ought to ϕ is a categorical imperative if and only if, Y would judge that there existed intrinsic reason for X's ϕ -ing.

On this definition, the question of whether a judgement is a categorical imperative is fixed by the opinion of the person making the judgement. Though the question of whether something constitutes an intrinsic reason or not is independent of the opinion of the person making the judgement, it will be argued (in 2.9) that the question of whether Y believes that there exists intrinsic reason for X's action

1. The independence of intrinsic reasons from desires is not related to the question of temporal priority. A desire may be attributed to someone on the basis of his formation of an intention, after coming to believe there exists intrinsic reason for him to act. But a person could also have a desire to ϕ , then come to see that there is also intrinsic reason for him to ϕ .

cannot be answered in any significant way independently of Y's view of the matter. This accounts for the appearance of 'Y would judge that...' in the definition instead of 'Y believes that....'

The definitions can easily be seen to be equivalent. If Y would judge that there is reason for X's ϕ -ing independent of X's desires, then he would not withdraw his judgement if he came to know that ϕ -ing was not conducive to the satisfaction of X's desires. Conversely, if Y judges that X ought to ϕ and would not withdraw the judgement if he came to know that ϕ -ing was not conducive to the satisfaction of X's desires, then his belief that there is reason for X's ϕ -ing would have to be a belief that there is intrinsic reason for X's ϕ -ing.

Before advancing it should be noted that some practical judgements are thought to contain elements of both hypothetical and categorical imperatives. On the above definition, a person is making a categorical imperative judgement when he would judge that there existed intrinsic reason for action, though he might also judge that there existed non-intrinsic reason as well. Thus a judgement designated categorical is one with some categorical element, but not necessarily purely categorical. The definition of a hypothetical imperative as a judgement that is made when a person would also judge that there existed non-intrinsic reason for action, could be regarded as a definition of a pure hypothetical imperative. A pure categorical imperative could be defined as a judgement that is made when a person would also judge that there existed intrinsic reason but no non-intrinsic reason for action.

Though it has been shown that the concept of categorical imperative does not pick up a vagueness due to the vagueness of the concept of desire, it does nevertheless suffer from a vagueness due to the difficulties facing Y, the person making

judgement, in deciding on his secondary judgement required in applying the definition.

Though equivalent, the two definitions call for slightly different secondary judgements. The withdrawal definition asks Y to imagine how his original judgement would be affected under certain conditions, whereas the intrinsic reason definition asks Y to consider whether he would judge that there existed reason for X's ϕ -ing independent of X's desires. It is likely that whenever Y has a strong intuition about the question of dependence upon desires, he will immediately know whether he would withdraw his definition under the proposed conditions. When intentions are not clear, the withdrawal definition may provide further chance of a secondary judgement by proposing a kind of thought-experiment in which Y tries to imagine the counter-factual situation obtaining. I shall examine the problems of applying the definitions by considering some examples, first moral, and then prudential.

2.7. Applications of the Definitions to Morality

If Y judges that X ought to ϕ and Y doesn't believe that ϕ -ing would be conducive to the satisfaction of X's desires, then it is obvious that Y cannot believe that the reason for X's ϕ -ing depends on X's desires, so that the judgement 'X ought to ϕ ' must be a categorical imperative. The clearest cases of this type are those in which Y knows that ϕ -ing would not be conducive to the satisfaction of X's desires. For example, I may judge that I ought to have visited my grandparents last summer, whilst knowing that at no time during that summer did I have any sort of desire, intention or pro-attitude towards visiting them. The clearest cases are often past tense first person judgements as in the example just given, though cases can arise in any person or tense. It might be disputed that first person

present tense judgements 'I ought to ϕ ' can be made by someone with no desire to ϕ and who forms no intention to ϕ . I will argue in 3.9 that there are cases of this type, though for the present purposes this is unimportant. Cases in which Y is not sure whether ϕ -ing would be conducive to the satisfaction of X's desires are principally second and third person judgements. For example, Y may judge that X ought to pay his grandparents a visit without knowing anything about X's attitude towards visiting his grandparents.

In the cases in which Y believes that ϕ -ing would be conducive to the satisfying of X's desires, in the broad sense (perhaps X has already ϕ 'd) both hypothetical and categorical imperatives are possible. It will be clear that the reason for X's ϕ -ing is independent of X's desires if the relevant desire arises after X has seen that there is reason for his ϕ -ing. But if the relevant desire arises before X has seen that there is reason for his ϕ -ing, the reason would be either dependent or independent of X's desires. There will be few cases in which temporal priority will help Y in classifying his judgement as a hypothetical or categorical imperative. In general, Y will know whether he takes the reason to be independent of X's desires. The most difficult cases arise when Y knows that ϕ -ing would be conducive to the satisfying of X's desires in the narrow sense. For example, I may judge that I ought to visit my grandparents, but because I enjoy visiting them, I don't know whether I'd judge that there existed intrinsic reason for visiting them. The withdrawal definition suggests that I imagine the counter-factual situation in which I have no desire in the narrow sense to visit, and try to decide whether I would withdraw my judgement. However, it may be that the difficulty I have in imagining myself without the relevant desire leads me to imagine a situation in which another person is substituted for myself. But this need

not lead to the correct verdict unless certain universalisability constraints are accepted as attaching to categorical imperatives.

2.8. Applications of the Definitions to Prudence

The task of distinguishing categorical from hypothetical imperatives is more difficult in the case of prudence. This is because the task of deciding whether the reason for action is independent of desires, which was the main problem in the moral case, is compounded by the problem of deciding when a desire exists (I shall now revert to writing ^{present} desire). Suppose, for example, that my supply of coffee is on the point of running out and I judge that I ought to buy another jar. Assuming the entailment from 'ought' to 'reasons' holds good for prudential as well as moral reasons, it may be deduced that I believe there is reason for my buying a jar of coffee. Now the question arises as to whether the reason stems from my future desire to drink coffee that I am now quite certain will arise tomorrow at about 2.00 p.m., or whether the reason stems from my present desire that that desire due around 2.00 p.m. tomorrow should be satisfied along with a set of similar desires placed at equal regular intervals in the future, or whether the reason stems from my present desire simply to have a supply of coffee in the kitchen, or whether the reason stems from my desire for happiness. In the first case the reason for action would be intrinsic and the judgement therefore a categorical imperative; in the next cases the reason would not be intrinsic, and the judgement accordingly a hypothetical imperative.

It is difficult to decide between these options because there seem to be no firm criteria from which to judge, and the choice seems rather unimportant except for the light it could possibly throw on the corresponding choice in the case of morality. The fourth of the above options should be rejected as conceptually unsound unless it can be shown, ^{in non-moral cases} λ when something counts as evidence that a person is not acting from a desire for happiness. The argument for intrinsic reasons often takes the form of an argument criticising the unnaturalness and implausibility of a desire for the

satisfaction of one's future desires. The third option of a present desire to possess a supply of coffee in the above example, often appears the most apt of the alternatives. Perhaps in the above example it would appear less natural to regard having a supply of coffee as one of my long-term interests if my interest is only awakened when a particular appetitive desire occurs. On the other hand if I take great delight in the appearance of my kitchen with the coffee jar along with everything in their correct places on the shelves, then it might appear less natural to regard my reason for buying some coffee as a desire for the satisfaction of future desires, or as those future desires themselves. The relevant desire in this case is the present one that something be done about the empty jar of coffee. To take another example, the reason for my buying a piano did not seem at the time to be a desire to satisfy future desires to play whenever they should arise. Nor did it seem as though those future desires alone gave me reason then for buying the piano. Rather it seemed as though the buying of a piano would be instrumental in serving the desire I then had to become as good a pianist as possible. The future desires to play would not have existed with such frequency had not the means been available for their satisfaction. This suggests another reason why future desires, directly or indirectly need not feature in a reason for action. Cases may exist in which no future desires D_i would be expected if I do not now perform act A. So the fact that I would have desires D_i which would be satisfied in the future if I now do A is insufficient for saying that those desires are my present reason for doing A.

Considering now the task of distinguishing categorical from hypothetical imperatives, the cases in which Y doesn't believe that ϕ -ing would be conducive to the satisfaction of X's present desires, when only the categorical imperative judgement is possible, are less easily recognised. The clearest cases would be those in which Y knows that ϕ -ing would not be conducive to the satisfaction of X's desires, and again these are most readily apparent in first person past tense judgements. For example, I may now judge that

I ought to have changed subjects at Cambridge ten years ago, knowing that there is no way that I could come to that conclusion based on my desires at that time, unless my desire for happiness is to be accepted. Again it is probable that cases occur in any person and tense in which Y would claim that his judgement that X ought to ϕ did not depend on X's desires. As an example of the case in which Y believes that ϕ -ing would be conducive to the satisfaction of X's present desires, and hence both categorical and hypothetical imperative are possible, take my judgement that I ought to have a game of squash this week. There is clearly a desire to play based on the enjoyment that the game will bring me, plus a desire to get a fair amount of exercise. Yet if both these desires were absent, I might still make the judgement, based on an idea that I needed exercise. The problem of deciding whether my judgement is a categorical imperative, i.e. that it could be withdrawn if playing squash were not conducive to the satisfaction of any of my present desires, is compounded by the prior problem of deciding what are to count as present desires.

The definition of categorical imperative has been shown not to be sensitive to the sense of 'desire' chosen. Nevertheless, the practical task of distinguishing the two types of judgement is affected by the problem of recognising when a desire in the broad sense is present, and this is more difficult in the case of prudence than that of morality. The exhibiting of clear cases in which a judgement falls into the category of hypothetical or categorical imperative will be used in the following section, as well as facilitating the enquiry of Chapter IV as to the extent of the use of categorical imperatives in a speech community, and those of Chapter V and VI.

2.9. Vagueness and Self-Deception

At this stage it is worth relating the conclusions reached so far to the problem mentioned in the introduction. The central belief that distinguishes the moralist from amoralist is the belief in the existence of intrinsic reasons. The moralist/amoralist distinction may therefore be presented in terms of the distinction between assenters and non-assenters to the proposition 'there exist intrinsic reasons'. There are three stages at which vagueness may appear in the moralist/amoralist distinction. First, the proposition itself might be vague. I have, however, argued that this proposition does not pick up a vagueness from the vagueness of 'desire'. Secondly, there may be uncertainty as to the truth of the proposition, and hence vagueness in knowing what degree of certainty is to count as assent. However, having presented a set of examples of categorical imperatives that are clearly distinguishable from hypothetical imperatives, it ought to be obvious to anyone considering these examples, whether he would make such judgements himself. Hence there should be no difficulty in deciding whether or not to assent to the proposition 'there exist intrinsic reasons', and therefore no vagueness in the distinction between assenters and non-assenters. Thirdly, the distinction between assenters and non-assenters may be thought not to coincide with the distinction between believers and non-believers in intrinsic reasons owing to some form of self-deception.

In choosing Y's own opinion as to whether he believes that there exists that there exists intrinsic reason for X's ϕ -ing as determinant of whether his judgement is a categorical or hypothetical imperative, I claimed (in 2.6) that there is no other significant sense in which it can be said that Y believes there exists intrinsic reason. It is now time to defend that claim, and I shall concentrate primarily on the moral case. The claim can be challenged,

either by denying that a person's sincerely asserting that he believes that p is sufficient for attributing to him a belief that p , or by arguing that a contrary unconscious belief that not p can more or less annul the effect of a conscious belief that p . The former option is unattractive as it would involve challenging the intelligibility of the words Y uses to express what he takes to be a belief that p , either by imputing to him some other belief, or by denying that he is expressing a belief at all. It may be that this latter approach would be adopted by non-cognitivists such as Hare. (He would claim that a person who takes the assertion X ought to ϕ as equivalent to the expression of a belief '∃ reason for X 's ϕ -ing' misunderstands the use of moral language terms.) It would challenge not the occasional assertion of a belief in the existence or inexistence of an intrinsic reason, but the view that anyone ever believes in the existence or inexistence of an intrinsic reason, by denying the intelligibility of the notion of intrinsic reason.

The second and more plausible way of challenging the view that Y 's own opinion is essential to an ascription to him of a belief in the existence of an intrinsic reason, is to argue that a contrary unconscious belief can coexist with the conscious belief. Only cases of unconscious beliefs conflicting with conscious beliefs need be considered since it is to be supposed, first, that the conscious belief is to be determined by what Y would sincerely judge if asked, (So there can be no cases in which Y has no belief merely because he has not thought about the situation); and secondly, that Y cannot withhold judgement unless he has insufficient information about X 's situation (But maybe he unconsciously knows all about X !)

If Y 's conscious belief is that there exists intrinsic reason for X 's ϕ -ing, a contrary unconscious belief may be attributed to him by comparing his behaviour with other

occasions on which he has a conscious belief that there exists an intrinsic reason, and with occasions on which he has a conscious belief that there exists no intrinsic reason. Further evidence for an unconscious belief might be sought by comparing Y's behaviour with that of other people on the occasions in which they express a conscious belief in the existence or inexistence of an intrinsic reason for action. This latter type of evidence might be adduced to claim that whenever Y has a conscious belief in the existence of an intrinsic reason, a contrary unconscious belief annuls its effect, i.e. that Y suffers from self-deception in assenting to the proposition 'There exist intrinsic reasons'. It is this systematic self-deception that is more plausible than isolated instances of self-deception.

If Y's conscious belief is that there does not exist intrinsic reason for X's ϕ -ing, a contrary unconscious belief may be attributed to him by comparison with his own behaviour and that of other people on occasions on which a belief in the existence or inexistence of an intrinsic reason is expressed. Again, the more plausible variety of self-deception would be systematic, in which Y claims that he never believes that there exist intrinsic reasons for action (i.e. doesn't believe in intrinsic reasons). In such a case, the behavioural evidence for imputing to him unconscious beliefs in the existence of intrinsic reasons obviously cannot include comparison with occasions on which he has such beliefs consciously.

The behaviour cited as evidence for an unconscious belief will seldom be linguistic behaviour since, in general, Y's linguistic utterances can be expected to be consistent with his expression of the contrary conscious belief. It is primarily certain of Y's actions and feelings that will be taken as evidence for Y's possession of an unconscious belief in the existence of an intrinsic reason for action. And correspondingly it will be the absence of those actions and feelings that will suggest an unconscious belief in the inexistence of an intrinsic reason for action.

Feelings that accompany thoughts about actions in the sphere of morality originate in complex ways through imitation of authoritative figures; Freudian mechanisms such as the internalisation of parent's standards and the acquisition of new motives in resolving the conflicts of the oedipal dilemma; and conditioning and reinforcement ranging from the bestowal and withdrawal of affection to the administering of pleasures and pains. But the facts of Y's early psychological development will play a role anyway alongside his beliefs, in explaining his judgements, actions, and feelings; and there will be nothing gained from duplicating their role by introducing a further unconscious belief. If an unconscious mental state were to be postulated, it would be difficult to distinguish an unconscious belief that there is intrinsic reason for X's ϕ -ing from an unconscious desire for X to ϕ .

Certainly it would be implausible to suggest that the existence of a set of feelings alone entailed the existence of a belief, as the phenomenon of residue guilt shows: The feelings associated with a belief that there is intrinsic reason to act, e.g. in accordance with considerations of etiquette, drop away rapidly but not instantaneously after a person consciously abandons that belief. To postulate a rapidly waning unconscious belief to account for the residue guilt would seem unnecessary if not absurd. I shall therefore continue to take a person's own testimony as the sole criterion of whether or not he believes in the existence of an intrinsic reason to act. The importance of this cognitive factor in accounting for his moral behaviour will be discussed in Chapter V.

2.10 An Objection to the Definition of Intrinsic Reason

The objection might be raised to the account of the distinction between two types of reasons for actions, as thus far presented, that the mere presence of a desire is not sufficient to guarantee the existence of a reason for an action which would be conducive to the satisfying of that

desire. A property must be ascribed to the object, perceptible to the prospective agent, whereby it becomes intelligible that that object could be the object of a desire. This amounts to a demand for a desirability characterisation of the object in Anscombe's sense (Intention § 37) It might then be claimed that the attribution of the property to the object may sometimes, if not always, serve as an intrinsic reason. The lack of firm criteria for deciding when this is so, indicates that the distinction may be too vague for the purposes for which it was intended.

Concerning the question of whether a desire needs amplifying with a desirability characterisation (i.e. a description of the object that provides a satisfactory answer to the question as to why the object was desired) in order for it to serve as a reason for action, it should first be remarked that Anscombe intended the possibility of finding a desirability characterisation to serve as a necessary condition for the putative desire to count as a desire. Once such a characterisation has been offered I see no reason why it should be denied that the desire provides a reason for action. The examples that might be advanced against this view, such as that of desiring another cigarette when trying to give up smoking, can easily be treated as cases where the desired object has something to render it desirable (such as that it would relieve irritation); and thus the desire provides a reason for action, though that reason is easily overridden by a more serious reason for refraining from that action.

As for the charge of vagueness, given that a desire requires the support of a desirability characterisation, it may be conceded that there is a vagueness concerning whether or not this desirability characterisation counts as an intrinsic reason. This renders vague some propositions of the

'There exists intrinsic reason for X's ϕ -ing' owing to a vagueness in the word 'reason'. But because some of these propositions will be non-vague, the general proposition 'There exist intrinsic reasons' will remain non-vague.

2.11. Modification to the Definition of Intrinsic Reason

A modification has to be introduced into the distinction between intrinsic and non-intrinsic reasons, in order that certain desires that are considered morally relevant should be allowed to appear essentially in an intrinsic reason. Mackie provides the example

You ought not to go in for school teaching if you are strongly attracted sexually to young children.

Ethics P.82, Penguin 1977

Here there is clearly a reading in which the judgement can be taken as a categorical imperative and yet the sexual desires of the agent, or the agent's sexual interest, occur essentially in the reason for action. Following Mackie, this can be accommodated by modifying the definition of intrinsic reason to 'that which is independent of any desire of the agent's to whose satisfaction the recommended action would be conducive.'

Further modifications to the distinction begin to appear necessary however when one looks at categorical imperatives associated with duties towards the self. It might, for example, be said of an overworked doctor that he ought to take a rest for the sake of his own health. This could be intended as a categorical imperative despite the fact that his own good health is certainly something that he desires. The definition yields the correct result, as the reason would be said to exist even if the doctor had no desire to protect his health.

If the desire in question cannot be regarded as a need but is particular to the individual, including his more mundane interests, the case is rather more difficult to accommodate. Suppose, for example, that a tennis enthusiast, non professional, has been practising for some local tournament he has entered and has been greatly looking forward to it. Then on the opening day, he remembers an invitation to a family wedding, at which his presence will be expected, and which, if he attends, will cause him to be scratched from the tournament. A friend might advise him to forget the wedding. Here the judgement 'You ought to play in the tournament' could be meant categorically, yet the reason for playing contains essential reference to a basic interest of his which would be furthered by the proposed action (provided that guilt and worry would not destroy his prospects of enjoyment). If the agent did not have this interest in tennis, the reason would not exist. What we have here is a complex reason for action constructed from two reasons, one intrinsic and one non-intrinsic, both of which, taken alone, could underlie an ought judgement. It is important that the non-intrinsic reason is asserted to be dominant as the same problem does not occur when the balance is reversed. If it is asserted that he ought to attend the wedding in preference to the tournament, then a full specification of the reason will contain reference to his desire to play tennis. Here the desire is thwarted by the action that is urged, so this example escapes misclassification on account of the first modification. It may also escape on account of a direct application of the original definition. If the person did not have the desire to play tennis, the judgement very obviously would not be withdrawn.

To accommodate such cases as the first judgement given in association with the tennis example, an intrinsic reason could be redefined as that which contains a partial reason which is independent of any present desire of the prospective agent; to whose satisfaction the recommended action would be conducive.

A partial reason may be defined as any part of the whole reason that could constitute a reason for an action by itself.

Further examples not captured by either of these modifications might be constructible in which a person has several desires and it is judged of him categorically that he ought to satisfy one in preference to others which in fact he would prefer to satisfy. Or it might be that someone has a desire but would prefer not to satisfy it in a way that someone judges categorically that he ought. Perhaps these cases are assimilable. An example of the second type might be something like 'If you enjoy mathematical manipulations you ought to become a mathematics teacher in a comprehensive school.'

As an example of the first kind, consider someone with two desires, which are in fact great interests of his such as horse riding and medicine. Suppose he has the ability to become a professional jockey or a doctor but has not the time to do both. Someone might judge of him that he ought to become a doctor, even though his preference is for becoming a jockey. Suppose, also that it is not his ability alone that prompts the judgement, and that it would be withdrawn if he didn't also have some interest in medicine. Now, the intrinsic reason that it is being claimed exists for the person's becoming a doctor contains essential reference to his desire to pursue medicine. As there is a desire to whose satisfaction the recommended action would be conducive, this example escapes correct classification after the first modification to the definition of categorical imperative. But is it correctly classified after the second modification? Is there any partial reason independent of the troublesome desire to pursue medicine? In this case his ability to become a doctor can be taken as the partial reason even though it doesn't suffice on its own as a

reason in an overall judgement, so the threat of a counter-example has been removed. The concept of a partial reason will receive further elaboration in the next chapter.

2.12 Comparison of Definitions of Modern Writers.

It may be helpful to conclude this chapter by comparing my distinction between intrinsic and non-intrinsic reasons that I take to underlie the distinction between categorical and hypothetical imperatives, with those of McDowell, Williams, Foot, Mackie and Nagel. The term intrinsic reason is borrowed from Mackie and used in the same sense as he uses it, except for the final modification mentioned above. The distinction stripped of the two modifications given above matches exactly that between external and internal reasons given by Williams (Internal and External Reasons, R. Harrison (ed.) Rational Action: Studies in Philosophy and Social Science. P.17) Nagel draws a distinction between objective reasons and subjective reasons in 'The Possibility of Altruism' which he later renames as a distinction between impersonal and personal reasons. He defines an impersonal reason as one which can be given a general form which does not include an essential reference to the person to whom it applies (Lecture I on Ethics delivered at Princeton 1979, p.6) Personal reasons he then takes as either autonomous, stemming from the agent's desires, or deontological, stemming from the claims of individuals with whom he has particular bi-personal relations. (Lecture II, p.3). Nagel's definition of personal, autonomous reasons thus corresponds to my unmodified definition of intrinsic reasons.

Foot's final conclusion (in 'Morality as a System of Hypothetical Imperatives', Philosophical Review LXXXI 1972) as to a definition that would allow moral judgements to be classed as categorical and the judgements of etiquette as

hypothetical is probably best derived from her statement,

Considerations of etiquette do not have any automatic reason-giving force,.....by contrast, it is supposed that moral considerations necessarily give reasons for acting to any man.

op. cit. p.309

As she talks of the possibility of statements and judgements as well as considerations as giving reasons for acting, some may interpret her as claiming that categorical imperatives are statements that themselves necessarily give reasons for acting to any man. As McDowell and Phillips point out, anyone thinking that his statement 'You should do this' itself gives reason to act, could not consistently see himself challenged by the retort 'Why should I?' He would be reduced to answering 'You just should'. Few people see such imperatives as themselves giving reason for action in this way. Even in the case of a command, the reason for obeying it is not the command itself, but more plausibly a fact concerning not just the command but also the relations of authority between the person issuing the command and the person receiving it.

It is more likely that Foot means to take a categorical imperative as a statement that expresses that there are considerations which necessarily give reason for acting to any man. The phrasing of this indicates that she takes general judgements as central. The interpretation of her use of 'categorical imperatives' here depends on the sense in which she uses 'necessarily'. If it can be conceded that 'considerations that necessarily give reason' may be replaced by 'fact that constitutes intrinsic reason' and that her account may be extended to cover particular judgements, then it will match that which I have offered. (I shall argue in the next chapter that Foot was mistaken in supposing that on this definition the judgement of etiquette would be rendered hypothetical imperatives.)

McDowell introduces a distinction between hypothetical and non-hypothetical reasons for action as follows:

We have then an apparent contrast between two ways in which an agent's view of how things are can function in explaining his actions. In one, exemplified by the case of taking one's umbrella, the agent's belief about how things are combines with an independently intelligible desire to represent the action as a good thing from the agent's point of view. In the other, a conception of how things are suffices on its own to show us the favourable light in which the action appeared.

Are Moral Requirements Hypothetical Imperatives? in P.A.S.S. 1978. p.22.

This contrast can be extended to cover a prospective agent's view of an action whether or not he eventually performs it, and to a person's view of other people's potential or realised actions. A hypothetical imperative then becomes a judgement expressing the existence of a conception of the circumstances which together with a suitable desire constitutes a reason for action. A categorical imperative becomes a judgement expressing the existence of a conception of the circumstances which by itself constitutes a reason for action.

From this passage it can be seen that McDowell takes a conception of how things are as basic and as constituting a non-hypothetical reason, and then defines an independently intelligible desire which is to serve in a hypothetical reason, derivatively or secondarily. This reverses the order of priority that I chose. I took the concept of desire as basic and defined an intrinsic reason as one which could be specified without reference to desires in the broadest sense of the term. McDowell's phrase 'a conception of how things are' is misleading for it offers no hint as to whether desires should in general be excluded from the conception, and as to the type of desires that may be included. The phrase 'independently intelligible desire'

likewise could be misleading because it would seem to exclude desires that are triggered off by an agent's conception of the situation. For example, a person's desire for an ice cream may not be intelligible independently of a conception of the situation that that person forms after hearing two and a half bars of Rudolf the Red-nosed Reindeer. Such desires are often unintelligible independently of a person's belief that the means are available to satisfy that desire. Until the notion of 'a conception of the situation' has been further refined, it would have to be taken to include such means/end facts. But this would sometimes render non-hypothetical a person's reason for buying an ice cream which would commonly be regarded as a paradigmatic case of a hypothetical reason. Thus I conclude that McDowell's distinction between hypothetical and non-hypothetical reasons makes use of a misleading conceptual priority of reasons over desires, and thus misses his intended mark.

CHAPTER III. THE VARIETIES OF REASON AND THEIR
ONTOLOGICAL STATUS

3.1. Aim of Chapter.

The analysis of categorical imperatives of the previous chapter dealt with variations of tense and person but very little mention was made of variations of generality and strength. In this chapter the relation between singular and general judgements will be discussed along with the relation between all-things-considered judgements and judgements made about an aspect of the situation, concerning what I called 'partial reasons'. To facilitate this analysis I shall first undertake an enquiry into the type of entity with which the reasons are to be identified, i.e. the ontological status of intrinsic and non-intrinsic reasons. This will also serve as an essential prelude to an assessment of Mackie's argument from queerness against the existence of intrinsic reasons.

3.2. The Aristotelian Syllogism

In examining what type of entity the reasons are to be identified with I shall look first at Aristotle's doctrine of the practical syllogism. Reasoning is represented by Aristotle as a consideration of two premisses which together yield a conclusion. In the case of theoretical reason the conclusion is a proposition which must be believed if the major and minor premisses are believed and there is no other theoretical syllogism whose premisses are believed that leads to a contrary conclusion. In the case of a practical syllogism the major premiss is a universal proposition connecting some clearly non-evaluative predicate with a predicate which is evaluative or indicates suitability for potential agents of a certain type, e.g. dry food suits any man. The minor premiss is a fact or set of facts indicating the obtainability of whatever was

mentioned in the major premiss. Then, if the person deliberating is of the type indicated in the major premiss as suited to the thing in question (and this may be mentioned explicitly in the minor premiss), he will arrive at the conclusion that in the absence of a syllogism leading to a conflicting conclusion, is an intention to act, which barring incontinence will result in action.

Looking to the practical syllogism for an identification of the reason presupposed in judgements of the form 'X ought to ϕ ', it would seem that neither premiss alone could constitute a reason, and that the only option available would be to identify the reason with a combination of major and minor premisses.

As the practical syllogism is offered as an explanation of action and as an account of the process of deliberation leading to action, it is designed to apply when the agent believes the premisses and has deliberated accordingly. For a complete discussion of reasons presupposed in judgements of the sort 'X ought to ϕ ', cases where X does not believe the premisses or where the premisses are false would need to be considered.

The doctrine of the syllogism can presumably be extended to cover those syllogisms which might be contemplated by someone who does not, in fact, believe or know the premisses, and also to cover syllogisms with false premisses that are sometimes acted upon or at least deliberated upon.

The doctrine of the practical syllogism, however, contains no mention of a distinction which could be used to separate intrinsic from non-intrinsic reasons. In Aristotle's examples, a desire in the broad sense engages with the object mentioned in the major premiss unless there is a conflicting syllogism. The doctrine could easily be extended to cases where there is no such engaging. But to isolate intrinsic reasons, a way of distinguishing major premisses that apply whether or not the object is desired in the broad sense is needed. The obvious way to do this would be to include reference to the potential agent's desire in one type of reason. But the combination of major and minor premisses

of an Aristotelian syllogism can contain no mention of anything's being desired. Hence Aristotle's account of practical reasoning is of little service for the purpose at hand.

3.3. Explanatory and Guiding Reasons.

Making a fresh start in searching for an identification of reasons, I shall begin with the non-intrinsic, or hypothetically imperative reasons. Assuming it has been granted that all genuine desires yield reasons (at least partial reasons) for action, it would most obviously be said that there exists reason for X's ϕ -ing, and X ought to ϕ , when ϕ -ing would lead to the satisfaction of one of X's desires, and X knows this. But what if ϕ -ing would lead to such satisfaction but X does not know this? In such cases the ought statement entails the reason statement as expected. An example can be found in a person's desire to recover from a particular ailment. It would be said that there is reason for him to take penicillin (say) and that he ought to take penicillin if this would cure him, even if he is unaware of the effects of the drug.

Now what if X merely believes that ϕ -ing will lead to the satisfaction of his desires when in fact it will not? Here I think the ought judgement is inappropriate. A person would, for example, think that he had been wrong in judging that he ought to go to the Albert Hall on discovering that the Promenade Concert he was interested in was to be held at the Roundhouse. Similarly he would not expect others to judge that he ought to go to the Albert Hall, or to tell him to do so merely on account of his misapprehension. But in both these cases it would be acknowledged that the person with the mistaken belief did have reason to act the way he did, despite it being contrary to his interests. In such cases, the ought statement 'He ought to have gone to the Albert Hall' appears not to entail the reason statement 'He had reason to go to the Albert Hall.'

This complexity can be accommodated by way of a distinction between X's reasons, the reasons that may be invoked to explain an action, and the reasons, the reasons whose existence is presupposed in ought-judgements. As such judgements may sometimes be used to guide action, I shall refer to these reasons as guiding reasons. Explanatory reasons may now be identified with some desire of X's combined with his belief that ϕ -ing will lead to the satisfaction of his desire. Guiding reasons may be identified with X's desire combined with the fact that ϕ -ing will lead to its satisfaction. This differs from the major and minor premisses of a practical syllogism in taking X's desire for an object as opposed to some statement indicating the desirability of objects of that type, and in taking a belief or fact that ϕ -ing leads to the satisfaction of that desire as opposed to a simple statement that ϕ -ing is an available option. It is natural now to stipulate 'X has reason to ϕ ', when explanatory reasons are available, and keep 'There exists reason for X's ϕ -ing' for guiding reasons. This also preserves the entailment from 'X ought to ϕ ' to 'There is reason for X's ϕ -ing', whilst denying the entailment from 'X ought to ϕ ' to 'X has reason to ϕ '.¹ As has been shown, there is no entailment either way between 'X has reason to ϕ ' and 'There exists reason for X's ϕ -ing'.

Though it has been shown that there is in general no entailment from 'There is reason for X's ϕ -ing' to 'X ought to ϕ ', there are no cases with non-intrinsic reasons where there exists reason for X's ϕ -ing which would count as demanding too much of X for the judgement 'X ought to ϕ ' to be asserted of him, corresponding to cases of superogation with intrinsic reasons. Hence the entailment holds for all but the cases where the reason is so slight that even in the absence of any conflicting reasons, the corresponding ought judgement would not be used, and the cases where the reason is easily overridden.

1 See note p.22

I shall now examine these logical relations between statements with regard to intrinsic reasons. The same distinction between guiding reasons and explanatory reasons can be applied. The two parts of the guiding reason can be telescoped into a fact that by itself constitutes a reason for action. Likewise, the explanatory reason will be a single belief about some feature of the situation.

It might seem, however, that there are circumstances under which false propositions may constitute intrinsic reasons for a person's acting. A person confronted by a beggar might come to believe the proposition that the beggar is hungry and penniless. Consequently, it might be said that this proposition constitutes a reason for the person to give the beggar some money, and hence that there exists reason for the person to give money, and that he ought to give money. Intuitions are probably divided as to whether these last two statements should be withdrawn if, unknown to the agent, the episode turns out to be the prank of some quite wealthy person disguised as a beggar. If it is thought that they should not be withdrawn, the reason for giving money could be identified with a fact involving the man's appearing to be a hungry penniless beggar. I believe it would be possible to construe all such cases in terms of facts about the appearance of the situation.

As with non-intrinsic reasons, 'X has reason to ϕ ' does not entail 'There exists reason for X's ϕ -ing', otherwise it could be said that X's belief concerning some feature of the situation, no matter how strange or repulsive, could constitute an intrinsic reason for him to act. Similarly, 'X has reason to ϕ ' does not entail 'X ought to ϕ ' in cases in which the explanatory reasons are neither supererogatory nor extremely weak, otherwise it could be said that X ought to ϕ whenever he believed that he ought to ϕ , no matter how strange or repulsive that belief.

The entailment from 'X ought to ϕ ' to 'There exists reason for X's ϕ -ing' was argued for with regard to reasons in general, and so the same argument applies a fortiori for intrinsic reasons. However, the reverse entailment in non-exceptional cases may be thought not to hold. To take Mackie's example, if it is to be supposed that the fact that someone is starving at your doorstep is a reason for doing something about it, it would seem that this is insufficient to say that you ought to do something about it, without adding a further condition, indicating that this fact could be known. This oddity may be acknowledged as a special property of intrinsic reasons, or it may be thought preferable to preserve the parallel between non-intrinsic and intrinsic reasons by requiring that the fact specify the prospective agent's role in the situation. This would mean rejecting such a fact as that someone is starving at X's doorstep as a reason for X's acting, unless it is filled out with some mention of X's position at the time. Mackie's other example, 'someone is writhing in agony before your eyes' would satisfy this requirement on facts which are to constitute reasons for action.

On the other hand, it might be thought that as the entailment from 'There is reason for X's ϕ -ing' to 'X ought to ϕ ' does not in general hold, it will not matter how the entailment fares in the non-extreme cases, and hence that it is perfectly acceptable to talk of facts which may constitute intrinsic reasons for X's ϕ -ing independently of X's knowledge of the circumstances. Following this line of thought, the entailment from 'There exists reason for X's ϕ -ing' to 'X has reason to ϕ ' fails, as in the case of non-intrinsic reasons, and the remaining entailment from 'X ought to ϕ ' to 'X has reason to ϕ ' can be seen to fail too. For it would certainly not be wished that 'X ought to ϕ '

entailed 'X believed there existed reason for his ϕ -ing' which would be a consequence of the prima facie plausible entailment from 'X has reason to ϕ ' to 'X believes there exists reason for his ϕ -ing' to be discussed shortly. If this line of thought is to be adopted, the six entailments that may be constructed from 'X ought to ϕ ', 'There is reason for X's ϕ -ing' and 'X has reason to ϕ ' are all invalid except for that from 'X ought to ϕ ' to 'There is reason for ϕ -ing', and this pattern holds for both intrinsic and non-intrinsic reasons.

3.4. Are Intrinsic Reasons Linguistically Expressible?

I shall consider now the entailment that was used unsupported above from 'X has reason to ϕ ' to 'X believes there exists reason for his ϕ -ing'. This is a particular instance of the more general entailment that if there are facts that constitute intrinsic reasons for action, then it follows that anyone believing such a fact must also believe that it does constitute intrinsic reason for action. Suppose that the entailment holds, and that 'a man is starving in front of X's eyes' constitutes an overall ^{1.} intrinsic reason for X to help the starving man. Now if X believes the fact, he must ex hypothesi also believe there is reason for him to help the starving man. If he then does not help he must be convicted of acting irrationally, an indictment that few find appropriate for the immoral man. If this conclusion is deemed unsatisfactory one of the premisses of the argument must be rejected. Suppose this to be the premiss that if X believes the fact, he must also believe there is reason for him to help. Then X's belief of the fact could not suffice as an explanatory reason for his action were he to help. If more were needed for a person to see reason to act than just a belief about some feature of the circumstances, then more would be needed by way of an explanatory reason.

1. For a definition of 'overall reason' see 3.5.

The more favourable alternative is to accept that anyone believing a fact that constitutes a reason for action must also believe that it does constitute a reason for action. Then the other premiss must be rejected, namely that 'a man is starving in front of X's eyes' constitutes an intrinsic reason for X to help. Assuming that there are overall intrinsic reasons, and that one is in the offing in this example, what can it be? The reason cannot just be a conjunction of the fact that a man is starving in front of X's eyes with the further fact that this constitutes a reason for action, as this would involve a form of self-referential incoherence. A final alternative is to reject the sentence 'a man is starving in front of X's eyes' as an expression of the fact that constitutes intrinsic reason, and as the argument would apply to any similar sentence, to reject any linguistic expression of such facts.

McPetridge explores one way of making sense of the idea of facts that cannot be linguistically characterised, namely by analogy with perceptual states.

We think of perceptual states as cognitive, as capable of representing how things are, while allowing that how, exactly, they represent things to be may exceed our capacity to state.

Are Moral Requirements Hypothetical Imperatives? in P.A.S.S. 1978. p.37

He then appears to put paid to this suggestion by adducing the evidence that agents frequently see circumstances as demanding a certain action from them after merely having been told of those circumstances. Platts explains this whilst holding onto the perceptual analogy by means of a distinction between concepts and conceptions.¹ Two people may agree on the facts of the situation yet disagree as to whether those facts constitute a reason for acting. Thus

1 'Moral Reality and the End of Desire' in 'Reference, Truth and Reality' in M.Platts ed. R.K.P. 1980

Platts would attribute to difference in depth of conception of some central concept appearing in the facts of the case i.e. to a difference in recognitional capacities of instances of the concept. So in both MacDowell and Platts, the intrinsic reason is to be identified with a fact about the situation. For McDowell (in 'Are Moral Requirements Hypothetical Imperatives?') this fact cannot be expressed in a sentence. For Platts, it can, though knowledge of the fact cannot be infallibly transmitted linguistically.

It is probably as a response to McPetridge's suggestion that a conception of a situation should be capable of linguistic expression that McDowell has developed his views into a two-tier system in 'Virtue and Reason'. At the first stage, the agent's conception of how to live manifests itself in the perception of a salient fact in a situation, this salient fact selecting one of the agent's concerns. The concern and salient fact then act as major and minor premiss in a practical syllogism. In seeing the fact as salient, the agent sees it as a reason for acting which silences all others.

Perhaps it will be possible on McDowell's view to fully formulate both major and minor premisses linguistically. What will not be formulable is the agent's conception of how to live which cannot be understood from an external standpoint.

Any attempt to capture it in words will recapitulate the character of the teaching whereby it might be instilled. Generalisations will be approximate at best, and examples will need to be taken with the sort of "and so on" which appeals to the co-operation of a hearer who has cottoned on.

Thus it is explained how different people's conceptions of how to live cannot be understood without being shared, yet the particular fact which a person sees as giving reason to act may be believed by someone else who does not see it as giving reason to act. This explains how someone may see reason to act after merely being told of the situation, which, as McFetridge pointed out, constituted a problem for the view that a conception of the situation which by itself gave reason to act was analogous to a perceptual state.

It may appear that McDowell means to identify the conception of the situation with the prospective agent's belief about the situation when he writes 'a conception of how things are suffices on its own to show us the favourable light in which the action appeared'. This is only because the quoted passage refers to an act which the agent has performed. To explain an act *ex post facto*, it is obviously insufficient to give a fact constituting a reason for the act without including the agent's belief of that fact. When McDowell's contrast is extended (in the manner suggested in 2.12) to cover cases in which Y is viewing the potential actions of X, a conception that suffices as a reason for X's action need only be regarded as a fact about the situation, not necessarily believed by X. I argued (in 2.12) that McDowell's notion of a conception of a situation that by itself constitutes a reason for action, was insufficiently spelled out to indicate the circumstances under which it may include desires. From now on, whenever I use this expression, it will be in the sense of 'fact that constitutes an intrinsic reason for action'. And in interpreting McDowell it should be borne in mind that he uses the expression both to mean 'fact that by itself constitutes reason for action' and to denote someone's belief of such a fact.

It will be convenient to use both accounts of McDowell's as the second is an elaboration upon rather than a

renunciation of the first. What McDowell initially regarded as a conception of the situation which constituted a reason to act but could not be linguistically formulated, is later split into a salient fact and concern which constitutes reason to act, and which can be linguistically formulated. As it is the conception of how to live that selects the particular fact as salient, this cannot be left out of the equation.

Though McDowell has in mind the virtuous person when talking of the agent's conception of how to live, it should be remarked that this type of explanation, if applicable at all, will apply equally to those whom the virtuous person would regard as having a slightly or indeed very warped conception of how to live. This is just an iteration of the familiar argument from relativism, with which any defense of intrinsic reason must deal. And it is also worth pointing out that on this view access to the requirements of morality is only gained through an elaborate training and self-education.

It is an important claim of McDowell's that someone need not manifest irrationality in failing to see reason to act as morality requires. By failure to see reason to act, I take McDowell to mean failure to come to the belief that there is reason for him to act.

Failure to see reason to act virtuously stems, not from the lack of a desire on which the rational influence of moral requirements is conditional but from the lack of a distinctive way of seeing situations.

Op. cit P.23

This point is contested by Williams:

What is it that one comes to believe when he comes to believe that there is reason for him to ϕ , if it is not the proposition, or something that entails the proposition that if he deliberated rationally, he would be motivated to act appropriately?

Op. cit. P.24

McDowell would have to reply, I think, that it is that given a suitably developed perceptual capacity or conception of how to live, an agent would be motivated to act appropriately if he deliberated rationally. An agent with this sort of perceptual capacity would certainly be motivated to deliberate, contrary to Williams' thought (op. cit. p.24), by his desire to find the right act in the circumstances, this desire being trivially attributable as a result of a belief of his about the world. McDowell would probably agree with Foot and Williams in taking claims that rationality may impose demands upon a prospective agent irrespective of his desires, as bogus or bluff. To support McDowell's position, the acquisition of the perceptual capacity still needs justification, as does the special status of the virtuous person's conception above all the near-misses and totally non-virtuous conceptions, to which I shall return in Chapter VI.

In conclusion, an intrinsic reason for action is to be identified with a fact that is either non-linguistically expressible, or it is linguistically expressible but is only recognised as an intrinsic reason for action by those with some special conception of how to live, or special perceptual capacity which itself is not linguistically expressible (like someone's driving style). In the moral case this fact will be a statement about the desires or needs of someone other than the agent, or a statement implying the presence of such desires or needs. In the prudential case it will be a statement about, or implying the presence of, some future desire or need of the agent. The possibility of further categorical imperatives that are neither moral nor prudential will be discussed in 3.11.

3.5. Overall and Conclusive Reasons

Returning to the question of the variation of strength of categorical imperatives, it was remarked in 2.2. that there are a host of linguistic forms which reflect different degrees of strength but which cannot consistently be ranked in order of strength. The analysis in terms of reasons

applies to all of these. The reason for action may be very weak and easily overridden, extremely strong and unlikely to be overridden, or anywhere between these extremes. In all cases the judgement will be taken as implying that there exists a fact about the situation which constitutes a reason to act. The analysis of reasons thus far as facts about the situation is well suited to accommodate cases where different reasons for and against a given action coexist. In such cases, different reasons are just facts about different aspects of the situation. The existence of reasons of this type is what is claimed to be asserted in a categorical imperative of what I chose in 2.2. to call the second kind.

The reason asserted to exist in an imperative of the third kind must, however, take all aspects of the situation into consideration and I shall call it a conclusive reason. The idea of an all-things-considered fact about the situation which is capable of being conceived and hence perhaps believed, is a far more recondite notion than an ordinary fact about the situation since there is no way of capturing it linguistically. However, since the ordinary facts under consideration have turned out not to be linguistically expressible on McDowell's first view, this constitutes no further objection to the analysis. On McDowell's second view, the fact that is singled out as salient by the conception of how to live is well-suited to be a conclusive reason, since other reasons have been overridden in selecting one fact as salient. It is more problematic on McDowell's second view to account for judgements concerning obligations that are easily overridden and to account for moral conflicts.

A judgement which is the best in the circumstances, i.e. that takes all things at hand into consideration, may have been arrived at after balancing judgements of the second kind, and thus can be regarded as a composite judgement. I shall refer to such all-things-at-hand-considered judgements as overall judgements, and the reasons for action whose existence they entail, I shall call overall reasons.

Overall judgements may be compatible with conflicting ought judgements based on evidence unavailable to the maker of the overall judgement. It may also turn out that there is no other evidence which would produce a conflicting ought judgement. Thus an overall reason may be a conclusive reason but the maker of the overall judgement will not know whether it is. 'Overall reasons' and 'overall judgements' are epistemic notions independent of the metaphysical notion of 'conclusive reason'.

I shall consider next the ways in which reasons combine to form composite reasons.

3.6. Conflicts Between Intrinsic Reasons

I have suggested as analysis of moral conflicts, or more generally conflicts between intrinsic reasons, that different facts about the situation may yield reasons for incompatible actions in that situation. But I left open the question of whether such reasons may be weighed against each other to produce a strongest reason, and hence a favoured course of action in the circumstances. Clearly there are cases where this does occur and these are best accommodated in terms of a merging of the two reason-giving facts to produce a single fact yielding a single reason for action. In experiences of such conflict situations the weaker obligation may not be completely annulled as the agent often feels some regret and tries to make up for his unfulfilled obligation. This is not absent from the analysis offered since there still exists the fact which gave reason for the weaker obligation, so that the outweighed reason still counts as a reason. (I shall be using 'outweighing' in the same sense as 'overriding'.) There are also cases where the two obligations seem equally balanced. Here, either it can be claimed that one must outweigh the other though it may be beyond the epistemic

capabilities of the agent to tell which, or it can be claimed that a single fact can yield reasons for a plurality of incompatible actions.

To accommodate such conflicts on McDowell's second view it would be necessary to allow the conception of how to live to pick out several salient facts, some of which may be stronger than others. To account for the overriding of some by others, the conception of how to live would have to pick out some facts as overriding and others as overridden but not silenced.

3.7. Conflicts between Intrinsic & Non-intrinsic Reasons.

I shall turn now to a discussion of conflicts between intrinsic reasons and non-intrinsic reasons. Beginning with the problem of forming an overall judgement in the conflict situation, McDowell suggests,

part of the point in claiming that the requirements of virtue are categorical imperatives may lie in a rejection of the possibility that the dictates of virtue may be outweighed by reasons for acting otherwise.

Op. Cit. p.26

But it is clear that the dictates of virtue are sometimes outweighed by non-intrinsic reasons. I might have arranged an appointment with someone who will have to spend some hour getting to and from the venue, and be totally unable to cancel the appointment. I may then view this as giving me reason to attend the appointment. This outweighs my slight desire not to go because of a toothache. But the situation can be adjusted by worsening my toothache up to the point where this outweighs my obligation to keep the appointment. Again this doesn't annul the reason for keeping the appointment. It can still truly be said of me, on this view, that I ought to keep the appointment. McDowell's suggestion could then be treated as a terminological point -the cla

that categorical imperatives imply the existence of only those reasons for action that are not overridden. But this would require finding a term for those judgements implying reasons for action that are overridden, so nothing would have been gained. The oddity of such a terminological move is clearer when general judgements are considered. The judgement 'people ought to keep their promises' is not thought non-categorical due to its sometimes having exceptions. So I shall treat non-intrinsic reasons as occasionally, but not usually, overriding intrinsic reasons.

A composite judgement is formed in situations in which it is judged that there exist both intrinsic and non-intrinsic reasons for possible actions, by combining the fact that constitutes the intrinsic reason with the prospective agent's desire and fact that the means to satisfying that desire are available. (The desire need not be for an action opposed to that for which there is intrinsic reason, as could be seen in the horse-racing doctor example.) The composite fact will then constitute an intrinsic reason, in most cases for the action enjoined by the partial intrinsic reason or the partial non-intrinsic reason. As in the case of conflicting intrinsic reasons, the two reasons may seem equally balanced. To accommodate this, either it can be regarded that a composite fact may yield reasons for a plurality of incompatible actions or it could be claimed that the verdict is only impossible to find in practice, and that the apparent need to accommodate plurality can be satisfied by reference to incomplete knowledge for the weighing to be done. One of the reasons will outweigh the other, the outcome arising immediately from the compound fact, so that no further explanation of why there is reason for the enjoined act is

necessary beyond the compound fact itself. Whichever partial reason loses out will still constitute a reason, as in the case of the weaker of two obligations. I have preferred the term 'partial reason' for a reason that may form part of a compound reason, to the frequently used 'prima facie reason' because the latter term has connotations of non-genuineness. The compound reason may itself be treated as a partial reason if other reason-constituting facts exist, known or unknown. Otherwise, the compound reason will be a conclusive reason. I shall continue to talk simply of reasons, only specifying whether they are partial or conclusive when necessary.

3.8. Silencing Versus Outweighing

On McDowell's second view, the conception of how to live takes into consideration the facts that recommend courses of action and singles one of these out as salient. Those facts that lose out are silenced and do not count as reasons for action at all. Silencing is also preferred to outweighing on McDowell's first view as can be seen from the following quote:

If a situation in which virtue imposes a requirement is genuinely conceived as such, according to this view, then considerations which, in the absence of the requirement, would have constituted reasons for acting otherwise are silenced altogether - not overridden by the requirement.

op. cit. P.26

Now it is not clear from this passage whether McDowell would take the dominant reason as silencing other reasons in all conflict situations including those in which virtue would impose requirements for conflicting actions (the standard moral conflict situation). If so, then he would leave himself no room to account for the emotional distress

that the virtuous person often feels as a result of having to neglect some alternative action that he would normally have felt to be required of him. If the silencing is only to occur in certain conflict situations, then some reason for this inconsistency is owed.

It should also be noted that McDowell talks in this passage of a situation in which virtue imposes a requirement, not of any situation in which a fact, or conception of the circumstances, imposes a requirement. Thus it appears that he is excluding considerations of etiquette from silencing without yet explaining the essential difference between etiquette and virtue. It might appear that he is also excluding considerations of self-interest (both prudential intrinsic reasons and non-intrinsic reasons) from silencing, until we read his caveat:

The reasons which silence are those which mark out actions as required by virtue. There can be less exigent moral reasons, and as far as this position goes, they may be overridden.

op. cit. P.29

This passage suggests that McDowell believes that overriding does occur in some conflict situations, namely those in which moral considerations are defeated by other considerations. But it is uncertain whether these considerations may also be moral considerations, or whether some non-moral considerations may be overridden. So it is still unclear on McDowell's first view, when the silencing comes into operation. More important is the question this passage invites as to how it is to be judged whether the moral reasons are exigent enough to silence others. The exigency of the reason can only be determined relative to the other considerations, hence a weighing-up must take place prior to the silencing.

This contrasts with McDowell's view that silencing replaces outweighing:

the dictates of virtue, if properly appreciated, are not weighed at all.

Op. Cit. p. 26

3.9. Incontinence

One of the merits McDowell claims for silencing compared with overriding, is that it provides a better account of incontinence. In a situation in which morality is thought to require some action and there are competing non-moral attractions as well, the continent and incontinent man's inclinations are aroused, as the temperate person's are not, by their awareness of competing attractions. The temperate man is also aware of the competing attractions but they do not engage with his inclinations because his conception of the situation, i.e. his belief of a fact that constitutes an intrinsic reason for action, silences the reasons for action which derive from those attractions. But now it would appear that the continent and incontinent man cannot share the temperate man's conception of the situation otherwise their inclinations would not engage either. So it would appear that the difference in behaviour between the continent or incontinent man and the temperate man cannot be explained in terms of a difference in conceptions.

McDowell replies:

The way out is to attenuate the degree to which the continent or incontinent person's conception of a situation matches that of a virtuous person. Their inclinations are aroused, as the virtuous person's are not, by their awareness of competing attractions: a lively desire clouds or blurs the focus of their attention on "the noble" !.

Op. Cit. P. 28

The attenuation McDowell refers to cannot be an attenuation of the extent to which other reasons are silenced. If a reason is merely muffled then it is not

silenced and it would have to be said that a conception which merely muffles differs essentially from one which silences. Hence in having the same conception, the two types of agent must believe the same reason - silencing fact. Where the two types of agent could differ is in the extent to which they attend consciously to, or actively think about, this reason-silencing belief of theirs. But it cannot be a difference in desires (a lively desire for the continent or incontinent man as opposed to a less lively one for the virtuous man) that explains the difference in attentiveness to the reason-silencing belief, because it is the lack of attention to the reason-silencing belief that allows the competing attractions to arouse the desires. Hence the silencing model does at least accommodate a coherent account of the differences in behaviour between the temperate man and the continent or incontinent man. The difference is explained neither by a difference in beliefs, nor by a difference in desires, but by a difference in ability to concentrate on the relevant reason-silencing belief. On McDowell's second view this would become a difference in ability to concentrate on the concern picked out by the conception of how to live.

Once the agent's concentration has lapsed and a desire has engaged his attention, all that is left of his categorical imperative judgement is the outweighing of the non-intrinsic reason by the intrinsic reason. So it appears that though the silencing model allows an account of temperate behaviour, the account of continent and incontinent behaviour reduces to that offered on the outweighing model. An outweighing model of categorical imperatives will be necessary in any case, it has been seen, to account for the way in which an outweighed reason may still count as a reason in producing certain types of emotion and compensatory action in conflict situations.

The silencing model also seems inadequate for other sorts of virtuous behaviour such as courage. When someone acts courageously he needs to be aware of the dangers and the reasons they provide for choice of action. A person who doesn't take full consideration of these reasons is thought foolhardy rather than courageous. A silencing of reasons after an initial careful weighing-up will seldom be appropriate in explaining courageous action either, in view of the uncertainty that often attaches to the danger and the need to take constant account of the reasons it provides for action during any courageous activity that is at all protracted.

Another reason for keeping the outweighing model is to account for the fact that some people occasionally decide to adopt a course of action whilst readily admitting that it is morally wrong, and having never intended not to follow that course of action. Such cases clearly cannot be assimilated to incontinence because no change of intention is involved. On such occasions, after incorporating his virtue-opposing desire and virtue-promoting fact into a composite fact and finding that there still exists an intrinsic reason to act virtuously, the agent chooses to act in accordance with his partial non-intrinsic reason rather than his overall intrinsic reason.

To account for this gap between judging that there exists overall intrinsic reason to ϕ and intending to ϕ , some way of balancing overall intrinsic reasons and non-intrinsic reasons must be found. As the non-intrinsic reason has already been integrated into the intrinsic reason, the only available option seems to be the balancing of the desire consequential upon the possession of the intrinsic reason with the desire opposing the intrinsic reason. Such an option would only be available on the outweighing model, since on the silencing model, the non-intrinsic reason

would have disappeared and hence would not be available for further balancing. This means that the person's ultimate decision as to what to do rests on the strength of the consequential desire, in some sense of 'strength' which an account along these lines would need to specify. It follows also, on this account, that when a person does decide to do what he thinks there is overall intrinsic reason for doing, this we must also attribute to the strength of his consequential desire. Thus all decision-making on the outweighing model appears to be ultimately a balancing of desires.

It is difficult to see how this position can be prevented from collapsing into that which dispenses altogether with intrinsic reasons. Though there will be a difference in judgements between someone using categorical imperatives who takes moral reasons as outweighing non-moral ones, and someone using only hypothetical imperatives, this difference in judgements will have no effect upon the decision taken as a result of deliberation, unless it can be shown that what determines the strength of the consequential desire for the virtuous action differs from what determines the strength of a desire in the narrow sense for the same action. But it is difficult to imagine how this could be so, since the strength of the consequential desire cannot be determined by the fact that constituted overall intrinsic reason as this would rule out the possibility of explaining how the overall intrinsic reason sometimes prevails over the opposing desire and sometimes loses out.

Having decided that there is overall intrinsic reason to ϕ , the temperate man's behaviour is best explained on the silencing model. The same decision made by the continent and the incontinent man will sooner or later best

be explained on the outweighing model, as a result of their failure to concentrate on the overall intrinsic reason. The temperate man makes a single categorical imperative judgement, from which his intention derives. The continent and the incontinent man make two judgements, the categorical imperative judgement from which their decision as to what they ought to do derives, and a hypothetical imperative judgement balancing the desire for the outweighed option with the purely consequential desire. The continent man judges that he still has overall non-intrinsic reason to \emptyset , whereas the incontinent man makes the opposite judgement. Incontinence is not now restricted to failures to act in the course of virtue. A person might have decided against the virtuous course of action yet his awareness of competing attractions could obtrude: a lively sense of obligation clouds or blurs the focus of attention on the ignoble.

For McDowell, it seems mysterious how the motivating potential of competing attractions can still have energy after being outweighed. This must be explained in terms of a difference in sense of 'strength of desires' between that involved in the forming of the original decision, when the desire may only have been anticipated rather than felt, and that involved in the changing of the decision. Neither is equivalent to the sense of 'strength' that is applied to the force of the phenomenological pull or urge, though the latter sense will be closer to it than the former.

On the silencing model, this change of intention might be thought not to require explanation in terms of difference in sense of 'strong desire'. The weighing-up prior to the silencing might be regarded as utilising a balance more favourable to the virtuous option than that used on the outweighing model. But then the incontinent man's overall judgement as to what he ought to do would change as he moves from the silencing to the outweighing model.

In conclusion, it has emerged that the formation of an overall ought judgement is at the same time the formation of an intention if the judgement is a hypothetical imperative, or if the judgement is a categorical imperative on the silencing model. The silencing model is appropriate for only a small section of virtuous behaviour, but a problem remained on the outweighing model in characterising the different senses of 'strength of desire'.

3.10. Variations of Generality.

The variations of generality of categorical imperatives can now be discussed in relation to intrinsic reasons. The judgements considered thus far have taken the form 'X ought to ϕ ' where 'X' denotes a person and ' ϕ ' is a description of an act. It was considered that singular judgements could be taken as primary, since they are used by those who consider them to be derivable from general judgements, as well as those who do not; whereas general judgements may be seldom used by those, such as McDowell, who do not take them as providing the foundation for singular judgements. Indeed, some might deny that there are any valid general judgements.

To this it might be objected that it does not matter that exceptions can always arise to any general judgement, the important point is that whenever a person makes a singular judgement he must always be able to pick out some general judgement, from which his singular judgement may be derived. For according to this view, it would be difficult to understand what could be meant by claiming that there is intrinsic reason for X's ϕ -ing without at the same time accepting that there is also reason for anyone in a similar position to X's to ϕ .

But as Winch has pointed out, (in 'The Universalisability of Moral Judgements' in *Ethics and Action*, Routledge & Kegan Paul, 1972) it is not obvious that first person ought

judgements are unintelligible unless they are universalisable. In cases where first person moral ought judgements conflict, what one person decides he ought to do need not be the same as what someone else in a similar conflict would decide to do.

A further reason for rejecting the universalisability of first person ought judgements is in accounting for the phenomenon of superogation. In cases of this, the agent invariably judges that he ought to do the act in question though he would often refrain from making the corresponding third person judgement with regard to people in similar situations. It will not be possible to explain this apparent failure of universalisability by claiming that the agent's judgements are inconsistent - that one (or both) of his judgements must be false. The fact that many language users would accept this discrepancy between first person and third person judgements means that it cannot be a logical feature of moral discourse.

If the intrinsic reason for X's ϕ -ing is specifiable not only without essential reference to X's desires, but also without essential reference to X, (i.e. it is an impersonal reason in Nagel's sense) then the judgements 'There is reason for X's ϕ -ing' and 'X ought to ϕ ' will be universalisable. In taking singular judgements as basic I am accommodating the possibility that not all intrinsic reasons are of this impersonal type, and that not all particular ought judgements are derivable from general judgements. General judgements themselves do not present a problem on this approach since truth conditions can be provided for them in terms of the truth-conditions of particular judgements derivable from them: A general non-conclusive ought judgement will be true if and only if all singular ought judgements derivable from it are

true. And a general conclusive ought judgement will be true if and only if most singular ought judgements derivable from it are true.

3.11. Etiquette

Having discussed the varieties of forms in which categorical and hypothetical imperatives occur, and noting that there are both moral and prudential cases of categorical imperative, I shall close this chapter with a discussion of the possibility of categorical imperatives that are neither moral nor prudential.

In 'Morality as a System of Hypothetical Imperatives', Foot argues that if categorical imperatives are to be construed as those that need not be withdrawn if the prospective action cannot be shown to be ancillary to the agent's interests or desires, then the imperatives of etiquette and club rules are let in on the categorical side. This point is contested by D.Z. Phillips, (in "In Search of the Moral 'Must': Mrs Foot's Fugitive Thought" *Philosophical Quarterly* XXVII, 1977) In the case of club rules he argues that the imperative 'You must obey the rules' is elliptical for 'If you want to join the club you must obey the rules.' As the imperative applies just as much to members as prospective members this should be extended to 'If you want to become or remain a member of the club you must obey the rules.' Foot's point is that the rules don't fail to apply to a member once he loses interest in his membership of the club. The imperative, as it is intended, is elliptical for 'If you are a member of the club you must obey the rules.' And here the conditional clause contains no reference to a desire, so the imperative need not be withdrawn if the agent is found to have no desire for club membership. Foot, therefore, seems correct in asserting that on this construal of the categorical/hypothetical imperative distinction, the imperatives of club rules and

etiquette¹ as well as morality and prudence are categorical imperatives.

However, this led Foot to reject the above definition of a categorical imperative as she considered it crucial that any rationale for the hypothetical/categorical imperative distinction should separate the dictates of etiquette from those of morality. Perhaps she inherited Kant's mistake of merging the notions of categorical imperative and true categorical imperative, thereby not having the option of classifying the dictates of etiquette as false. Or perhaps she overlooked the distinction between categorical imperatives asserting the existence of reasons for action and those asserting the existence of conclusive reasons for action, thereby losing the option of taking etiquette as providing non-conclusive reasons.

McDowell expresses a similar view:

Many actions performed for reasons of etiquette can be explained in terms of bewitchment by a code. There may be a residue of actions not explicable in that way. It does not seem to me obviously absurd, or destructive of the point of any distinction between categorical and hypothetical imperatives, to suppose that such residual actions might be most revealingly explained in terms of non-hypothetically reason-constituting conceptions of circumstances. One can attribute such conceptions to others without being compelled oneself.

op. cit. P.24

It should first be noted that when agents act in accordance with rules, principles, or a code, whether or not they might be regarded as bewitched, their reasons are still of the type thought of as deriving from a conception

1. There are other non-moral, non-prudential categorical imperatives that could be added to this class, for example imperatives based on traditions such as the judgement that Owen Wingrave ought to join the army, discussed by Williams in 'Internal and External reasons'.

of the situation irrespective of their desires. McDowell doesn't himself believe that the dictates of morality can be derived from a few basic rules. (Etiquette, too, is not usually thought of by those who take it seriously to be a matter of a few rules. The person who gets the rules right but lacks the style may be ridiculed.) Nevertheless, this should not prevent him from classifying the sincere moral judgements of those who do believe this as categorical imperatives.

The other feature to note arises from the last sentence of this quotation. What can it mean to attribute a conception to others, if such a conception is not linguistically expressible? McDowell must mean that it is possible to entertain precisely that proposition that is accepted by someone else as constituting intrinsic reason to act, without oneself believing it. But this opens up the possibility of a person's entertaining a virtuous conception of a situation and yet rejecting it, a possibility that McDowell would seem to want to block in the following passage:

It would be wrong to infer that the conceptions of situations which constitute the reasons are available equally to people who are not swayed by them, and weigh with those who are swayed only contingently upon the possession of an independent desire. That would be to assimilate the second kind of reason to the first. To preserve the distinction, we should say that the relevant conceptions are not so much as possessed except by those whose wills are influenced appropriately. Their status as reasons is hypothetical only in this truistic sense: they sway only those who have them.

op. cit. P.23

McPetridge shares this interpretation of McDowell and sees it as presenting too happy a view of the moral outsider:

For it rules out the possibility that someone might know how the world is taken to be by those committed to morality but simply hold that they were mistaken in thinking that things were like that. One might think that a cognitivist account of morality ought not simply to rule out such a possibility.

op. cit. P.40

If, following this interpretation, the gap between entertaining a conception and embracing it is closed, it must be assumed, for consistency, that two people do not have access to the same conception of a situation involving etiquette if one sees reason to act in accordance with etiquette and the other does not.

But McDowell also writes (last line of p.19)

....one can understand a moral outlook without sharing it.

op. cit. P.19

This suggests the alternative view, that a conception could be understood, available or entertained without actually having, possessing, embracing or believing it. If it is believed then it must also be seen as giving reason to act. This seems a more faithful interpretation of the beliefs of the user of categorical imperatives.

Thus I conclude that there are uses of categorical imperatives in domains such as those of etiquette and club rules as well as morality and prudence. Prudential intrinsic reasons are readily distinguishable from moral intrinsic reasons, as the existence of the former depends on the welfare only of the prospective agent whereas the existence of the latter depends also on the welfare of others. I shall not

undertake to distinguish other uses of categorical imperatives from moral ones. In general they derive from matters independent of anyone's welfare. Though a case might be made out for assimilating them to moral cases. It may cause considerable distress to those who live by the conventions of e.g. etiquette, club rules, or family tradition when they see these rules being flouted.

CHAPTER IV. DO ALL, SOME, OR NO PEOPLE USE
CATEGORICAL IMPERATIVES?

4.1. A Simple Answer

As a consequence of making the speaker's opinion the sole determinant of whether or not he is using a categorical imperative, the task of this chapter would appear to have been made too simple. In order to argue that some, but not all, people make use of categorical imperatives in their normal judgement, all I need do is produce someone who swears he does use them, and someone else who swears he does not. This I can easily do by citing myself and a friend of mine. This need not be seen as a *reductio ad absurdum* of the view that there is no significant sense in which unconscious beliefs in the existence or inexistence of intrinsic reasons can be spoken of. The thesis that some but not all people use categorical imperatives can be given further support by examining the views of philosophers on this question and the findings of experimental psychologists. The experimental data will also be of relevance in comparing the behaviour of those who do use categorical imperatives with those who do not. The question is best examined by treating the cases of prudence and morality separately. I shall begin with a brief examination of prudence.

4.2. The Prudential Case

The clearest cases of categorical imperatives are those in which Y judges that X ought to ϕ , knowing that ϕ -ing would not be conducive to the satisfying of any of X's present desires in the broad sense (i.e. knowing also that X does not even intend to ϕ). And the case in which I suggested that Y is most likely to have such knowledge is the case in which he is making a first person past tense judgement. Now it might be thought that users of prudential categorical imperatives might easily

be picked out by asking people whether they ever make judgements of the form 'I ought to have ϕ 'd, knowing that at the time they had no interest that ϕ ing would have promoted, and had no intention of ϕ ing, as could be said of my judgement 'I ought to have changed subjects at Cambridge years ago'. To this it might be objected, however, that such judgements are the mere expression of unrealistic wishes and not the assertion of the existence of past reasons. This is because any such putative reasons would have been unknowable then, as future interests could only have been predicted with a low degree of certainty. In reply to this, it could be pointed out that reasons are thought to exist in the present even when people are unaware of them, as was suggested in the example of someone suffering from an ailment, for whom there exists reason to take penicillin even though the sufferer is unaware that this would cure him. A rejoinder to this could be to challenge the view that there exists reason when that person could not possibly discover the reason. But why should the possibility of discovery implying a high degree of certainty, be insisted upon? There seems to be no obvious reason for this, and hence no reason why past tense first person ought-judgements should not be thought of as asserting something that entails the existence of past reasons. But as complications have arisen, such judgements cannot be taken as providing an acid test for the recognition of prudential categorical imperative users.

Perhaps a clearer test case would be the making of second person present tense judgements that a person ought to do something for the sake of his future welfare which would not further any of his present interests, and which he has no intention of doing. A possible example of this would be 'You ought to take a cold shower every morning'. By considering examples of this kind it should

be easy for people to recognise when they do use categorical imperatives in prudential judgements; and after an extensive consideration of examples, those that do not find any such judgements that they would recognise as making themselves, ought to become convinced that all their prudential judgements are hypothetical imperatives.

It is difficult to decide whether all, some, or no people use categorical imperatives in making prudential judgements because of the problem of deciding whether desires for future welfare or happiness are to be accepted as legitimate, i.e. non-vacuous. If hypothetical imperatives based on such desires are ruled out, then it would seem as though categorical imperatives are a universal feature of prudential judgement. Otherwise, it would be difficult to explain how someone could regard their former conception of their own happiness to have been warped, or someone else's conception of his own happiness to be warped.

4.3. The Views of Some Modern Philosophers.

Mackie offers an account of what he calls moral scepticism that appears to match what I have in mind by amorality. The task at hand is to find a precise description of what the sceptic is sceptical of. Mackie provides an over-simple account of this by calling the issue an ontological one: The moral objectivist believes in the existence of a special kind of entity glossed as objective value and including moral values and intrinsic reasons for action. The moral sceptic does not. There is an equally simple way of choosing between these two positions: if these entities do exist then moral objectivism is correct and moral scepticism must be rejected. Mackie himself argues that the moral sceptics have got things right, and introduces the term 'error theory' for his view that many moral judgements entail the false proposition that there exist objective

values and hence intrinsic reasons. I shall not beg the question in favour of moral scepticism by adopting the term 'error theory' but note this as an example of a view that moral judgement is bifurcated within a speech community, and in which the objectivist and hence the categorical imperative users are in the majority.

Williams too can be seen as adhering to an error theory in Mackie's sense, when he suggests that

external reasons statements, when definitely isolated as such, are false or incoherent, or really something else misleadingly expressed.

op. cit. P.26

The tenor of the paragraph suggests that he'd rather not think of such judgements as false, but as misleadingly expressing some sort of evaluative judgement. I can't help feeling that Williams is rductantly drifting towards a more subjectivist position. In denying the connection between ought-judgements and reason statements he avoids committing himself to saying that all categorical imperative statements are based upon false belief, yet his arguments have bought him virtually to this position.

McDowell talks of the outside gaining entry to the moral outlook by having certain natural desires that lead him to act similarly to the moral man. This outsider is clearly someone whose judgements are all hypothetical imperatives. Yet it is not clear to me whether McDowell's outsider is a theoretical construct, or whether he believes such outsiders are actual members of society. So it is not obvious whether McDowell would consider all or some people to use categorical imperatives in their moral discourse.

For Foot's view I return to what she describes as the central supposition

that moral considerations necessarily give reasons for acting to any man.

As noted, two possible interpretations of the notion of categorical imperative may be drawn from this. Either it is to be taken that a categorical imperative is a statement that itself gives reason for acting, and Foot herself agrees that it is implausible that all (should) statements give reasons for acting, or it is to be taken that a categorical imperative is a statement expressing the existence of reasons for acting. Foot comments on the supposition above,

the difficulty is, of course, to defend this proposition which is more often repeated than explained.

op. cit. p.309

Her conviction that the proposition cannot be defended leads her to conclude that there are no categorical imperatives. This would be an acceptable conclusion on the first of the two interpretations of her notion of categorical imperative mentioned above. However, on the more plausible interpretation argued for in the previous chapter, her view that moral considerations do not necessarily give reasons for acting to any man, leads naturally to the conclusion that categorical imperatives are false judgements.

Foot does recognise a distinction between judgements which may be regarded as hypothetical imperatives and statements which purport to express categorical imperatives. Someone adopting this position of Foot's will, she writes,

recognise in the statement that one ought to care about these things (liberty, justice etc.) a correct application of the non-hypothetical moral "ought" (NHU) by which society is apt to voice its demands.

op. cit. p.315

But these NHUs Foot treats non-cognitively:

Suppose that what we take for a puzzling thought were really no thought at all but only the reflection of our feelings about morality?

I claim that Foot's treatment of NHUs is non-cognitive because she attributes to them an emotive meaning whilst regarding them as having no truth-value. She differs from non-cognitivists in general in that she would certainly wish to provide a cognitivist analysis of hypothetical imperatives in contrast to most prescriptivists and emotivists who tend to treat categorical and hypothetical imperatives together and regard them as essentially non-information-providing. Foot also differs from emotivists in regarding such non-cognitive utterances as serving no useful function, whereas most emotivists regard these utterances in virtue of their capacity to express and engender moral feelings, as constituting the essence of moral discourse.

According to Foot's emotivist sketch of the NHU, we feel a binding force in association with moral judgements which derives from the strictness of our moral education. She claims that one can feel as if one has to do what is morally required without believing oneself to be under physical or psychological compulsion. Her failure to find any alternative form of compulsion then led her to reject belief in compulsion as a basis for moral judgement. Foot compares moral feelings to the feeling of falling, which she correctly claims can be experienced without actually believing oneself to be falling. By this analogy, it looks as though she is suggesting that people can feel compelled without actually believing themselves to be. But those who could sincerely claim to believe themselves to be compelled, and I take them to be the majority, must, according to the argument of 29 be accredited with that belief.

Another difficulty facing an analysis of moral judgements that dispenses with moral beliefs, arises in accounting for the differences in a person's feelings that accompany what that person thinks is a change in his beliefs. Feelings of compulsion may rapidly drop away once a person no longer thinks he believes himself to be compelled (e.g. by the standards of etiquette). Residual feelings may often remain after a person thinks he has abandoned a belief, but it is difficult to explain what may sometimes be a dramatic change in feelings without reference to a change in beliefs. Feelings may also strengthen or even originate in adulthood long after the principle psychological factors involved in inculcating feelings have had their effect.

Thus I think Foot is mistaken in supposing that certain judgements that purport to be expressions of moral beliefs are in fact merely the expressions of feelings. In asserting that moral judgements are hypothetical imperatives, Foot often appears to hold that all moral judgements taking the standard linguistic forms for practical judgements are hypothetical imperatives. I have suggested that she is more plausibly interpreted as claiming that certain utterances involving the NHU which might normally be construed as categorical imperatives, are not proper judgements at all, and that all authentic moral judgements are hypothetical imperatives.

It is difficult to say with any certainty what answers should be attributed to existentialist writers on this question. It is likely that they would regard those involved in inauthentic existence as using categorical imperatives, but in building an authentic existence it often appears as though intrinsic reasons reappear for actions having such characteristics as honesty, courage, and faithfulness to one's chosen set of principles. If an initial act of commitment is taken as providing reasons for

action, then it is most plausibly that act itself which should be regarded as the reason, as in the case of a promise, rather than any antecedent interest that that act of commitment might have been founded upon. On such a view, reasons are independent of desires and hence intrinsic. On the other hand the act of commitment might be thought of as initiating an interest which might cease at any time in the future, in which case reasons for actions would be regarded as non-intrinsic.

Thinkers based in Taoist and Hinayana Buddhist traditions should be regarded as non-users of categorical imperatives and hence will be likely to regard some but not all people as non-users of categorical imperatives. An example of a modern philosopher influenced by oriental thought who can be regarded as a non-categorical imperative user is Raymond Smullyan, as may be seen in his book 'The Tao is Silent', Harper Row, 1977. (see especially p.70, Taoism Versus Morality).

4.4. Kohlberg's Experimental Enquiry

At the beginning of this chapter, I expressed the hope that empirical investigation might throw light on the question of the extent of the use of categorical imperatives in moral judgement. The best known psychological work on moral judgement is that of Piaget and Kohlberg. Piaget's work concentrated mainly on the moral judgement of young children, whereas that of Kohlberg extends across all age groups, social classes, and a variety of cultural groups.

Kohlberg's central concern is to show that age development in moral judgement parallels development of non-moral cognitive faculties, and that it proceeds through an invariant sequence of six stages, each of which integrates and replaces, rather than adds to, the preceding stage. Kohlberg's data is based on a system for scoring any moral judgement in any context, but he uses it primarily in analysing responses to ten hypothetical moral dilemmas that were presented to the subjects.

Unfortunately, none of the six stages can be taken as one in which the categorical imperative is not used at all. Stages four, five and six involve explicit recourse to the concept of duty, rules and principles and hence contain only responses in the categorical imperative mode. Stage one is based on obedience and punishment. Stage two is summarised as follows:

Naively egoistic orientation. Right action is that instrumentally satisfying the self's needs and occasionally others. Awareness of relativism of value to each actor's needs and perspective. Naive egalitarianism and orientation to exchange and reciprocity.

'The Cognitive-Developmental Approach to Socialisation' in Goslin (ed.) Handbook of Socialisation, Theory & Research, Rand McNally, 1969.

Stage 3 is summarised:

Good-boy orientation. Orientation to approval and to pleasing and helping others - Conformity to stereo-typical images of majority or natural role behaviour, and judgement by intentions.

Stages two and three, therefore, seem the likeliest groups into which a moral judgement using the hypothetical imperative could fall. However, it is probable that judgements viewed as exhibiting 'naive egalitarianism' in stage two, and 'judgement by intentions' in stage three would be categorical imperatives. This mixture of hypothetical and categorical imperative judgements in these stages is apparent from the example offered of one of the moral dilemmas with a positive and negative 'standardised' response (invented by Kohlberg) appropriate to each stage.

'In Europe, a woman was near death from cancer. One drug might save her, a form of radium that a druggist in the same town had recently discovered. The druggist was charging \$2,000, ten times what the drug cost him to make. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about half of what it cost. He told the druggist that his wife was dying and asked him to sell it cheaper or let him pay later. But the druggist said, "No". The husband got desperate and broke into the man's store to steal the drug for his wife. Should the husband have done that? Why?

Op. Cit. p.379

It is clear from the way that the question is phrased that a categorical imperative answer is being invited. There is no alternative means available for the man to satisfy his desire to save his wife, as the story is presented, so the hypothetical imperative answer is obvious. Nevertheless, if a non-categorical imperative user were to be presented with such a question, he would give the obvious affirmative answer, basing his reasons on Heinz's love for his wife, and the action taken being his only hope of saving her. Two of the standardised stage 2 responses given are:

'It's all right to steal the drug because she needs it and he wants her to live. It isn't that he wants to steal, but it's the way he has to use to get the drug to save her'

Op. Cit. p.379

'If you do happen to get caught you could give the drug back and you wouldn't get much of a jail sentence. It wouldn't bother you much to serve a little jail term if you have your wife when you get out.'

Op. Cit. p.381

Again, the way the first of these standardised answers is phrased by way of the 'It's all right' renders it unquestionably a categorical imperative, though if this were replaced by something like 'He should', this and the second response could be either hypothetical or categorical imperatives. As Kohlberg also supplies negative standardised responses to the question at all stages, including two and three, it is clear that he assumes that there will be categorical imperative answers.

Responses to moral dilemmas of this sort could be used in investigating empirically the extent to which moral judgements are made with hypothetical imperatives. It would be important to include cases in which verdicts were likely to differ in accordance with choice of hypothetical and categorical imperative, as well as cases in which the hypothetical imperative judgement required careful means/end reasoning so that it would be clear when an answer neglected non-hypothetical considerations.

The cross-cultural survey showed a preference for stage three and four judgements by the age of sixteen in middle class urban boys in Taiwan and Mexico, and for stage 1, 2 and 3 judgements in isolated village boys at that age in Turkey and Yucatan. Unfortunately, the published results indicate for each population only the percentage of judgements at each stage, so that if the entire population used both stage 3 and stage 4 judgements with equal frequency, this would appear the same as the case in which half the population uses only stage 3 judgements, and the other half only stage 4. Thus it is impossible to tell from Kohlberg's presentation of his data, what percentage of the populations use only stage two or stage three judgements, which would give some indication of exclusive use of hypothetical imperatives. Even so, the high percentage of stage 2 and stage 3 responses in some populations could be taken as slight evidence for the viability of a morality based on hypothetical imperatives, and its commonness in certain cultures. It could be expected especially in oriental cultures, such as those

influenced by Hinayana Buddhism, though unfortunately, none of these feature in Kohlberg's experiments.

Kohlberg does present some information about individuals whose responses fall with regularity into only one or two stages. But these examples are picked out as exceptions. He analyses some statements made by Adolf Eichmann, showing that they fall predominantly into stages 1 and 2, and he also talks of 'the con-man morality' as a crystallisation at stage 2. However, these examples need not be taken as indicating that all characters whose judgements would predominantly be classified as stage 2 need have hatreds for large sections of humanity, or no concern for other people. The only case recorded of a 'regression' to a lower stage was found amongst twenty percent of middle-class boys at the end of high school to mid-college, but they were all found to have returned at least to their previous stage by their late twenties.

Though Kohlberg argues that development is, with this one exception, from lower to higher of his six stages, it seems likely that someone growing up in a culture in which categorical imperative moral judgements are prevalent would first adopt such a form of judgement before rejecting it. Though Kohlberg's investigations show no sign of permanent shifts to a lower stage, his analyses were directed at different issues, and so need not be taken as counting against the view that some members of certain cultures use only hypothetical imperatives.

Against the view that there could be cultures in which categorical imperatives are scarce or non-existent, it might be argued that children must pass through the stage one morality based on authority and obedience, and hence will first come into contact with words like "should" in connection with the demands of authority, usually the parents, to perform actions that are clearly at variance with the desires they recognise as their own. Yet unless the child can distinguish between 'You should' and

'I want you to' it cannot be said of him yet that he understands and makes judgements that are categorical imperatives. Thus it has not been shown that no society could flourish with a system of moral education that dispensed with categorical imperatives. On the other hand, no positive evidence has been found that there exists or has existed such a society.

In summary, I hope I have shown that empirical evidence could establish the frequency within a community, of persons using only hypothetical imperatives in making their moral judgements; or the existence of a community in which categorical imperatives are not used at all. Modern philosophical scepticism about categorical imperatives suggests that a morality without them is a genuine practical possibility. In the following chapter I shall discuss the differences between a morality based on hypothetical imperatives and one based on categorical imperatives, and some arguments concerned with the justification of categorical imperatives.

CHAPTER V. THE DIFFERENCES BETWEEN USE AND NON-USE
OF CATEGORICAL IMPERATIVES

5.1. Plan of Chapter

An attempt will now be made to look at the differences in behaviour between those who use categorical imperatives and those who do not. The differences in behaviour will be differences in language use, action, and emotion.

In the introduction the thought was raised that the difference between moralist and amoralist, which I equated with a difference between the user and non-user of categorical imperatives, might turn out to be merely a difference in language use. It has emerged that the difference will not be easy to spot since both moralist and amoralist make use of judgements with and without conditional clauses. What determines whether a practical judgement is a hypothetical or a categorical imperative is a secondary judgement that the person making the judgement could make if requested. There are thus not two distinct linguistic forms concerning which the question of the existence of simple translation rules from one into the other could be raised. Accepting that there are sometimes practical difficulties in distinguishing categorical from hypothetical imperatives, the differences in language use are at least easy to chart: In situations in which someone uses a categorical imperative, the non-user will either refrain from making a practical judgement altogether, or he will make the linguistically equivalent hypothetical judgement, or he will make an alternative hypothetical judgement.

The approach this suggests that will be adopted is to look at the situations in which a person may make a categorical imperative judgement, dividing these up according as the judgement is made in the first, second or third person, present or past tense, and then examine the linguistic alternatives and consequent differences in action and emotion for the non-user. As there will be some cases in which a non-

believer in intrinsic reasons will nonetheless use a categorical imperative, though not believing that it can be supported by an intrinsic reason, I shall not use the terms 'user' and 'non-user of categorical imperatives'. And in view of the similar clumsiness of the terms 'believer' and 'non-believer of intrinsic reasons' I shall revert to using the terms 'moralist' and 'amoralist', apart from in the short section on prudential judgements. These terms should be understood in the sense indicated in Chapter I. The amoralist need not be an egoist as will emerge in the following paragraphs.

It may be thought that many of the moral decisions facing the moralist just would not arise for an amoralist. However, this need not be the case. An amoralist may have a great stock of interests which would be considered moral by the moralist, and a moralist may be insensitive and largely oblivious of the moral considerations that moralists usually see as providing intrinsic reasons for action. Camus' *Outsider* provides an example of this. He shows us that he accepts the institution of categorical imperative making and occasionally acts upon it, as exemplified in his visiting his mother although he regarded it as 'a fag.' Another example in which he uses a categorical imperative is his wondering whether he ought to smoke whilst sitting beside his mother's coffin. But he is largely unaware of the desires and feelings of others and neither sees these as intrinsic reasons nor incorporates them into his own projects.

A person making first person moral practical judgements by means of hypothetical imperatives will do so, according to Foot, only in so far as he accepts certain moral ends such as justice and liberty. And his second and third person judgements will be conditional upon the acceptance in the prospective agent of certain moral ends. The range

of such hypothetical imperative judgements can be extended to include judgements conditional upon the desire that the agent's actions have a certain moral character. To illustrate this extension, the judgement that a person should tell the truth on some particular occasions may be construed hypothetically either as conditional upon that person's desire to create some of the benefits that honest dealings bring to men, or as conditional upon that person's desire to act honestly. What is meant here by having a moral end such as justice or honesty is desiring that people reap the rewards of just or honest actions. In the occasional case in which the just act or the honest act would be expected to bring less benefit overall, when a person judges hypothetically that he ought to perform such an act, it would not be said of him that he accepts justice and honesty as ends. It could, however, be said of him that he desires that his actions possess the virtue of justice or honesty.

At this point it might appear as though the moralist and amoralist could sometimes be indistinguishable. There will be a difference though in the way the virtue is grounded. The moralist will have cultivated the virtue through disciplining himself to perform virtuous acts because of a belief about the world, though the early stages of this cultivation would have been controlled by others especially his parents. The amoralist will generally have come to possess the virtue through a drifting of his preferences. The person who develops virtuous dispositions like the moralist and later comes to reject his belief in intrinsic reasons whilst retaining his preference for virtuous action must be designated an amoralist though admittedly his actions may match those of the moralist completely. Differences between these two types of virtuous agent may still be expected however in judgement of the actions of others, and in the experiencing of emotion. Though an amoralist may act in the same way that a conventionally virtuous man would act, it is to be stressed that such a resemblance would be exceptional.

Foot recently gave as a non-moral example of a judgement she would have regarded as categorical, 'I ought to pick these things up from the floor because they don't belong on the floor'. As a result of her domestic training she had come to believe in the existence of an intrinsic reason for tidiness and keeping everything in 'its place'. After abandoning that belief she still retained a preference for tidy floors, but her reason for picking things up from the floor should now be regarded as non-intrinsic.

5.2. First Person Moral Judgements.

Looking now at moral first person past tense judgements, 'I ought to have ϕ 'd', four cases can be distinguished depending on whether the agent did or did not ϕ , and on whether his judgement prior to the time at which the act was to be performed was that he ought to ϕ , or that he ought not to ϕ .

In the case in which X ~~did~~ ϕ , and his judgement prior to acting had been that he ought to ϕ , it is unlikely that there will be any systematic difference in the tendency a person has to making such judgements corresponding to a difference between use and non-use of categorical imperatives. The feelings accompanying the making of a categorical imperative of this type are often described as self-righteous and contrasted with those attending the corresponding hypothetical imperative judgement. In describing the emotional difference between the moralist and the amoralist in such circumstances as a difference between self-righteousness and self-congratulation, I have merely indicated the change in cognitive component. Unfortunately, there are no terms available in ordinary language with which to isolate the cognitive from the affective - a fact which reflects our epistemological difficulties in making the distinction rather than a lack of interest in such a distinction.

To give a fuller account of the emotional differences, a comparison of bodily sensations would be necessary. This could be attempted by way of a statistical comparison of such physiological gauges of emotion as heartbeat, respiration, brain rhythms, galvanic skin response etc. (These remarks apply throughout this chapter to redescriptions of such emotions as guilt, indignation and resentment in which nothing more than the change in cognitive element can be expressed.)

The second and third cases in which a judgement of the form 'X ought to have ϕ 'd' is made are those in which X acted incontinently. In one of these, X's present judgement agrees with that which he made prior to his acting incontinently. In the other, his judgement differs and he is looking back on his incontinence as a piece of good fortune. In the fourth case, X's present judgement is different from the past judgement which he acted upon. This may have been due to his changing his mind just before acting, or to his having made the wrong decision altogether.

Beginning with incontinence, it is interesting to note that not everyone would claim to act in a way that meets what is often termed a pretheoretical description of the phenomenon. Wiggins for example writes:

Almost anyone not under the unfluence of theory will say that, when a person is weak-willed, he intentionally chooses that which he knows or believes to be the worse course of action when he could choose the better course.

(Weakness of Will, Commensurability,
and the Objects of Deliberation and
Desire in P.A.S. 1979, P.1.)

In a recent seminar in which Wiggins was discussing this topic, there was a significant interruption of objections, not to his ensuing analysis, but to the idea that the given

description of weak will was of general applicability. About half those present felt that in situations of weak will they would consider themselves for a time at least to have changed their mind as to what was best, as well as their intention. Still I shall preserve the pre-theoretical description by restricting the terms 'incontinence' and 'weakness of will' to cases where the agent changes his intention but not his judgement as to what was best.

A puzzle over incontinence can now be seen to parallel closely the problem of Chapter IV concerning categorical imperatives. Does the pre-theoretical description of incontinence fit behaviour in which all people, some people, or no people engage? As with the corresponding problem as to whether all people, some people, or no people use categorical imperatives, the option of 'some people' has not been popular amongst theorists. As with the categorical imperative case, the question is simply answered in favour of this option if the agent's own testimony as to whether his behaviour ever fits the description of incontinence is taken as sufficient. And similarly there are great problems in finding a set of criteria under which such testimony can be overruled.

In fact it turns out that the division between people that sometimes intentionally act contrary to their best judgement and those that never do, can be shown to be related to the division between those believing in the existence of intrinsic reasons and those who do not. I have claimed that a person making an overall hypothetical imperative judgement that he ought to ϕ will of necessity intend to ϕ , since the judgement is his balancing of his present desires. This would suggest that his actions could not possibly satisfy the pretheoretical description of incontinence, and hence that he would be someone who would

claim to have changed his mind, i.e. his judgement as to what was best, at the last moment. Thus the amoralist cannot act in accordance with the pretheoretical description of incontinence.

When a person judges overall categorically that he ought to ϕ this judgement may automatically constitute an intention to ϕ or it may not. I have suggested that the silencing model best accommodates an explanation of the behaviour of the man for whom the overall categorical imperative judgement automatically constitutes an intention to ϕ . He will be a possessor of the virtue of temperance, prone neither to incontinence nor to changing his judgement at the last moment. The behaviour of the continent and the incontinent man is best explained on the outweighing model, in which it is possible for the agent's categorical imperative judgement as to what is best, to diverge from his hypothetical imperative judgement from which his intention derives. It is the agent's hypothetical imperative judgement that changes at the point of incontinence and this need not be accompanied by a change in his categorical imperative judgement. Thus the moralist may sometimes act in accordance with the pretheoretical description of incontinence, and may sometimes change his mind at the last moment as to what is the best course of action.

I shall now make some suggestions as to the effect upon action and emotion that such a difference in conceptualising may have.

The moralist will sometimes act incontinently, and feel some guilt or shame at having failed to do what he judged was best. Such emotion will be stronger than that following his decision to change his judgement which may be a sort of shame that his desires were so strong as to outweigh other considerations that would have led to a different action.

The main difference between the moralist who changes his judgement as to what is best and the amoralist who never acts incontinently is that the former person is likelier to change his mind again after acting. The latter person can only change his judgement afterwards if he thinks that better means were available to furthering the interests that he had at the time of acting. If he sees no such mistake he may still feel some sadness, shame or embarrassment at the desires or interests that were his at the time.

Immediately prior to acting, the moralist is less likely to change his judgement as to what is best, but more likely to change his intention, i.e. to act incontinently. This is because the amoralist will always be balancing interests and desires that are his own, and these are likely to be more stable than the moralist's desire which is consequential upon belief in the existence of an intrinsic reason that enters his intention-forming judgement.

I shall now discuss the case of past tense judgements in which a person does not perform the act he judges he ought to have, through having failed to reach the right decision. The only retrospective reaction likely from the amoralist is shame or annoyance at his own inability to calculate or foresee consequences, with perhaps a noting of the experience as a guide to future judgement. The moralist on the other hand is judging that something should have been done which would not have been conducive to the satisfying of any of his desires at that time. In many situations of this type, the amoralist would not make a judgement at all. The situations in which he makes a judgement that he ought to have ϕ 'd when he did not, are invariably situations in which the moralist would also make a judgement, either the corresponding categorical imperative or another hypothetical imperative. Thus the moralist will more frequently find himself feeling displeasure over a

poor judgement made in the past than a amoralist.

Considering together the four cases of past tense judgements 'I ought to have ϕ 'd', the moralist is likely, in his behaviour towards others, to be more sensitive to his own past success in acting in accordance with his present judgement, than the amoralist. He will, for example, tend to accept the reproofs of others and feel less able to condemn others when he considers them to be behaving wrongly, if he himself frequently fails to act in accordance with what he judges to be best. The amoralist will have a different manner of dealing with criticism and influencing the actions of others. His success in influencing action may well depend on his own past record of action, but his efforts and keenness to do so will not. Whether or not the agent is a moralist should not affect his expectation of response from others, as it will be the conceptualising of his audience that will be relevant to the response. The moralist might be more likely to modify his own future conduct as a result of his retrospective judgement of his own actions on account of the greater number of cases in which he can make a present judgement that he ought to have ϕ 'd when in fact he did not ϕ .

Concerning first person present tense judgements, there will be two types of decision that the moralist makes that are qualitatively different from those made by the amoralist. These will be the conflicts between two intrinsic reasons for incompatible actions, and the conflicts between intrinsic and non-intrinsic reasons. In the moral conflict situation, the person who conceptualises in terms of hypothetical imperatives will not feel guilty or blameworthiness over his leaving one of the acts unperformed. He may, however, feel horrified at the unfortunate circumstances that forced the decision on him, and feel compassion for the victims, though this would be no different from the compassion he might feel for the victims if the injury had been inflicted by someone else. As well as being qualitatively different, the emotion experienced by acting by the moralist is likely to be stronger than that of the amoralist. Whilst deliberating

and prior to acting, the emotional state of the agent is unlikely to depend on the way he conceptualises the decision. The thought of leaving either of the actions unperformed may be just as unpleasant, and the decision just as excruciating, for the person who formulates the judgement as a hypothetical imperative.

I have discussed the differences between the moralist's experience of conflicts between intrinsic and non-intrinsic reasons and the experience of an amoralist who encounters a similar situation in connection with past tense judgements. An interesting particular case of such conflicting reasons which has not been discussed is that of supererogation.

Acts of supererogation are those regarded as not required by duty, yet of supreme moral worth. There are two main types into which such acts fall. One type consists of acts that would be considered as both of value, and required by duty, were it not for some considerable inconvenience or danger to the prospective agent in the situation. The moral value attaching to the action is considered to be greatly enhanced as a result of inconvenience or danger, and in situations in which there is danger, the act would be regarded as possessing to a high degree the virtue of courage. The other type of supererogatory act consists of acts in which no special inconvenience or danger is present, such as may be found in some examples of running a non-profit-making organisation or accepting a position of leadership.

The principal difficulties facing an analysis of supererogation are in explaining the divide between evaluation and reason-giving, and in explaining the enhancing of the value attaching to an action by the presence of danger. Beginning by considering the supererogatory act motivated by a categorical imperative user, it will not be too difficult to see how the agent came to judge that there was overall

intrinsic reason for him to act, commensurate with the favourable light in which he saw the act. What will be more difficult to comprehend is how others would regard the act as highly praiseworthy, yet not make the third person judgement that there was intrinsic reason for the agent to have acted the way he did. The same problem arises in accounting for the making of a first person practical judgement that there is no overall intrinsic reason to act courageously together with an evaluative judgement that the courageous act would be better than the non-courageous one, in situations in which the person making the judgement decides not to adopt the courageous choice of action. Perhaps the failure of entailment from the evaluative statement to the overall reason statement can remain as an unexplained feature of an ordinary moral judgement. More problematic is the apparent compatibility of 'There exists overall intrinsic reason for X's ϕ -ing' with 'There does not exist overall intrinsic reason for X's ϕ -ing' when the first judgement is made by X, and the second by someone else. One of these judgements would have to be considered false if 'intrinsic reason' is to remain an objective notion. The most acceptable of these options would be to claim that the saint or hero's judgement that he had overall intrinsic reason to act was false. This need not detract from the high moral worth with which the action is regarded.

Sometimes, however, the agent will see the supererogative act as not required of him, just as it would not be required of anyone else. If he then decided to go ahead with such action, this must be explained in terms of an interest (e.g. in some charity or form of social work, or, in the case of courageous action, in the welfare of an individual or group of individuals), in which case his judgement will be a

hypothetical imperative. Such a hypothetical imperative judgement to do the supererogatory act may be made by someone who normally uses categorical imperatives in his moral judgment, i.e. the moralist, as well as by the amoralist. Thus supererogatory acts are not restricted to moralists, and it will often be easier to understand the decision prior to acting as a hypothetical imperative.

5.3. Second Person Moral Judgements.

Turning to moral judgements given in the second person, a moralist may be advising in the genuine belief that the prospective agent may take the advice, and that there is overall intrinsic reason for him to perform the recommended act. Or he may be stating a belief in the past tense that cannot be utilised in any future actions, but that is nevertheless felt to be worth pointing out for the benefit of the might-have-been agent. People have an interest in the past that is not totally dependent on their interest in the future. Then there is the hackle-raising use, where the judgement is offered not in any belief that the advice will be adopted, but as an expression of indignation intended merely to outrage the recipient. Finally, there is the use in which a person, Y, aims to guide another, X, purely for his, Y's, own ends. Mackie claims that the speaker's attitude can help constitute the reason for the proposed action, and he labels this type of reason egocentric, in contrast with intrinsic and hypothetically imperative reasons (op. cit. P.25). As I have defined intrinsic and non intrinsic reasons, they are exhaustive, and parallel Mackie's intrinsic and hypothetically imperative reasons. It is true that when Y asserts to X, 'You ought to ϕ ', Y's attitudes, principles and interests may be partly accessible to X and may contribute to the existence of a reason for X's ϕ -ing. But they would contribute by means

of an intrinsic reason deriving from Y's authoritative position over X, or from X's desire to conform to Y's demands based on his liking of Y or his fear of him.

Thus if reasons deriving from a speaker's attitude play any part in validating 'You ought to ϕ ', they do so as intrinsic or non-intrinsic reasons. The term 'egocentric reasons' may be adopted so long as it is realised that these do not constitute a third category of reason.

Of the uses just sketched that are available to the moralist, those connected with advice are invariably cast in categorical imperative form. This is because the believer in intrinsic reasons will see the aspect of some moral consideration that is conducive to the satisfying of the agent's own desires as irrelevant, except as bait where necessary to lure the agent into doing the right thing. The amoralist will always address his advice to some moral interest of the prospective agent's and hence the scope of his advice will be restricted in comparison with that available to the moralist. This advice will on average have a greater chance of being heeded than that of the moralist, since it addresses the interests of the prospective agent.

The purely informative non-advisory past tense imperative will be unavailable to the amoralist as it indicates something the moralist takes to be a fact worth pointing out which is of interest irrespective of one's interest in the future. To the amoralist there are no such facts. The antagonistic or hackle-raising use in both present and past tense relies for its effect on a belief that the statement could be true. Thus the amoralist may use a categorical imperative knowing it to be invalid, purely because of the effect he expects it to have on its recipient, if the recipient is himself a moralist. An amoralist is less likely to want to make use of such a weapon, since his motive for so doing cannot be indignation.

Finally, the judgement 'You ought to ϕ ' may be used as a means to satisfy the speaker's own desire that the prospective agent ϕ 's. Both hypothetical and categorical imperatives may be used to this end, and will be available to both moralist and amoralist. As in the antagonistic case, the speaker need not believe the categorical imperative judgment but he may nevertheless issue it in the hope that the recipient's belief in intrinsic reasons will contribute to the likelihood of his ϕ -ing.

5.4. Third Person Moral Judgements

Ought judgements are also made in the third person where there is no chance that the judgement can have any effect on the action under discussion. Take as examples: The 1980 Olympics should not be held in Moscow. The Shah should return to Iran. Or the more mundane: Mrs Lewis should not travel on the train today without paying. Similar examples occur in the past tense: The Americans should not have bombed Hiroshima. The Nazis should not have set up concentration camps. Or the personal and fictional: Hamlet should not have killed Polonius.

Some of these examples can be construed as cases of which generalised forms can influence the speaker in deliberation over his own actions, or influence the actions of those to whom he asserts the judgement. The judgement that Mrs Lewis should not travel on the train without paying is obviously a case of this. A person making such a third person judgement, would be expected also to make the general judgement that people should not travel on the train without paying. Hence it might be thought that the purpose of making such a third person judgement is to influence action. It seems highly implausible that all third person judgements should be made for this purpose. Admittedly, part of the purpose of making judgements in the third person, it might be claimed, is to prepare oneself and others for whatever slight action they can muster in the hope of influencing

the course of some similar future action. E.g. perhaps organise some sort of protest action if it looks as though the U.S. is about to use its nuclear weapons again. But people continue to make third person judgements with categorical imperatives even though they feel that the making of such judgements can have no influence over events. It is, for the moralist, an essential part of the interest he takes in part of the external world that extends beyond his control.

How, the moralist will wonder, can the amoralist be so narrow-minded as to take an interest only in the events which he can influence? And how can the amoralist even influence the choices of others indirectly unless he uses third person categorical imperative judgements? To the second of these questions it may be replied that the amoralist may use a categorical imperative in the second person case if he wishes to influence the choices of another person if he thinks his audience's belief in intrinsic reasons will make his utterance effective in guiding action. Although the amoralist's chief means of influencing the actions of others will be by suggesting consequences of a prospective action in the hope that the prospect of those consequences' obtaining will itself encourage or deter the agent from performing the action.

To the first of these questions it may be replied that the amoralist certainly has at his disposal the conceptual means to assess various possible events over which he has no control. He may hope that Reagan becomes president, or want Reagan to become president, just as he may hope that Ali loses his last fight, or want Ali to lose his last fight. He may also make judgements of objective value concerning such events, from which his subjective preference may depart. It will be discussed in the next Chapter whether there is an entailment from 'It would be good if X ϕ 's' to 'There is intrinsic reason for X's ϕ -ing'.

Comparing the moralist and amoralist from the emotional point of view, the third person categorical imperative judgements of the moralist will often be accompanied by indignation. And this will not be present when the amoralist makes his hypothetical imperative judgements, or feigns categorical imperative judgement. But it must be repeated that the term 'indignant' is used to refer to an emotion supported by a particular cognitive element, so that in discussing the effect this cognitive element has on emotion nothing has been gained by saying that the difference concerns the presence or absence of indignation.

So far the use of categorical imperatives and hypothetical imperatives has been considered primarily from the point of view of the speaker, or person making the judgement. In cases in which the judgement is asserted to an audience, the way in which that judgement is interpreted will correspond in the main to the way a moralist would classify it if he were himself to make that same judgement. Thus moralists will interpret practical judgements as categorical imperatives whenever plausible just as they would tend to use categorical imperatives in preference to hypothetical imperatives whenever plausible. (An exception to this has been noted in the case of supererogation). An amoralist too is likely to interpret a judgement as a categorical imperative whenever plausible, having grown up in a society in which this use is widespread. The only exception to this would arise when the audience knows that the speaker is an amoralist.

A final point of comparison between the hypothetical and the categorical imperative in situations in which either could be asserted concerns the way in which they are supported. The amoralist will aim to spell out the conditional indicating the interests of the agent to which the action

would be ancillary, and the causal route between action and furthering of the agent's interest. The moralist will aim to enunciate the fact that is to constitute the intrinsic reason for action, and indicate why the fact does constitute intrinsic reason for action. The task of elaboration is more clear cut for the amoralist but this will not deter the moralist from making his judgements when he sees them as correct and of value.

To many moralists, the belief in intrinsic reasons is an essential part of what galvanises them into action, and what makes life meaningful, and without it they would be at a loss. Something external and objective is felt necessary as a justification of action. As well as a need for intrinsic reasons for actions, there is also a need for intrinsic reasons against actions, especially those that have already been performed. Moralists often have greatest need for intrinsic reasons as a means to justify their feelings of hostility towards the self and to justify their need for an enemy.

5.5. Prudential Judgements

I shall now briefly consider the effect upon prudential judgement of the use of categorical imperatives. As has been remarked, there are difficulties in distinguishing categorical from hypothetical prudential judgements, especially in the first person. What was said concerning incontinence in the moral case will apply also to weakness of will in the prudential case. The existence of prudential cases of incontinence that conform to the pretheoretical description can in fact be taken as an argument for the categorical imperative status of certain prudential judgements.

Prudential categorical imperatives are perhaps more readily distinguishable from hypothetical imperatives in the second person, and I shall consider now some uses of prudential categorical imperatives.

One common use of past tense judgements is to advise for the future. E.g. You should have asked someone else to do the proof reading for you. Clearly this can have some effect on future behaviour. But not all examples are of this kind. The judgement 'You ought to have changed subjects at Cambridge' may be offered as the statement of a belief which can no longer have any bearing on the person's future judgements. There is also the case of a statement made purely to raise hackles because it is known that the person to whom it is addressed will strongly disagree, whether it be over something that is potentially serviceable as advice or not - Such judgements assume a categorical imperative basis if they have any pretensions to being factual, i.e. are not merely taunts. The hypothetical imperative form can only accommodate them if it allows for a vague present desire for a general flourishing.

Present tense prudential judgements can only be advice-offering or hackle-raising, though sometimes it is difficult to tell which. The judgement 'You shouldn't eat white sugar' may be intended as advice directed at the person's concern to eat the right kinds of foods, in which case it would be construed as a hypothetical imperative. Or it may be offered despite knowing that the person is a firm believer that the body will cope with virtually anything it is given as fuel. Such a remark assuming the form of a categorical imperative is essentially antagonistic and can perhaps be seen as the speaker's affirmation of his own set of values - in order to feel confident about a first person practical judgement he needs to believe that the judgement is universalisable. Alternatively, the remark may be made from a sincere interest to prevent any harmful effects from afflicting the recipient. The judgement is then either categorical, or hypothetical based on a desire for future health.

There will be few instances in which users and non-users of categorical imperatives can be distinguished with any certainty with prudential judgements. Some judgements, it has been seen, can only be construed as hypothetical imperatives if some very general present desire for future welfare is assumed. The believer in intrinsic prudential reasons is likelier to be given to making pronouncements to others than the non-believer.

CHAPTER VI. JUSTIFICATION OF THE USE OF CATEGORICAL
IMPERATIVES

6.1. Types of Justification.

In this chapter I shall offer a rather cursory examination of the justification of categorical imperatives. I shall not consider prudential judgements, and I shall not consider any justifications of moral categorical imperatives that aim at justifying general rules from which singular judgements may be derived. This rules out all forms of contractarian argument, along with the arguments of Kant.

The types of justification I shall consider may be conveniently characterised ^{as} /theological, pragmatic and metaphysical.

6.2. Theological Justification.

Obviously it would be inappropriate to discuss the plausibility of Christianity here. I want only to suggest that those acting in accordance with what they take to be a divine command, or God's will, need not be seen as acting from fear of Hell or hope of Heaven. If this were their sole reason for obeying a divine command, then their judgements that various actions ought to be undertaken would be hypothetical imperatives, as Kant claimed. But as most Christians would agree that divine command would give them reason to act even if there were to be no personal rewards, this would render their judgements categorical imperatives according to my definition. There may even be some Christians and other religious people who believe in a deity that issues commands or prohibitions without any rewards or punishments beyond the joys of acting in accordance with the will of the deity and the guilt incurred in disobeying.

The notion of a command which cannot be enforced may appear a little suspect. If the commands and prohibitions of the deity are replaced by preferences for certain actions, the role of the deity becomes more intelligible. Acting in accordance with God's will may be thought of as acting as God would wish, or as acting in a manner pleasing to God, rather than obeying God's command. Now it might be thought that the reason for acting in accordance with God's will is a desire to please God, or a love of God, which would again be a non-intrinsic reason. Again, I think that most Christians would consider that God's will would provide them with reason to act even if they had no love for God or desire to please God. If it now be objected that no sense can be attached to the idea of a reason for acting in accordance with God's will, if it is not based on love or punishment and reward, the Christian reply may be rather similar to the atheist's portrayal of intrinsic reasons: the intrinsic reason is God's will, just as X's need just is an intrinsic reason for Y's action.

A familiar objection to the view that intrinsic reasons may stem from the will of the deity arises in Plato's Euthyphro. Does God will an action because it is good, or is an action good because God wills it? If the latter, then 'good' means 'willed by God' and we discover whether an action is good by discovering whether God wills it, for example, by divine revelation or the reference to holy scriptures. On this view it could turn out that good actions are unpleasant to the self and others. Why then should there be reason for a person to follow the will of the deity? Perhaps as a result of the deity's responsibility for bringing him into existence. There are few, however, who would regard this as a satisfactory solution to the Euthyphro dilemma, for it is generally thought that it cannot just be a piece of luck that good

actions serve human needs. For goodness to derive from God's will, most would require that God necessarily has a concern for human interests. But now it is no longer clear whether 'good' or 'willed by God' is more basic.

If the former direction of dependence is chosen as answer to the Euthyphro dilemma, it might be argued that the will of the deity cannot be taken as the reason for action, since it is the prior property of goodness in an action that determines whether it is willed by the deity. To this it might be responded that if value or goodness is granted to be fixed independently of God's will, then unless goodness turns out to be what every individual always desires, there need not always be reason for someone to bring about value, or perform actions that are good. By willing that good actions are performed, the deity could then fill the gap between an action's being good and there being reason to perform the action, i.e. between 'It would be good if X ϕ 'd and 'There is reason for X to ϕ '. If good actions will provide reasons that are unquestionably based on divine will, if it provides reasons at all, but again there may be qualms about accepting the will of a merely contingently benevolent deity as reason-providing.

It should appear, therefore, that there are solutions to the Euthyphro dilemma on which divine will yields intrinsic reasons, if 'goodness' and 'God's will' are only contingently related. Most Christians would regard divine will as also providing the foundation of intrinsic reasons if 'goodness' and 'God's will' are necessarily related.

6.3. Pragmatic Justifications

By pragmatic justifications of categorical imperatives I mean justifications that seek to establish that, whether or not a metaphysical justification is available, society

could not flourish without such an institution. Such justifications include the "Noble Lies" of those who do not believe in the existence of intrinsic reasons but argue that it is necessary that the prevalence of such a belief be sustained for the coherence of society.

One pragmatic argument for the justification of categorical imperatives is the claim that it is a necessary part of moral education. This could be the claim that no other linguistic means is available of controlling or guiding the actions of children, or that insufficient interest in the welfare of others could develop across a whole society without the institution of categorical imperatives. It was argued in Chapter IV that empirical evidence could establish the existence of a society in which no categorical imperatives were used, or a society in which categorical imperatives were used only as a passing phase in moral development. On the negative side, empirical evidence could only render implausible not disprove the existence of, such a society. I could not offer positive empirical evidence, as there was little experimental work of the right kind available. But I argued that the basic commands that a child has to obey are not categorical imperatives, and I see no reason why other judgements made by the child and to the child need be categorical imperatives.

The view that insufficient benevolent inclinations would develop without categorical imperatives to provide a sane, happy society that people would choose to live in could, this time, be refuted, but not established, by empirical evidence, but again there is no evidence of the right sort available. The factors that lead to the thriving of a society from the viewpoint of all, or the vast majority of, its members, are complex and depend on political, economic, technological, religious, and artistic factors. My claim

is that the use of categorical imperatives is not part of any political, religious or social institution that is necessary for that thriving. In suggesting some differences between moral thought incorporating categorical imperatives and that based on hypothetical imperatives in Chapter V, I hope to have shown that a non-user of categorical imperatives, or more strictly a non-believer in intrinsic reasons, need not be parasitic upon a society in which belief in intrinsic reasons is the norm, and hence that a society of such non-believers is not a practical impossibility.

Another pragmatic argument for the justification of categorical imperatives derives from a consideration of the view that morality has a particular purpose. Mackie expresses this view as follows:

Protagoras, Hobbes, Hume and Warnock are all at least broadly in agreement about the problem morality is needed to solve: limited resources and limited sympathies together generate both competition leading to conflict and an absence of what would be mutually beneficial cooperation.

Op. Cit. p.111

Surprisingly, as he has argued against the existence of intrinsic reasons, Mackie claims that moral sentiments and dispositions, the respect for various obligations, plus more formal moral rules and politico-legal devices for law enforcement and the making of positive law all contribute to the purpose of morality, and may continue to contribute in a society whose members accept the inexistence of intrinsic reason and invent their own moral code.

But can Mackie consistently claim that such a moral code can operate without a belief in intrinsic reasons? When a person abandons a belief in intrinsic reasons, moral sentiments and dispositions are likely to drop-away though

there will be some residual feelings for a time, and maybe a certain stable core will remain. These feelings are likely to acquire a different character as a result of the change in cognitive component, and to lose some of their severity. It is possible that a person may acquire a new set of feelings on launching his chosen moral code, and this might be a fitting description in some cases of the committed existentialist.

In inventing a new moral code, Mackie thinks that some of the old rules will be retained and some new ones added. Just as there are moral agents who believe in intrinsic reasons but who do not base their moral decisions on rules, it is also to be expected that there will be non-believers in intrinsic reasons who see no place for moral rules in their invented morality. On abandoning a belief in intrinsic reasons, a moral agent is likely to retain his preference for or against deliberation in terms of a set of rules. But without intrinsic reasons, what reason will there be for acting in accordance with the newly chosen or newly endorsed set of rules? Mackie suggests that it is the moral feelings that we all have that we can fall back on, (op. cit. p.191) but such feelings will motivate action without the mediation of moral rules, and it is unclear how feelings can be attached to the observing of unanchored rules. Most people's adherence to a set of moral rules seems to rest on an expectation that other people will adhere to a similar moral code. When the whole of society has abandoned belief in intrinsic reasons, expectations that strangers will share one's moral code will, initially at any rate, be reduced. Though Mackie himself recommends adopting a set of moral rules that accord with his three stages of universalisation, he accepts that others will not be bound to do so. A set of unanchored

moral rules over and above the agent's feelings and dispositions could serve as a check on limited sympathies, but it will not be necessary that people adopt rules in order for society to cohere. This is fortunate, for it is difficult to see what reason a person would have for keeping to his chosen rules. Mackie himself suggests (op. cit. p.191) that we may have to abandon the attempt to show why a person should observe his own rules and suggests that one can more easily reconcile oneself to this by reflecting on the doubts about the rationality of prudence.

It is thus dispositions and feelings which will provide the foundation for moral behaviour when belief in intrinsic reasons has been abandoned. The division of dispositions, feelings or sentiments into moral and natural, for a believer in intrinsic reasons, will reflect whether or not the feeling is supported by this belief. It would be a mistake to think of these natural feelings as innate, and as including a person's natural sympathies, and to then think of the moral feelings as a necessary adjunct to counteract the deficiency of natural sympathy. Feelings are inevitably moulded by experience, so that it makes little sense to talk of natural sympathy as though it were an innate feeling. Everything thus depends on how a person's feelings and sympathies develop in a society in which moral education is conducted without a cognitive structure based on a belief in intrinsic reasons.

Though a morality based on hypothetical imperatives need not contain an effective set of moral rules, this need not cast doubts on the possibility of a legal system, and enforcement procedure. The inflexibility of the Hobbesian state can be attributed to the lack of basic sympathetic feelings in the (hypothetical) members. Whether or not a legal system and enforcement procedure are stable depends principally on the extent to which the laws and punishments are recognised as of overall benefit to the community. The stabilising

factor resulting from a feeling of moral obligation to obey the law will be relatively insignificant, though it might be argued that a strong sense of moral obligation will limit the need for an elaborate policing system. However, it is unrealistic to suppose that a state could ever dispense with a means of enforcing its laws. A desire for enforcement of laws is a basic essential, and the size, and hence cost, of this device will depend on the scale on which infringements occur. Again this is a matter that cannot be argued without empirical evidence, but I believe that the existence of a moral feeling of obligation to obey the law will make very little difference to the extent of its observance.

Hence I conclude that pragmatic justifications of categorical imperatives cannot carry much weight devoid of relevant evidence. In any case it would seem that metaphysical backing for moral beliefs would be required in order to support a pragmatic justification. The cost of sustaining poorly supported belief in a community is appreciable. (Consider, for example, the cost of upholding a belief in the literal truth of the contents of the bible.)

6.4. Metaphysical Justifications.

I shall now take a brief look at Mackie's arguments against intrinsic reasons that would serve to undermine, rather than justify, the use of categorical imperatives. The argument from queerness is unlikely to have much persuasive power, though by examining the ontological status of intrinsic reasons and concluding that such reasons must be non-linguistically expressible facts, I hope to have shown that the target for the argument from queerness has at least become more conspicuous. The argument from relativity has been described by Mackie as follows:

The argument from relativity has as its premiss the well-known variation in moral codes from one society to another and from one period to another, and also the differences in moral beliefs between groups and classes within a complex community.

The success of such an argument will obviously depend on the items of the moral code that are picked out for comparison between and inside societies. Yet it is precisely this line of reasoning that is adopted by many to argue for the reverse conclusion. The anti-objectivists will seize upon some aspect over which there is widespread divergence of opinion to argue their case. The objectivists will pick some aspect over which there is considerable agreement, or towards which opinion has converged over a period of time, to argue for the existence of some moral feature in the world.

Arguments from relativity (or consensus) usually aim at disputing (or establishing) the objectivity of some value property or evaluative predicate. The anti-objectivist will invariably take a general predicate like 'good' whereas the objectivist will consider more distinctive predicates like 'courageous' to be central. (cf. Platts op. cit. p.2) However, the issue at hand is the existence of intrinsic reasons and the extent to which there is consensus over the practical judgements which imply the existence of such reasons. The argument may proceed either directly to an assessment of the truth of a practical judgement; or in two stages, the first of which would be an assessment of the truth of some evaluative judgement concerning a potential action, and the second would be an assessment of the validity of the implication from the evaluative judgement to the practical judgement. I shall discuss first the second approach because of the light it throws on the relation between evaluative judgements and practical judgements. Evaluative predicates of actions, such as 'good' and 'courageous' are objective if and only if, for all possible actions X, the corresponding evaluative judgements 'X is good' and 'X is courageous' are either true, false, or indeterminate, and this truth-value is fixed by the way the world is constructed.

The possibility that the truth-value may be indeterminate is included to allow for the existence of objective vague predicates such as 'slow'.

If these predicates are objective then two people cannot diverge in their assessment of the corresponding judgements as true, false, or indeterminate, without one of these people being in error.

Assuming that both understand the meaning of the terms 'good' or 'courageous' the error would be explicable in terms of epistemic limitations, either insufficient information to hand, or insufficiently developed recognitional capacities on the part of the speaker.

Thus if a predicate such as 'good' or 'courageous' is objective, a person may be mistaken in judging that a certain action has that property. And if a person believes that such terms are objective, then he will believe that he may be mistaken in making such a judgement and that his judgement may diverge in its positive/negative aspect from his own preference in the matter. If the predicate is not objective, then a person cannot be mistaken in making a judgement involving the predicate provided he knows the meaning of that predicate. And if a person believes the term is not objective, then he will believe that there can be no difference between giving his opinion as to whether an action satisfies that predicate, and judging that the action satisfies that predicate. He will also believe that there can be no divergence between the positive or negative aspect of the judgement, and his own favourable or non-favourable assessment of the act in question.

How then is the objectivity of a predicate to be assessed? One point at which it might be thought the anti-objectivist could attack the objectivist, is the claim that the truth of the judgement 'X is good' is fixed by the way the world is constructed. This may also be seen as the claim

that evaluative predicates denote properties in the world. The anti-objectivist might demand that if this were the case, then the evaluative predicate ought to be reducible to other predicates that clearly do denote properties in the world, such as the predicates of some natural science. This is a modern way of putting part of Moore's argument that goodness cannot be a natural property. However, it has been argued that many predicates such as 'is red', 'is a strawberry', 'is a car', are not reducible to a set of predicates of natural science, yet they are fixed by the way the world is constructed in the sense that no two entities could differ in respect of their satisfaction of such a predicate without differing in respect of their satisfaction of some physico-chemical predicate. And once it is seen that properties like redness may be fixed by the way the world is constructed without 'red' being reducible to physico-chemical predicates, then the irreducibility of 'good', 'courageous' etc. cannot be seen as counting against their objectivity.

I can see no way in which this argument can be reversed in favour of an objectivist view of evaluative predicates by arguing that such predicates are fixed by physico-chemical predicates in the manner described above. Even if someone picked out a precise set of physico-chemical predicates that he claimed were sufficient to fix an evaluative predicate, someone else who disagreed as to the extension of the predicate could produce a different fixing class of physico-chemical predicates and it could be claimed by the anti-objectivist that both are equally valid. The problem as to whether evaluative predicates have a unique extension would be no closer to solution.

How then is it to be determined whether divergence over the truth-value of such judgements as 'X is good'

involves error, and hence whether 'good' is objective? Wiggins has proposed a slightly different account of what it is for a predicate to be objective that promises greater hope of applicability. According to him, a predicate, P, is objective if and only if, there is intersubjective agreement over the truth-value of statements of the form 'X is P' and the best explanation of such consensus is the fact that X is P. This differs from my definition in demanding intersubjective agreement as a necessary condition for objectivity. Whether or not this is too strict cannot be established until it is known how great a consensus must be to count as intersubjective agreement. On the face of it, the problem of determining whether the best explanation of the consensus is the fact that X is P, would appear to be the same as the problem of determining whether divergence of opinion as to whether X is P entails error. The difference, I think, is that Wiggins' account places the burden of proof with the anti-objectivist in providing an alternative explanation of the consensus. He is thus taking intersubjective agreement as not just necessary for, but also as prima facie evidence for, objectivity. The weight of evidence in favour of the objectivity of a particular predicate would depend on the extent of the consensus and the extent of the convergence over time, however these are to be gauged, and the plausibility of alternative explanations of the consensus.

A first step towards the establishment of intrinsic reasons and justification of categorical imperatives would be the establishing of the objectivity of such predicates of actions as 'good', 'honest', and 'courageous' from an examination of consensus, convergence or some other argument. This would establish the existence of facts of the form 'X is good' etc. The next step would be to argue that such facts constitute partial reasons for action, i.e. that 'If X is good then there is reason to perform X' is analytic.

Analyticity claims can ultimately only be settled by consensus amongst the speech-community, in the light of a thorough search for counter-examples. The third step would be to argue that such facts constituting partial intrinsic reasons to perform x may sometimes combine with all relevant non-intrinsic reasons to produce conclusive intrinsic reasons to perform x. It is difficult to know how this would be argued for. When people are educated in using categorical imperatives or thinking of intrinsic reasons for action, they are invariably introduced first to putative cases of overall judgments which presuppose the possibility of the existence of conclusive intrinsic reasons. There would not be much point in categorical imperative judgements that could never constitute overall judgements, i.e. take into account all that is known about the agent's desires. But this cannot count as an argument against the sceptic.

The argument for the existence of intrinsic reasons may be disputed at any of these three stages. The objectivity of 'good' (it need not matter what sort of perfectionist or naturalist theory of the good is adhered to) will be less readily accepted than the objectivity of less general predicates like 'courageous'. But then the entailment from the fact that x is good to the existence of intrinsic reason for performing x will be more readily accepted than the corresponding statement with 'courageous'. Someone accepting a consensus argument for the objectivity of some evaluative predicate but rejecting the step to intrinsic reasons may be amenable to a consensus argument directly to intrinsic reasons. Such a direct argument might also seem preferable in that a wider range of facts may now be considered as candidates for intrinsic reasons, rather than just those facts that attribute some moral property to the prospective action.

A direct argument might begin with some situation in which there is consensus of opinion over a practical judgement. For example, in a situation in which a man is starving at X's doorstep, there might be consensus over the judgement 'X ought to give this man food' or the less colloquial 'There is reason for X to give this man food'. But as before, the arguments that aim at placing the burden of proof with the opposition carry little persuasive power. And the opposition can make several suggestions concerning any cases of consensus over evaluative judgements and practical judgements that might be taken as evidence for objectivity. A principal suggestion might be that consensus centred around false beliefs that were kept in place by rigid theological practices. The view that Christianity provided the rationale for categorical imperatives has been presented forcefully by Anscombe (in 'Modern Moral Philosophy' Philosophy'58) and is also accepted by Foot. Mackie proposes a number of ways in which belief in the objectivity of evaluative predicates and intrinsic reasons may be explained from an anti-objectivist stance (op. cit. pp.42-46). Some of these I would clarify as pragmatic, others bear directly on the problem of explaining the appearance of consensus.

6.5. Conclusion.

In attempting to present a minimally vague characterisation of what it is to believe in intrinsic reasons I hope to have facilitated discussion of the existence of intrinsic reasons. The metaphysical arguments that I have considered here have been inconclusive, but such arguments would in any case be expected to carry little sway in changing a person's belief in the existence or inexistence of intrinsic reasons. More effective in changing such beliefs will be reasoning based on personal experience or empirical

evidence. People need to come into contact with individuals and social groups in which categorical imperatives are not used in order for them to become convinced that the world need not be a hideous place without categorical imperatives. There will be great psychological difficulties in practice in changing one's belief in the existence or inexistence of intrinsic reasons. Such beliefs are moulded by early education and are difficult to reverse. Even if someone has reached a stage at which he has no strong preference for or against a morality based on categorical imperatives, i.e. has no need to cling to one or the other position, it would be difficult for him to actually adopt the opposite belief, just as in the case of theistic belief, it is difficult for an agnostic to know how to make a leap of faith even if he wants to. I recently decided to try such a leap of faith with intrinsic reasons. I planned to take a moral weekend, but I soon realised that I could not sincerely claim to have the belief any more than an agnostic could take up Pascal's wager and opt to believe in the Christian God. A belief in the existence or inexistence of intrinsic reasons can only change involuntarily as with any other belief. It cannot be changed by decision.