Logic, Facts and Representation

An Examination of R. M. Hare's Moral Philosophy

Rønnow-Rasmussen, Toni

1993

Document Version:
Publisher's PDF, also known as Version of record

Logic, Facts and Representation

An Examination of R. M. Hare's Moral Philosophy

Toni Rønnow-Rasmussen

STUDIES IN PHILOSOPHY 6
Editors: Bengt Hansson and Göran Hermerén

Lund
University
Press

STUDIES IN PHILOSOPHY 6
Editors: Bengt Hansson and Göran Hermerén

Logic, Facts and Representation

An Examination of R. M. Hare's Moral Philosophy

Toni Rønnow-Rasmussen

Lund
University
Press

Tot era d'una estranya i inesperada flonjor

L'Estació

J. V. Foix

Lund University Press
Box 141
S-221 00 Lund
Sweden

Art nr 20283
ISSN 1100-4290
ISBN 91-7966-228-5 Lund University Press
ISBN 0-86238-324-4 Chartwell-Bratt Ltd

Printed in Sweden
Studentlitteratur
Lund 1993

Contents

Preface

I wish first to thank Ingmar Persson, for whose penetrating and helpful criticism I am highly indebted. More than anyone else he has helped me find my way around in the world of philosophical ideas. It has been a most exciting and inspiring guidance. For this, and for his willingness to offer help whenever assistance was requested I am greatful.

My thanks goes also in particular to Göran Hermerén. Besides providing valuable comments on the final draft and on numerous earlier drafts, he has over the years given me much useful advice on how to write a philosophical essay.

Throughout my work on this thesis I have benefited from discussions with several persons. Special thanks are due to Dan Egonsson, who read the entire manuscript and provided me with many helpful comments. I have also profited much from Peter Sandøe's comments on earlier, rough drafts of chapters I and III.

I owe thanks to my fellow participants at the seminars held at the University of Lund. I have learned much from them. I also record my gratitude to Bengt Åke Hoff for valuable assistance and support.

Thanks are also due to the Erik and Gurli Hultengren Foundation, and the Fil. Dr. Uno Otterstedt Foundation who enabled me to visit Oxford. Alan Crozier polished the English in this thesis, for which I am grateful.

Stefan Berglund deserves my warm thanks for spending hours discussing philosophy with me. To my wife, Elly, I owe more than I can say. She has been a constant support these past years, and she provided me with much help in getting the manuscript into a printable state.

Last but not least, I would like to thank Richard M. Hare, from whom I have learned so much. He scrutinized an earlier rough draft of chapter III. For this and for his valuable answers to the many questions that I raised during our meetings in England, Denmark, and Sweden, I am deeply grateful. The warm welcome which I received from Richard Hare and his wife during my visits to Ewelme and Oxford is a most pleasant memory.

I dedicate this work to my father for encouraging me to raise philosophical questions, to my daughters Anna and Sofia for making it clear

to me that some of these questions are more vital than others, and to Elly for making me realize the necessity of answering other than philosophical ones.

Toni Rønnow-Rasmussen

Abbreviations

For the sake of convenience I will use the following abbreviations to refer to Hare's major works (see bibliography for further details):

LM = *Language of Morals* (1952)
FR = *Freedom and Reason* (1963a)
PI = *Practical Inferences* (1971a)
MC = *Essays on The Moral Concepts* (1972a)
MP = *Applications of Moral Philosophy* (1972b)
MT = *Moral Thinking, Its Levels, Method, and Point* (1981)
HC = *Hare and Critics: Essays on Moral Thinking with Comments by R. M. Hare* (1988)
ET = *Essays in Ethical Theory* (1989a)
PM = *Essays on Political Morality* (1989b)

However, when referring to an article for the first time or when it is of special interest, I will let the abbreviation be followed by the year of publication. Thus, '(*PI*, p. 45, 1968)', for instance, should be read as saying that the quoted passage is found in *Practical Inferences*, on page 45, and is from an article originally published in 1968. Moreover, the following convention for quotation marks will be used in this work: to name an expression I put it between single quotation marks. To quote a passage I use either double quotation marks, or set it as a block. Finally, all quotation marks occurring either within quotation marks or block quotations are single.

Introduction

For more than forty years the numerous writings of Richard Mervyn Hare have been a dependable source of intellectual and philosophical stimulation. The attention which Hare's work has attracted is tremendous, and probably unparalleled among moral philosophers of today. In most cases the attention has been of a critical kind, and W. D. Hudson is surely right when he says that "few living philosophers have attracted more critical comment than [Hare]".

My aim in this book is to contribute to this very lively discussion between Hare and his critics, by examining a method proposed by Hare for solving our moral problems. This method consists in making a choice under the constraints imposed by the non-moral facts, and by the logical properties of the moral concepts. To make such a choice is what Hare in *MT* calls *critical moral thinking*. My examination will focus on two issues, namely how he defends his central metaethical claims and how he uses them to arrive at a utilitarian position.

For introductory purposes, let me outline what Hare's theory of critical moral thinking amounts to (for a recent and admirably clear survey of the development of Hare's ideas over the years, see W. D. Hudson 1988).

Let us begin with the metaethical part. Hare has focused attention notably on the three following logical properties, which he claims moral judgements have: (i) prescriptivity, (ii) universalizability, and (iii) overridingness. They are logical or formal in the sense that they have been established, according to Hare, on the basis of our understanding of how the moral words or concepts are used (cf. *ET*, p. 177, 1989f).

The claim that moral language in its primary sense is prescriptive, was argued for in Hare's first book *The Language of Morals* (*LM*). At the time the idea was not new that we are doing something else than asserting that something is the case, when we express a moral judgement. In the first half of this century, a value-theory had developed, viz., emotivism, that has been linked over the years perhaps especially to the works of the American philosopher Charles Leslie Stevenson (see especially 1944). In brief, Stevenson maintained that by issuing a moral judgement, such as 'This x is good' the speaker expresses his personal approval of x. There was a further

element to his analysis. Besides expressing approval, the speaker strives to evoke the approval of the hearer. The sentence 'This is good' means, according to Stevenson 'I approve of this; do so as well'. The former part of the analysans, 'I approve of that', constituted the descriptive meaning of 'This is good', whereas the latter, imperative part constituted what Stevenson referred to as its emotive meaning. The analysis presented in *LM* shows some affinities with that of Stevenson. Both theories mount an attack on descriptive analyses that take moral judgements to be logically equivalent to factual statements, and both theories maintain that by virtue of the meaning of moral terms, the speaker will not only be in a cognitive state, but also in a motivational state. However, Hare criticizes not only descriptivism but also emotivism, which he thinks mistakenly took the function of moral judgements to be persuasive or influencive.[1]

A central idea in *LM* is that we are doing something else than asserting or stating that something *is* the case, when we issue moral judgements. Value-language, in general, has a different function, namely to be *action-guiding*. By saying to someone, for instance, 'You ought to visit your sick mother', or 'This painting is good' we are typically guiding, or advising the listener to do something, viz., to go and see his mother, or to choose this painting rather than some other painting. In order to account for this prescriptive part, Hare argues that we must analyse moral judgements as entailing at least one imperative.

The classification of prescriptive language which he presents in *LM* consists of two major groups, namely imperatives and value-judgements. The former contains the subgroups (i) singular imperatives, and (ii) universal imperatives. Value-judgements, in their turn, are divided into (iii) non-moral judgements and (iv) moral judgements.

In this book I will focus nearly all attention on (iv). Moreover, most of the time I will be concerned with one special kind of judgement, namely 'ought'-judgements. The reason for this choice is that it is the logical properties of this word that Hare has invoked in order to secure his claim that taking facts and logic into consideration will lead to conclusions that correspond to those of a certain form of utilitarianism.

There is a another side to the prescriptivity-thesis, other than that moral judgements typically are used in order to guide the listener's actions. It

1. Hare's criticism of emotivism is found notably in Hare (1949) and in "Freedom of the will" published in *MC*, pp. 1-12. Cf. *MC*, p. 50, and his comment in Hare 1975, p. 202.

concerns what is involved in assenting to a moral judgement. This can be brought out by contrasting what is involved in assenting to a factual statement that something is the case, say, 'Lund is a university town in Sweden'. In such a case we are said to be sincere in our assent if and only if we believe that it is true that Lund is a university town in Sweden. But if moral judgements are not, as Hare argues, logically equivalent to factual statements, but must rather be analysed as entailing an imperative, it is the conditions for assenting to the latter that will be informative. Now, Hare maintains that "it is a tautology to say that we cannot sincerely assent to a second-person command addressed to ourselves, and at the same time not perform it, if now is the occasion for performing it and it is in our (physical and psychological) power to do so" (*LM*, p. 20). Moreover, since Hare explains what it is to have a desire in terms of assent to an imperative, we may, as Hare himself does, put this point another way: Whereas statements of fact express the beliefs of the speaker, prescriptions express the desires of the speaker.

The idea that moral judgements express the desires of the speaker gives rise to questions about the rationality of moral thinking. A widespread idea that stems at least from David Hume's *Treatise*, and which finds its perhaps most well-known expression in the phrase "'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger", is that desires, except for those instrumental ones that are appropriate as means, essentially are excluded from the realm of rationality. People have different desires, and if there is no way we can rule out some as being irrational, how can we aspire to show, as we will see that Hare claims in *MT*, that moral disagreement in this world will be entirely a matter of disagreement about factual or logical questions? Will there not in the end be a genuine moral disgreement that derives from the fact that the disagreeing parties have conflicting desires? The attempt to unite the prescriptivity-thesis with the claim that moral thinking is rational in the strong sense of closing the door to conflicts of desires, creates a tension in Hare's work, which I personally find most exciting.

In Hare's second book *Freedom and Reason* (*FR*) he sets out to explain, among other things, why there is no real conflict between the prescriptivity-thesis and the claim that moral thinking is rational. He does so by giving a detailed account of the second logical property of value-judgements, namely that they are universalizable. However, the attempt in *FR* to show what the requirement to universalize our prescriptions leads to, was not

entirely successful. There was room in *FR* for moral disagreement that did not derive from logical or factual mistakes. This possibility, he maintains, is excluded in his third major book, *Moral Thinking, Its Levels, Method, and Point* (*MT*).

What Hare means by universalizability can roughly be expressed in the following way: If we make a moral judgement about one situation (person or act), we must logically make a similar judgement about any other situation (person or act), real or hypothetical, that is exactly similar but for the fact that they are numerically distinct. Suppose I think that Fred ought to give money to Oxfam. It then follows, Hare thinks, that I must, on pain of committing a logical mistake, also prescribe that I ought to give money to this organization, if I am in exactly the same circumstances as Fred, including having his desires and beliefs. But to assent to such a prescription I must assent to the imperative which is entailed by this judgement. I must, in other words, actually desire that I should give away money in the situation imagined. This desire should be formed in light of logic and facts. In order to achieve knowledge of whether I am ready to prescribe such an act, I should represent to myself what it is like to be in Fred's shoes.

Now, suppose my representation of what it is like to be in Fred's situation is properly conducted. In such a case I will, according to Hare, acquire a preference that is exactly similar to the one Fred has but for the fact that mine concerns a hypothetical situation. This acquired desire or preference – what I will call a person's *conditional desire* or *conditional preference* – is not a hypothetical desire but an actual desire which the person has for some hypothetical situation.

Now, having performed such an act of representation, my decision concerning whether Fred ought to give money away, will be determined, Hare claims, by the strength of my original preference, and the acquired, conditional preference concerning the hypothetical situation. Suppose that Fred's desire is stronger than mine. In such a case I will, if I have represented his situation correctly, prescribe that Fred ought not to pay. I will in effect answer the question as a preference act-utilitarian would have done.

The method for solving our moral problems which I have outlined briefly above is what Hare, as may be recalled, calls in *MT critical moral thinking*. According to Hare, there is a further level of moral reasoning – what he calls "intuitive thinking". I will in this work say next to nothing about the latter. My reason is not that I think it is in no need of discussion, but rather

that the more fundamental, critical level contains sufficient material for one work. The same applies to the third logical property, "overridingness", which in itself is far from unproblematic. However, since it plays no vital role for the above argument, I will in this work leave the discussion of it to one side. Let me therefore here say a few words about these two topics.

Hare tells us that there are good reasons for not all the time conducting one's moral thinking on the critical level. To begin with, the moral principles of critical thinking - namely the principles we arrive at by universalizing our moral judgements - will be highly detailed. The complexity of critical moral principles can reach such a degree that it becomes impossible even to formulate them verbally in sentences of manageable length. If we want principles to be practical guides, they must be sufficiently unspecific, to cover a variety of cases, which all have certain salient features in common. A further reason for relying much of the time on such general principles is that "if we do not, we expose ourselves to constant temptation to special pleading" (*MT*, p. 38). When we reason in accordance to such, what he calls, general prima facie principles, we are on the intuitive level.

Acting in accordance with "intuitive" principles, which Hare claims we do most of the time, is an essential part of moral thinking. Such principles derive from our upbringing and our experiences in the past. However, those principles are not self-justifying. We can, according to Hare, "always ask whether the upbringing was the best we could have, or whether the past decisions were the right ones, or, even if so, whether the principles then formed should be applied to a new situation, or, if they cannot all be applied, *which* should be applied" (*MT*, p. 40). To appeal to intuitions in order to solve these problems will not do. This would be circular, since our problems arise because our intuitions are called into question.

To solve the above problems, Hare maintains, we must avoid making any appeal to intuitions other than linguistic. We must do some critical thinking.

Critical thinking has a further function besides solving, for instance, conflicts between different intuitive principles. Since it is helpful to have such intuitive principles, critical thinking will help us find the right principles by setting a standard. The set of prima facie principles to choose are those whose "acceptance yields actions, dispositions, etc. most nearly approximating to those which would be chosen if we were able to use critical thinking all the time" (*MT*, p. 50).

The prima facie principles differ from critical principles in more than

being non-specific. In contrast to critical principles, the prima facie principles lack the property of overridingness. Principles and considerations, other than moral ones, can in case of conflict override prima facie principles. This is not the case with critical thinking and its principles, which override all other principles when they conflict with them.[2]

Now, when I set about writing this essay, it was with the ambition of giving a thorough critical account of all the central issues in Hare's moral philosophy. However, realizing along the way the wealth of ideas that Hare has addressed over the years, it became evident that I had to lower my ambitions. To limit the scope of issues, I decided therefore to focus on what Hare has called critical thinking. Moreover, since Hare's views of the logical properties of moral discourse range over a number of different issues, it became advisable to concentrate on those of his views, the understanding of which is required in order to be able to assess his claims about what follows if we reason in accordance to logic and facts.

My examination will, in other words, by no means exhaust all there is to Hare's moral philosophy, not even all the issues raised by his theory of critical moral thinking. However, I hope that the importance of the claims discussed will make the reader bear with the fact that the present book does not entirely live up to its subtitle.

One of the issues that I found especially difficult to leave out concerns Hare's views on the problem of weakness of will. This is not the only topic, however, the examination of which would be worthwhile to engage in. Another such subject, for instance, concerns Hare's innovative work in the field of imperative logic. Still, I do not see how I could, within the frame of the present work, do justice to these matters. Moreover, I am not confident that I could contribute to the discussion in either field.

This work falls into two parts, viz., one major part concerning Hare's metaethics, and one that focuses on his theory of moral methodology. The latter, as Hare has explained, concerns "how moral thinking ought to proceed, or how moral arguments or reasonings have to be conducted if they are to be cogent" (*ET* p. 144). The three first chapters belong to the former part. Of these the two first ones will deal with Hare's thesis that

---

2. The claim that moral judgements override any other consideration has been questioned by several writers. See for instance Williams (1976), Foot (1978) and Michael Slote (1983). More recently Thomas Spitzley (1989) has questioned Hare's claim that the prima facie principles of intuitive thinking are universal and prescriptive. See also Garrett Thomson (1987, pp. 118-121).

evaluative language is primarily prescriptive. In chapter I I consider his arguments for why he thinks we must analyse a moral judgement as entailing at least one imperative. I consider in addition two arguments which Hare has put forward against attempts to analyse moral judgements as factual or descriptive statements. In chapter II I go on to examine in greater detail how he characterizes the speech act of prescribing. Chapter III, in its turn, will focus on his universalizability-thesis. In the subsequent chapter I present a preliminary answer to the question of how logic and facts, according to Hare, constrain what moral judgements we can endorse. That is, having examined what Hare means by, as well as the reasons he offers for his claim that moral judgements have the two logical properties of universalizability and prescriptivity, I outline how taking these two properties, together with the facts of the case, into consideration will constrain our moral reasoning. Finally, I examine in chapter V a claim concerning the meaning of 'I' which plays a crucial role in Hare's derivation of utilitarianism. These chapters are then followed by concluding remarks, where I will summarize my main conclusions.

Another way of bringing out the above division is to say that chapters I–III concern the *logic* of moral terms, chapter IV how Hare thinks that *certain facts* and the *logical properties* of moral words will constrain our moral reasoning, and chapter V focuses on a further conceptual claim made by Hare, namely that 'I' has prescriptive meaning.

Let me at this point make some general methodological remarks. The aim of this work is, as mentioned, to critically examine Hare's theory of critical thinking. However, is it at all possible to speak of one theory? The fact that Hare's works range over such a considerable period of time gives rise to the question of whether it is in fact possible to integrate his various writings in a homogeneous theory of moral reasoning. When Hare in *MT* speaks, for instance, of universalizability, which he considers to be one of the three logical properties of moral terms, is it really the same idea as when he spoke of this alleged logical property in earlier works?

However interesting a detailed examination and presentation of the evolution of his various idea would be, I have in this essay not made it my aim. Rather, I have, in trying to answer my two main questions, borne in mind the fact that Hare has himself commented in more than one place that it is not the case that his later views have changed substantially from his earlier ones (cf. e.g. *MT*, p. vi; *HC*, p. 201).

As mentioned above, my aims do not extend to an encompassing

assessment of Hare's metaethics. Nor will I try to show whether or not Hare's theory of critical thinking gives rise to moral conclusions that correspond to my own "moral intuitions". I have in this work refrained as far as possible from stating any external criticism of a moral kind. The reason for this is intimately linked to Hare's controversial thesis that, given that we accept the requirement to take facts into consideration, he will not need to make anything but *formal* claims in order to show that an answer to, say, 'What ought I to do in this situation?' will correspond to the answer given by a preference act-utilitarian. Moreover, he understands such formal claims, as mentioned, as being claims that are established only by appeal to considerations that "can be established on the basis of our understanding of the words or concepts used" (*ET*, p. 177; see also Geach 1981, and Hare 1981b). Given the nature of Hare's claim, an examination of it must reasonably take priority over any moral criticism. My aim has therefore been to concentrate on whether what Hare says about the logic of moral terms, and what follows if we take logic and facts into account when answering our moral questions, stands up when it has been subjected to questions such as: Is Hare's view of the logical properties of 'ought' correct? Given that it is correct, and that we take facts into consideration, will we all then be bound, as Hare claims, to agree to the same moral 'ought'-judgements?

We may find it likely or unlikely that taking facts and logic into consideration would serve us as outlined above. Most of us would, I venture to say, welcome a method by which we could answer our moral questions, and solve our moral disagreements with others. If taking facts and logic into account is what we have to do in making up our minds about moral issues, it would seem to simplify our problems a great deal.

I. Hare's Argument for the Prescriptivity of Moral Language

1.0 Introduction

The topic of this and the following chapter is Hare's claim that moral language, at least in its primary sense, belongs to the language of prescriptions. Hare has over the years done much to elucidate this prescriptive feature of moral language. However, I will not be concerned with all aspects of his claim but will centre my attention mainly on two fundamental issues.

The first bears on his argument for why he thinks that moral judgements must be analysed as carrying a prescriptive or imperative element in their meaning. I will in the present chapter present and examine this argument in sections 1.3–1.4. However, before entering into this issue, a word must be said about Hare's view of the very subject matter of moral philosophy. I will do this in section 1.1. Section 1.2 contains a presentation of some central expressions and concepts which Hare applies, and which it is important to be clear about. Having examined Hare's argument, in section 1.5 I will suggest a way of understanding 'ought' that does not require that we analyse 'ought'-judgements as carrying a prescriptive meaning-element. Although I think this analysis is promising, my main reason for introducing this idea here is to show that such an analysis would fulfil one strong requirement which Hare's linguistic intuition places on an analysis of 'ought'. I have been unable to present and defend, within the frame of this work, an alternative to Hare's analysis of moral language. Finally, I will in the remaining sections of this chapter consider two objections which Hare puts forward to what he regards as opposing theories to his own.

The second issue, which I will bring up in the next chapter, concerns his characterization of the speech act involved in issuing a prescription.

Now, a customary way of defending and supporting a thesis or theory in philosophy, is to bring the defects and fallacies of opposing theories to the surface. Hare is here no exception. Over the years he has put forward several penetrating objections to the various opposing theories to his own

one. Hare actually construes the central issue in moral philosophy as one between two opposing claims about the meaning of moral terms, namely, what he calls "descriptivism" and "non-descriptivism". Thus, the rationale of my examination of the two earlier mentioned arguments which Hare directs against descriptivism, is the following: since both arguments are aimed at the refutation of non-prescriptivist analyses of moral judgements, the success of either of these could easily, I imagine, have the effect of lending some credibility to a prescriptivist (or emotivist) analysis. Since I do not think these arguments are as conclusive as Hare thinks they are, and since I will put forward objections to his prescriptivism, it will be important to evaluate the success of these two arguments. This will be done in sections 1.7–1.8. Moreover, since one of these arguments bears a strong resemblance to G. E. Moore's "open-question" argument, I will in section 1.6 briefly comment on the pros and cons of this argument.

Let me now begin by presenting the divisions of moral philosophy which Hare makes.

1.1 Moral Philosophy

To restate the various divisions of ethical issues which Hare has discussed, it will prove illuminating at this stage to bring in the generally recognized distinction between first- and second-order questions of moral philosophy. Although I know of no good attempt to spell out the necessary distinction between these levels, the questions are sufficiently distinguishable for my present preparatory purpose.

First-order questions are often said to all concern matters of substance. A more instructive way of explaining the underlying notion is to give examples. Thus, the expected answer to such a first-order question will tell us, for instance, what particular action we morally ought to do, or what particular properties a good person will have. 'Ought I to break my promise to Jim, in order to help Charlie?', 'Is racial discrimination wrong?', and 'Am I a good man?' are all such first-order questions. In contrast, second-order moral questions were for a period dominated by questions of the meaning or logic of the terms involved in answering first-order questions. Hare makes use of this distinction in several of his works (see especially his article "Ethics" reprinted in *MC*. For a similar but more recent account see *ET*, pp. 175–176).[1]

1. Hare sometimes adds a third kind of question, namely what falls under the heading of "descriptive ethics", which concerns "questions of facts about people's moral

As has been hinted at already, and as will be shown later on, Hare construes the controversy between his own prescriptivist account and competing theories as a question about what is involved in assenting to a moral judgement as well as an issue about what kind of speech act is performed in issuing a moral judgement. In other words, Hare regards the central second-order moral questions as bearing on the meaning of moral terms. He divides theories about the meaning of moral terms into the following two fundamental groups:

(I) (i) descriptivism vs. (ii) non-descriptivism.

Hare conceives of descriptivism as the view that the logical character of moral judgements "is similar to that of other descriptive statements" (*ET*, p. 18, 1976b). Non-descriptivism, on the other hand, appears to be the view that moral judgements do not (primarily) have a logical character similar to descriptive statements. Thus, in one place he suggests that "non-descriptivism" is the most perspicuous term "for those who do not think that [the] central function [of moral judgements] is to state facts at all" (*ET*, p. 34, 1979c).

Often the controversy in (I) is said to concern the kind of speech act which is involved when we issue moral judgements. The issue in (I) then takes the form of a discussion about whether the function of moral judgements is to

(I') Tell someone that something is the case vs. to tell someone to make something the case/to convince someone to make something the case (cf. e.g. Hare 1969).

Sometimes Hare construes this issue as a psychological thesis concerning whether moral judgements express (cf. *ET*, p. 83, 1985)

(I") Beliefs of the speaker vs. Attitudes of the speaker.

Among the non-descriptivist theories, Hare makes the further division between what he thinks are the two major forms of this kind of theory:

---

opinions" (*MC*, p. 39). See also e.g., 1989f.

(I:ii) Emotivism vs. Prescriptivism (cf. *ET*, p. 101)[2]

Turning to descriptivism, Hare subdivides this kind of theory in more than one way: one recurring division is the one between

(I:i) Naturalism vs. Non-naturalism/intuitionism

In a recent paper he describes these views in the following way:

> We have, first, various so-called 'naturalist' theories. The distinguishing feature of these is that they hold that moral judgements are equivalent in meaning to factual statements of some ordinary non-moral kind (*ET*, p. 100, 1987c).

> Secondly, we have various kinds of 'intuitionist' theories, according to which moral judgements are statements of fact indeed, but statements of a special kind of fact which is not open to ordinary methods of discovery, but requires a special kind of moral thinking to ascertain it (*ET*, p. 100–101).

Hare sometimes discusses another division, namely, the one between "cognitivism" and "non-cognitivism", which he thinks are notions referring to different types of ethical theory (*ET*, p. 18). Consider the following passage, for instance:

> There is first of all the division between what are sometimes called cognitivist and non-cognitivist theories, but which I prefer to call descriptivist and non-descriptivist theories [. . .]. The pairs of terms do not mean precisely the same: ethical cognitivism is, presumably, if we are to rely on etymology, the view that moral judgments can be *known* to be true; ethical descriptivism is, rather, the view that their logical character is similar to that of other descriptive statements or judgements (*ET*, p. 18, 1976b).

Some lines further on he introduces a further division which it is important to notice:

> The second division that I wish to note is the division *within* descriptivist or cognitivist theories between those which are subjectivist and those which are objectivist. I repeat, with all the emphasis I can muster, that this is a division within descriptivist theories. It is not the same as the division between descriptivist and non-descriptivist theories (*ET*, p. 18, 1976b).

---

2. This is, in turn, most often conceived of as a controversy about whether moral judgements have a function to get people to do something vs. to tell them to do something. More about this in chapter II.

> According to a *subjectivist* descriptivist theory, [moral judgements] are descriptive of states of mind, dispositions, etc., of people (usually those who make them); whereas an objectivist descriptivist theory holds that moral judgements are descriptive of states of affairs other than states of mind, dispositions, etc., of people (*ET*, p. 19, 1976b).

The latter passage suggests that he conceives of objectivist as well as subjectivist forms of naturalism. Non-naturalism, on the other hand, would be an objectivist form of ethical theory.[3]

Of the various theories that have been criticized by Hare, ethical naturalism fills a central place. In addition to the main aim of this chapter, I will, as mentioned, end it by examining what is, if I have understood Hare correctly, his central and supposedly strongest objection to what he calls naturalism. This objection applies in fact equally well to any theory that analyses moral judgements as statements of facts – whether these are natural or non-natural facts, or whether these are facts about the speaker or not about the speaker. The objection, it is reasonable to say, is in other words an objection to descriptivism.

Now, given this dichotomy – 'descriptivism/non-descriptivism' – a critic of Hare's (non-descriptive) prescriptivist account is apparently stranded with the dilemma (or option) of becoming either some kind of a naturalist or a non-naturalist. But a moment of reflection discloses the rashness of such a conclusion. My primary concern in this and the next chapter is to examine some essential aspects of Hare's prescriptivist account, not to choose sides in a controversy, which, for all I know, eventually may well be regarded as being non-exhaustive or simply wrongly construed. It is plausible, for instance, that more work in philosophy of meaning will shed light on the issue between prescriptivists and descriptivists.

At this juncture let me comment on a fundamental issue that lurks in the background of the controversy between the various theories which Hare calls descriptivism and his own prescriptivism, namely what the proper subject matter of moral philosophy is. The size and complexity of this matter is far too great to be dealt with inside the present framework. Nevertheless, the issue deserves to be commented on.

Hare regards, as mentioned, the central second-order moral questions as bearing on the meaning of moral terms. Now, the issue to which I referred

---

3. That non-naturalism would be an objectivist form of ethical theory gets support from Hare's discussion of different theories in *MT*. See e.g. p. 72. Relevant is also p. 76 where he speaks of a "sub-species of naturalism", namely, what he calls "'old-fashioned subjectivism'".

above concerns whether second-order questions should deal exclusively with matters of logic and meaning. Is there room for more than logical analyses and questions about meaning on the second-order level? Should ontology/metaphysics and epistemology occupy the itinerary of the second-order moral philosopher? Hare, if I have understood him correctly, considers that ethics (which he understands as second-order moral philosophy) does best to avoid ontology and metaphysics (e.g. *ET*, 1985 and *ET*, 1984b).

Thus, the apparent attempt above to make the division between descriptivism and non-descriptivism correspond to the one between cognitivism and non-cognitivism, reflects Hare's scepticism about the place of epistemological questions in second-order moral philosophy. This scepticism extends also to ontological questions. There are mainly two ways of understanding Hare's reluctance to deal with these questions. On the one hand, he has in some articles expressed the view that *in ethics* ontological as well as epistemological problems are really conceptual ones (cf. *ET*, p. 85, 1985). On the other hand, sometimes his reasons seem rather to be that the views of those who claim that value properties/facts exist and that we can have knowledge of them, are very implausible. This is, for instance, the approach which he appears to take in one place against the (Moorean) view that, say 'goodness' is a *sui generis* non-natural property, such that all things that are good, necessarily (but not analytically) have this *sui generis* property. Such a view, he thinks, "leaves the relation between descriptive and moral properties looking queer". Moreover, he maintains that such a view is a "sitting target for Mackie's 'argument from queerness'" (*ET* p. 72, 1984b). Mackie's argument (1981, pp. 38–42), it should be noticed, proceeds from the assumption, which Hare is well aware of (*loc. cit.*), that the question of whether there are such non-natural properties is not a conceptual but a real ontological issue.

I am inclined to think that there are genuine ontological and epistemological problems in moral philosophy. However, let me emphasize that by suggesting this I am not implying anything about what epistemological or ontlogical position to take. My inclination derives rather from the following contention: When we make up our minds about what ontology to endorse, linguistic evidence must reasonably count as only one kind of evidence. Mackie's so-called "error theory" deserves to be mentioned here. To put it briefly, the "error theory" suggests that in their ordinary use, words such as 'ought', 'good', and 'right' connote objective

properties of actions, people, and things. The error which people who use these words commit, according to Mackie, is a factual and not a conceptual error, viz., to think that those acts, etc., which they call morally obligatory or good, etc., have such moral properties. But this is a mistake, Mackie argues, since there are no such properties.

Mackie's "error theory" may be more or less plausible.[4] I mention it here since I wish at least to vent my intuition that there are other fundamental questions besides the one concerning the meaning of moral terms.[5]

1.2 Evaluative Moral Judgements

With his prescriptivist account, Hare takes sides in a most central discussion of analytic moral philosophy, namely, how to analyse and characterize the seemingly existing relation between assenting to a moral judgement and trying to act upon it. Following Falk (1947-48) it has become customary to describe this discussion as a controversy between 'internalist' and 'externalist' conceptions of this relation. As recent discussions have been held in these terms, let me say a few introductory words about what is the crux of the matter.

Generally speaking, the internalists argue that the relation between assent to a moral judgement and trying to act upon it, must be of a logically necessary kind, whereas the externalists deny this. From an internalist standpoint, assenting to, say, 'I ought to do *x*' logically commits me to do *x*, i.e. if I do not do *x* it is a sign either that I have been insincere, or that I have used 'ought' idiosyncratically or suffered from weakness of will. Externalists, although not denying that we are typically motivated to do what we think we ought to do, claim that the desire or intention to do that action which we think we ought to do, forms no part of what constitutes a sincere assent to value-judgements. The desire to do what we think is our moral obligation is viewed rather as being empirically correlated to such an assent.

---

4. Hare comments further on Mackie's "error theory" in Hare (1985), for instance.

5. Consider an example from religious language. It does not seem an unreasonable point of view to say that one can endorse one of a number of views (prescriptivist, emotivist, descriptivist) about the meaning of sentences such as 'God spoke to mankind' or 'I will do what God demands me to do', without having to deny that the question 'Does God exist?' concerns a real ontological issue. Bruno Schüller (1986) has an interesting discussion concerning to what extent Christian discourse about moral rules is consistent with Hare's metaethical prescriptivism.

Hare's reason for maintaining that moral judgements, if used evaluatively, must entail imperatives is, as will become clear, closely connected to his "internalist" view of the relation at issue. However, since Hare does not himself use this terminology I will most of the time avoid it, save in those cases where I think it will be helpful in order not to create any confusion.[6]

A frequently recurring idea in Hare's works is that we can use words such as 'good', 'right', and 'ought' in several distinct ways. Throughout his works, he has mainly focused on two genera of such uses, namely what he calls prescriptive or action-guiding speech acts, and what he sometimes calls assertive speech acts (occasionally he simply uses 'statements'). To perform a speech act belonging to the latter genus is to *tell someone that something is the case*, whereas to utter a prescriptive judgement is to *tell someone to do something.*

Now, according to Hare, moral language is in its primary sense prescriptive. However, he does not deny that moral judgements sometimes are used in order to make assertions. These are what he refers to as "inverted commas" uses. Such uses, he argues, are nevertheless logically secondary to the prescriptive use. By using 'good' or 'ought', for instance, in an inverted commas way, we are "not making a value-judgement ourselves, but alluding to the value-judgements of other people" (*LM*, p. 124). With regard to the sentence "I ought to do *x*" Hare suggests that it can be treated in three different ways:

> (1) 'X is required in order to conform to the standard which people generally accept' (statement of sociological fact)
>
> (2) 'I have a feeling that I ought to do X' (statement of psychological fact)
>
> (3) 'I ought to do X' (value-judgement) (*LM*, p. 167).

(1) and (2) are what he calls conscious and unconscious inverted commas uses respectively of the sentence at issue. Besides the inverted commas use, Hare distinguishes in *LM* between two further non-evaluative uses, viz., what he calls "conventional" and "ironic" uses. However, since the

---

6. For a recent discussion of the relation between assent to moral judgements and action, held in terms of the distinction between externalism and internalism, see Brink (1989). See also Silverstein (1983), who calls *LM* the "*locus classicus*" of internalism, Boatright (1973), Frankena (1958), and Falk (1947–48). Hare has recently applied the distinction at issue: "As an 'internalist' I hold that to utter sincerely a moral or any other prescription is to be (morally or otherwise) motivated" (*HC*, p. 252).

important feature here is that they are, like the inverted commas uses, logically secondary to the evaluative uses, I will henceforth concentrate on the inverted commas uses. It should, however, be understood that there are these further uses, according to Hare, besides the above mentioned ones.[7]

Now, if I say 'I ought to do *x*' I will be making either a statement of a sociological or psychological fact, or a value-judgement. Hare also seems to think there are combinational possibilities, so that 'I ought to do *x*' sometimes could be "(1) and (2), or all three, or some other combination" (*loc. cit.*). Unfortunately, Hare does not further explain just how he conceives of such combinations.[8] Although it is clear that we should regard this tripartite division as a blueprint, it would have been desirable if he had said something further about how one can make, say, a *statement* of sociological fact and at the same time be said to have made a *prescription*.

Now, early in *MT* (p. 22) as well as in *FR* (p. 26) he makes it true by virtue of definition that a value-judgement or evaluative judgement is a judgement which is used prescriptively. Thus, he says in *MT*:

> I shall be following my usual practice (FR 2.8) of using the words 'evaluative' and 'value-judgement' to apply to utterances which are both prescriptive and universalizable (or actual universal). I shall also use the terms 'evaluative meaning' and 'prescriptive meaning' for the element in the meaning of value-judgements which gives them their prescriptivity. The

7. That these uses shade into one another has been argued by for instance Carter (1982, p. 58 ff.). Hare further distinguishes between words in which the evaluative meaning is secondary to its descriptive meaning (see *LM*, 125 ff; *FR*, p. 24) – e.g., 'industrious', 'honest' and 'courageous'. In *LM* and in Hare 1957 he also considers a class of words which he calls "functional words": "A word is functional if, in order to explain its meaning fully, we have to say what the object it refers to is *for*, or what it is supposed to do" (*LM*, p. 100). Examples which he gives of such words are 'auger', 'knife' and 'hygrometer'. Such a word "gives us a partial specification of the virtues required" (*LM*, p. 139). For a critical examination of this issue, see e.g. Conway (1972), Nickel (1973), Margolis (1971, pp. 104–111) and Schüller (1986).

8. Although what Hare here says about the combinational possibilities may suggest the contrary, there is nothing elsewhere in his work to indicate that a single speech act can be both a case of someone stating a fact and prescribing. These are clearly two different speech acts. Relevant here is Hudson who argues against an alternative interpretation by Kerner (1966). Thus, Hudson (1978, p. 246) says "It is no part of Hare's position that a *single* speech-act may be both descriptive and evaluative (or prescriptive)". Hudson uses 'descriptive' where I would have used assertive (see chapter II). But the point is the same. Contrast Carter (1982, pp. 71, 249), who appears, in saying "that speech acts can be of greater or lesser complexity", to hold the possibility open that a single speech act can both be prescriptive and assertive. Hare returns in *MT* (p. 58) to the second kind of statement, claiming that we have to distinguish between "the judgement *that we have* the feeling, and the judgement which is the *expression* of the feeling". Unfortunately, Hare does not explain whether he conceives of 'expressing' as an illocutionary act. Relevant here is also the discussion in section 2.9.1.

> presence of this element does not prevent their also having a descriptive element in their meaning, so that it can be said to be in some sense right to apply them to certain kinds of things and not others (FR 2.1 ff.). But I use the term 'descriptive judgement (or statement)' only for judgements lacking any such prescriptive or evaluative meaning, i.e. those whose meaning is purely descriptive (*MT*, p. 22).

As Hare himself notes, this practice differs somewhat from how he used "evaluative meaning" in *LM*.

> In [*LM*] I used the words 'evaluative meaning' for the prescriptive meaning of evaluative expressions. This had some advantages, as being a less question-begging expression, which did not presuppose that what gave these terms their evaluative meaning was their prescriptivity; but in the end it turned out to be in the interests of clarity to make this, in effect, true by definition (*FR*, p. 27).

The account given in *LM* is interesting in that it is only late in that book (p. 168) that he makes it true by definition that value-judgements are prescriptive. This fact is noteworthy for the following reason. *LM* contains an argument to the effect that among the different ways in which we can use a judgement such as 'This painting is good' or 'He was in his right to stay home' or 'I ought to be more helpful', one is such that we must analyse it as entailing at least one imperative. If I have understood Hare correctly in *LM*, and in the light of the above passage, it seems reasonable to say that the argument at issue does not contain among its premisses the analytic definition in question.

Hare states explicitly his reason for why in the first place he introduces this definition in *LM*. Thus, he says that he thinks it will be impossible to "prove or even render plausible" (*LM* p. 168) that value-judgements, when they are held not to entail imperatives, are used in an inverted commas way. To overcome the difficulty of knowing whether a person has made a value-judgement or not, Hare suggests the following:

> I propose to get over this difficulty in the only possible way, by making it a matter of defintion. I propose to say that the test, whether someone is using the judgement 'I ought to do X' as a value-judgement or not is, 'Does he or does he not recognize that if he assents to the judgement, he must also assent to the command 'Let me do X' (*LM*, pp. 168/169).[9]

For the sake of clarity, I suggest we ascribe the following theses to Hare:

---

9. As was mentioned in the introduction, what distinguishes moral evaluative judgements from non-moral evaluative judgements, is the fact that the former are treated as overriding.

E1: An utterance which is both prescriptive and universalizable (or actually universal) is an evaluative or a value-judgement.

E2: 'Evaluative meaning' and 'Prescriptive meaning' are used for the element in the meaning of value-judgements which gives them their prescriptivity.

E3: A moral judgement which has no evaluative meaning is a moral judgement used in an inverted commas (or in a conventional or ironic) way.

E4: A descriptive judgement is a judgement which lacks any prescriptive or evaluative meaning. Its meaning is purely descriptive.

Hare is customarily interpreted as maintaining that moral terms and moral judgements have descriptive as well as evaluative/prescriptive meaning (e.g. Hill 1974, p. 60). However, Hare does seem to leave the door open for there being moral judgements that entirely lack descriptive meaning (see e.g. *FR*, p. 27; cf. Carter 1982, p. 65). Still, he offers no example of what would count as such a judgement. What exactly Hare means by 'descriptive meaning' is a matter I will bring up in chapter III, where I examine his universalizability-thesis. For the moment it suffices to restate Hare's own summary of what he means:

> a statement or judgement is descriptive if its meaning (including its reference) determines uniquely its truth-conditions, and vice versa (*ET*, p. 18, 1976b).

Hare's use of 'entailment' is likewise important to notice. As will be seen later on in section 2.8, Hare thinks there can be logical relations between imperatives. How we should characterize these relations in the light of the fact that the values 'true' and 'false' are not available, is a difficult problem that we need not go into here. In the case of the entailment relation, Hare suggests the following:

> A sentence P entails a sentence Q if and only if the fact that a person assents to P but dissents from Q is a sufficient criterion for saying that he has misunderstood one or other of the sentences. 'Sentence' here is an abbreviation for 'sentence as used by a particular speaker on a particular occasion'; for speakers may on different occasions use words with different meanings, and this means that what is entailed by what they say will also differ (*LM*, p. 25).

Let us now turn to Hare's reason for maintaining that value-terms are used prescriptively, as action-guides, and that they can be action-guiding only if they entail imperatives. I venture to say that it is hard to get a grip of his argument. The reason for this is that Hare has not sufficiently well explained the relation between the following theses:

(*) Value-judgements entail at least one imperative.

(**) Value-judgements are action-guiding.

1.3 Outlining Hare's Argument

The argument at issue rests on the following rule or principle (in the literature sometimes referred to as Poincaré's principle):

> (*S1*): No imperative conclusions can be validly drawn from a set of premisses which does not contain at least one imperative (*LM* p. 128).

This rule, Hare maintains, stands in need of qualification. There are so called hypothetical imperatives that can be deduced from indicatives. Moreover, in subsequent works, for instance in *MT* (p. 223), Hare claims that further qualifications must be undertaken. Notwithstanding, given that moral judgements are a kind of prescriptions, such inferences from indicatives to imperatives are in Hare's opinion "exceptions to the rule that there are no valid inference from facts to prescriptions" (*MT*, p. 223).[10]

Hare's idea seems to be that the rule holds good, and that the logical problems which arise in connection with it concern problems of how to define and correctly formulate it, rather than the rule in itself. Of course, this remains to be shown. However, let us sidestep this complicated issue and grant Hare that something to the effect of *S1* is valid. We may then concentrate on the very argument at issue.

---

10. Hare's rule has been questioned by several writers. For a recent objection, see Hamblin (1987, pp. 90–91). *LM* contains a further rule, viz., "No indicative conclusion can be validly drawn from a set of premises which cannot be validly drawn from the indicatives among them alone" (p. 28). Hare maintains (*loc. cit.*), however, that it is only the other rule with which he will be concerned. For a plausible counter example to the above rule, see Hudson (1978, p. 234), for instance. For an attempt to derive prescriptions from statements of fact, see Geach (1977). See also Ibberson's (1986) successful criticism of this derivation. For an objection to Hare's formulation, see Castañeda (1963) in Castañeda and Nakhnikian, 1963, p. 229. Sen (1966) and Morris (1966) are also of interest.

The second important part of the argument originates from Hare's contention that moral judgements primarily are action-guiding. The following passage expresses well the backbone of the argument:

> If we admit, as I shall later maintain, that it must be part of the function of a moral judgement to prescribe or guide choices, that is to say, to entail an answer to some question of the form 'What shall I do?' – then it is clear, from . . . [*S1*] just stated, that no moral judgement can be a pure statement of fact (*LM* p. 29).

Another passage brings out this relation between (*S1*) and the action-guiding nature of value-judgements:

> Since a large part of my argument hinges on the assumption, hitherto not fully defended, that value-judgements, *if they are action-guiding*, must be held to entail imperatives (*LM*, p. 163, my italics).

The latter as well as the former passage give the impression that Hare means something more by action-guiding than merely 'entailing at least one imperative'. If not, at least the second passage would lose its force, in that the passage would turn out to be trivially true.

The former passage suggests that we may ascribe the following thesis to Hare:

(*S2*) Moral judgements are primarily used for action-guiding.

The word 'primarily' is motivated, as may be recalled, by the fact that Hare thinks moral judgements can be used in ways that do not need to be analysed as being action-guiding.

Now, Hare does not further explain what he means by "action-guiding" other than what he states in the former of the two above passages. There he seems to give an explanation, namely when he maintains that the function of moral judgements is to guide choices "*that is to say,* to entail an answer to some question of the form 'What shall I do?'" (*LM*, p. 29. My italics). "Action-guiding" is apparently explained in terms of being answers to questions of the form 'What shall I do?', or as he puts it in *FR*, 'What ought I to do?'(cf. *MC*, p. 7).[11] Hence, we may reasonably expand (*S2*) into:

11. Thus, in *FR* he adjusts the picture from *LM*: "The question to which 'ought' gives an answer is not that asked by a man who is wondering what to do, but that asked by a man who is wondering what *he ought* to do" (*FR*, p. 56). These two questions are related in the sense that "unless the practical question ['What shall I do?] arises, the 'ought'-question cannot arise, if 'ought' has its full force (as it must have, if it is to imply 'can')" (*FR*, p. 56. See also p. 59). A critical examination of what Hare means here by "arise" is

(*S2*) Moral judgements are primarily used for action-guiding, i.e., they are primarily used in order to entail answers to questions of the form 'What shall I do?'

Basing himself on what I have called (*S1*) and (*S2*), Hare mounts in *LM* an attack on what he calls naturalistic analyses:

> A statement, however loosely it is bound to the facts, cannot answer a question of the form 'What shall I do?'; only a command can do this. Therefore, if we insist that moral judgements are nothing but loose statements of fact, we preclude them from fulfilling their main function; for their main function is to regulate conduct, and they can do this only if they are interpreted in such a way as to have imperative or prescriptive force (*LM*, p. 46).

In due time I will consider his argument against such naturalistic analyses. Here I wish merely to draw attention to a third premiss in his argument, which this passage contains. By interpreting "regulate conduct" as being equivalent to 'guide action', we may ascribe the following to Hare:

(*S3*): To be action-guiding, i.e., to answer the question 'What shall I do?', a sentence must entail an imperative.

In other words, if the sentence (i) 'I ought to do x' is used to guide action, and it is analysed as being equivalent to some indicative, the analysis fails to account for the action-guiding nature of the moral judgement. The reason it fails, it seems, is that only an imperative or command can answer the question 'what shall do?'. Moreover, since only an imperative can account for the action-guiding use of (i), and given that (*S1*) is valid to the effect that we cannot derive imperatives from indicatives alone, we must analyse (i) as having imperative force.

This, I suggest, constitutes the premisses of the argument in Hare's works, for why certain uses of sentences containing moral terms must be understood as entailing imperatives.

1.4 Examining the Argument

Let us recapitulate the theses involved so far:

(*S1*) No imperative conclusions can be validly drawn from a set of premisses which does not contain at least one imperative (*LM*, p. 128).

---

given by Tännsjö (1974, pp. 122–124).

(*S2*) Moral judgements are primarily used for action-guiding, i.e., they are primarily used in order to entail answers to questions of the form 'What shall I do?'

(*S3*) To be action-guiding, i.e., to answer the question 'What shall I do?', a sentence must entail an imperative.

Now Hare draws two conclusions in *LM*. From (*S3*) and (*S2*), he infers that:

(*C1*) Moral judgements entail at least one imperative.

Moreover, from (*C1*) and (*S1*) he further concludes:

(*C2*) Primarily, moral judgements are not purely descriptive.[12]

The crucial claim involved in these derivations is (*S3*). The question is why moral judgements must be said to entail imperatives in order to be action-guiding. After all, it is not as if moral philosophers, who think that a proper analysis of moral judgements does not require that we ascribe to these an imperative part, deny that they are action-guiding. Boatright lays his finger on the crux of the matter:

> for Hare the reason why no moral judgement can be a statement of fact is that the answer to the question 'What shall I do?' must be an imperative and no imperative can be derived from a pure statement of fact. But even if we accept this latter claim, we may still ask why only imperatives can be answers to questions of this sort. Why can only imperatives be practical? (Boatright 1973, p. 319).

In *LM* Hare offers no reason, as I can see, for (*S3*). However, he has recently commented:

> I used to argue, and would still argue, that since the main point of moral judgements is to guide actions, and since they cannot do this without having imperatives as their consequences, they cannot themselves be derivable from pure non-moral 'is'-statements, since *these* do not have imperative consequences [. . .]. Though this argument is valid, I gave up using it because objectors tended to deny one of its premisses, namely that moral judgements entail imperatives. I still think they do, if used prescriptively (this is indeed a tautology); but the matter obviously needed more discussion which I duly gave it (Hare 1987b, pp. 82/83).

This passage supports the above interpretation that there is a sense of

12. A similar argument is found in Nowell-Smith (1954) and Ibberson (1986, p. 3).

'action-guiding' such that Hare evidently takes it for granted that using a moral judgement in this sense must entail imperatives.

(*S3*) has been questioned by a number of writers (e.g. Nesbitt 1973; Hancock 1963; McGuire 1961).[13] Examples have been produced, for instance, of alleged action-guiding statements of facts. Versényi (1973) is here a case in point:

> Now if we take such action-guiding character by itself as our criterion for distinguishing between prescriptive and descriptive statements, statements of value and statements of fact, it seems to me that the distinction between description and prescription, fact and value, collapses. For purely descriptive statements, statements of fact, can be just as action-guiding as commands, imperatives and oughts (Versényi 1973, p. 22.).

An example which shows this, Versényi thinks, is the following:

> If I make the descriptive statement 'this is poison' and someone proceeds to drink the stuff while cheerfully protesting that he does not want to die, his action manifests that he did not understand or accept my statement, or else was lying about wanting to live (*ibid.*, p. 22).[14]

However, Hare's argument is not so easily refuted. That there is a sense in which a statement of fact can be intended to change the behaviour of the listener, was seen early on by Hare. Thus, he says:

> If you want a man to take off his trousers, you will more readily succeed by saying 'A scorpion has just crawled up your trouserleg' than by saying 'Take off your trousers' (*PI*, p. 21, 1949).[15]

However, such uses of statements, which in later works he calls "action affecting" (*MC*, p. 32), is not what he has in mind, when he speaks of

---

13. These writers see that Hare gives no reason for what I have called (*S3*). McGuire's account differs, though, from my account. According to McGuire, Hare aims at establishing that there are "imperative entailing uses", whereas I think Hare assumes that there are "action-guiding" uses, the analysis of which requires an imperative analysis. Nesbitt puts forward other objections. However, these objections constitute no serious threat to Hare's account. For instance, Nesbitt claims that imperatives "cannot" answer 'What shall I do?'(*op. cit.*, pp. 256, 257), since an appropriate answer is an indicative such as "I command you . . .". Hare would maintain that this is in fact a prescription. In chapter II I consider objections similar to those of Hancock (1963)

14. For another example, see Railton (1990, p. 151). Relevant here is also Sumner (1967, p. 788).

15. See here also Hare's reply to Geach's objection in *MC*, p. 32.

"action-guiding". If we turn to *MT* there is a passage that suggests why Hare does not consider statements to be action-guiding:

> The mere fact that what we say could be given as a reason for acting in some way does not make it prescriptive. That the hotel faces the sea could be given as a reason for taking a room there, but to say that it does is not to say something prescriptive; for somebody who did not like looking at the sea could sincerely assent, and yet not take a room there (*MT*, p. 21).[16]

The sense of 'action-guiding' which Hare thinks of, is such that there is a necessary relation between assenting to the moral judgement and acting on it. This is, as mentioned before, the central claim of internalism. What Hare argues is that value-judgements, analysed as being statements of facts, can only establish a contingent relation between assenting and acting.

1.5 Action-Guiding Statements of Facts

Above I attempted to show what Hare's reasons are for claiming that some uses of moral judgements must be analysed as entailing imperatives. His reasons, as we saw, are derivable from his internalist position. Of course, not all writers share Hare's internalist intuitions. Boatright, for instance, maintains that although moral judgements are action-guiding (he uses the term "practical") "there is no reason as yet why [the relation] must be a logically necessary one" (Boatright 1973, p. 320. Recently O. Brink 1989, expresses a view to much the same effect. See for instance p. 8).

However, even if we shared Hare's internalist linguistic intuitions and held that assent to a value-judgement commits one logically and not merely causally to action - my own language intuituions make me part company with Boatright and the externalists - the question remains whether we must bring imperatives into the picture as Hare does. Are there statements of facts, assent to which logically commits us to acting? On the face of it, many statements of facts would seem to qualify in this respect. Consider for instance the following sentence 'You have a (dominating) desire to do X'. Among a great many philosophers, including Hare, a common way of

---

16. According to Hare there is a sense, viz., a prescriptive one, of "better than", such that we cannot say of a hotel that it is better than another hotel, and then "when faced with a choice between the two hotels (other things such as price being equal) choose the other hotel" (*MT*, p. 21). For a criticism of this example, see Williams (1985, p. 125). Relevant here is also Rennie (1968), who objects to the definition which Hare gave of "better than" in *LM* (p. 184), viz., "'A is a better X than B' is to mean the same as 'If one is choosing an X, then, if one chooses B, one ought to choose A'"

analysing 'desire' is to say, in brief, that to have a desire is to be disposed to act. On such a view, assent to the above sentence would appear to commit the one who assented to action in the required sense.

It is beyond the scope of this work to give a detailed account of whether the relation between assenting to a statement of fact and acting on it can be action-guiding. However, I will in a few words at least outline one promising path to take, which analyses moral judgements as being equivalent to statements which seem clearly descriptive, without losing sight of the logical tie between assent and action. However, a caveat must be mentioned here. Such an analysis will be internalistic in a wider sense than the one favoured by Hare. But this version of internalism is no less plausible than the one endorsed by Hare.

On this wider version of internalism there is the further proviso advanced to the effect that what the person assents to must be *correct*. For instance, I may sincerely assent to 'You have a dominating desire to do *x*' and still not do *x*, because it was not correct that I had such a dominating desire to do *x*. Hare claims there is a logical connection between assent and acting, under the condition that the assent is sincere. The wider notion adds that the speaker must not only be sincere, but that what he assents to must also be true.

Consider the following sentence:

(s) '*a* has a stronger reason R to do *x* than to avoid *x*'

Now if we could show that there is a necessary relation between the notion of having a want and the notion of having a reason, we would have come a long way towards an explanation of how an assent to a statement of fact, such as (s), can be conclusively action-guiding. Such a view has been put forward by Persson (1981), who argues that describing "a proposition as a reason for acting involves a reference to attitudes of wanting" (*op. cit.*, p. 129), and where 'wanting' is analysed in terms of dispositions to act.

Suppose that there is such a necessary link between having a reason for acting and having dispositions to act. If *a* understands by (s) that he is supposed to have reasons for doing *x*, then assenting to (s) would also mean that he had a want for *x*. Given a dispositional view of wants, it follows that if (s) is true then *a* will, *ceteris paribus*, perform the action *x*. If *a* did not do

x, it would be a sign that *a* had either changed his mind, or that he was insincere (or suffered perhaps from weakness of will).[17]

Is (s) a reasonable *analysans* of 'I ought to do *x*'? It is likely that (s) would need to be qualified somehow, in order to distinguish moral reasons from the reasons which a prudent and/or altruistic and/or egoistic person would have. We would perhaps require that the reasons were universalizable and/or overriding. In addition, certain wants and preferences should perhaps be excluded. Nevertheless, how and to what extent the *analysans* stands in need of qualification will not change the fact that an assent to a statement of fact, such as (s) above, strongly appears to logically commit one to action.

To this it might be objected that, although an assent to (s) necessarily commits one to action, the person who says 'You, *a*, have a stronger reason R to do *x* than for avoiding X', cannot be said to be guiding *a*'s action. The person cannot be said to tell *a* to make something the case. But why must issuing a moral judgement be an instance of telling someone to make something the case? Is it that clear that maintaining that someone ought to do something requires something of the speaker? Consider the two following judgements:

(x) 'You ought to do X, but don't'

(y) 'You ought to do X, but you have no stronger reason to do X than to avoid X'.

If I have understood Hare correctly, to issue the judgement in (x) would be a logical misuse of 'ought'. On the other hand, (y) would be consistent with a prescriptive analysis of 'ought'.[18] But is it that obvious that (x) but not (y) is an example of something we logically cannot say? My own

---

17. In *MT* (p. 71) Hare puts forward an objection which he thinks could be directed against any descriptivist theory. Thus, he says "Descriptivism of any sort, if absorbed and practised by any section of society, will lead to the adoption by them of a 'So what?' morality; they will be able to say 'Yes, I know it would be wrong: so what?". However, given the above analysis, not all forms of descriptivism will be troubled by the "So what?" question. Only an irrational person would not do what he has most reason to do.

18. See Nagel (1988) and Hare's comment (both in *HC*). Nagel puts forward an alternative to Hare's analysis, which my own alternative resembles. In Nagel's view, to say that someone ought to do something entails that the person has a reason to do it, even if he does not happen to recognize that reason. Hare replies as could be expected "I can certainly say to someone 'You ought to do it' without implying that *he* has a reason or motive (yet) for doing it" (*HC*, p. 251). See also Wolff (1989, p. 206).

linguistic intuition makes me depart from Hare's view here. The judgement in (y) appears to me logically incorrect.

Hare has no argument, as I see it, that sanctions the claim that they *must* be instances of someone telling someone to make something the case. It will not do to invoke *S2* here. In other words, the following reasoning is erroneous: Since moral judgements entail imperatives, we must conclude, in the light of the principle *S2* that the moral judgement is in the imperative mood – where this is understood as that mood in which we are telling someone to make something the case.[19] If – as suggested above – there is no need to bring in imperatives in order to explain the logical tie between assent to a moral judgement and a logical commitment to acting upon it, Hare's second conclusion from above is invalid. That is, the conclusion:

(*C2*): Primarily, moral judgements are not purely descriptive.

which Hare derives from

(*S1*): No imperative conclusions can be validly drawn from a set of premisses which does not contain at least one imperative (*LM*, p. 128).

and

(*C1*): Moral judgements entail at least one imperative.

contains the unsubstantiated premiss *C1*.

If the above suggestion holds and there indeed could be cases where assent to a statement of fact logically commits the speaker to action, there will be uses of 'ought' that fall outside the scope of Hare's tripartite[20]

---

19. Although I am not sure that I have understood MacNiven (1972, p. 177) correctly, his idea that evaluative language has as "its function to convey information of a special kind which is directly relevant to practical problems" seems to fit well into the suggested pattern of analysis. However, when MacNiven claims in one place that "The function of conveying information . . . cannot be used to distinguish evaluative and descriptive language" – the main reason being that "Descriptive language conveys information for a purpose as well", the sense of "action-guiding" used here (viz., "conveying information for a purpose") is not the sense in which Hare uses "action-guiding". Moreover, that evaluative language also conveys information is nothing that Hare denies. Various writers have suggested that 'ought' judgements should be analysed in terms of reasons. For a recent account see Fumerton (1990). See also Taylor (1965, p. 282) who offers an analysis of 'good', to which my own proposal is similar.

20. Nor will such uses necessarily be what Hare calls conventional or ironic uses.

division above in section 1.2. Moreover, such cases will create serious difficulties for Hare's claim that the prescriptive function is a meaning-element in words such as 'ought'. I will follow up this claim in chapter II, where I examine Hare's view on the speech act of telling someone to make something the case.

I will now end this chapter by considering a matter that is closely related to the issue discussed above, viz., Hare's objection to descriptivism and ethical naturalism.

1.6 The 'Open-Question' Argument

Ethical naturalism has been criticized from various directions over the years. The most well-known argument against it is probably G. E. Moore's so-called "open-question" argument. Hare's own central objection to naturalism bears a close resemblance to it, and it is clear that Hare himself regards his argument as a refurbishment of Moore's point (*LM* p. 84). Let me therefore briefly comment on the pros and cons of Moore's argument, in order to examine in section 1.7 whether Hare's version will meet the standard objections put forward against Moore's point.

What is it then about naturalism that both Moore and Hare object to? The fact that the term 'ethical naturalism' is being used in the literature to refer to a great variety of theories, makes it somewhat difficult to answer this question. Moreover, sometimes it is hard to avoid the impression that the controversy between prescriptivists and naturalists contains a considerable element of shadow-boxing - at least on the part of the prescriptivists. It is hard to find anyone who actually embraces the kind of naturalism which, for instance, Hare objects to.

Now, Moore set out to refute those who attempted to give a definition of 'good'. The term 'good' stands, according to Moore, for a simple property (of a special nature) and it is by virtue of this fact that 'good' is indefinable. Those who did not realize this and who defined 'good' were accused of committing the *naturalistic fallacy*. Moore sought to establish this fallacy by means of an alleged *reductio ad absurdum*, viz., his so-called "open-question" argument.[21]

---

21. According to Moore himself, only Sidgwick had anticipated him. However, see Hudson (1978, pp. 72–73) who argues that Richard Price defended a *reductio ad absurdum* which paralleled Moore's own one. Cf. Ibberson (1986, p. 109).

Letting 'C' stand for a natural property, or a set of natural properties, we can use the following as an example:

(1) A good X is an X which is C.

To reduce hypothesis (1) to an absurdity, we have now to ask whether the two following questions are self-answering:

(2) Is an X which is C (always) good?

(3) Is an X which is C (always) C?

Moore's point was the following: Whereas the question in (3) is self-answering, the one in (2) is "intelligible", and this shows, according to Moore, that we have two "notions" in mind when we ask (2), namely goodness and 'C-ness'. 'Good' and 'C' cannot therefore be synonyms.[22]

The "open-question" argument has been criticized by a number of writers. It has been held that Moore neither proved that goodness is distinct from natural properties, nor established that the naturalistic fallacy is a fallacy in the first place (Frankena 1939; Nakhnikian 1963; Smart 1984; Williams 1985; Falk 1986).[23] The standard objection, which I think is valid, has been that the open question is question begging. In reply to the alleged fact that (2) above expresses an open question, naturalists have countered by claiming that it must be shown that it is indeed open – not simply be taken for granted.

As outlined earlier, Moore's open-question argument was construed to show the non-synonymity between moral terms and terms referring to natural properties. However, much of the recent discussion of the open-question argument (and the related naturalistic fallacy), concerns rather whether the argument in addition succeeds in establishing the non-identity

22. This account of Moore's naturalistic fallacy, and the related issue of the open-question argument should be taken for what is is, namely an introductory note to an argument of Hare's. My interpretation is consistent, I believe, with the view which is most often ascribed to Moore. For an interesting account, see Rohatyn (1987), who argues that there is a "plethora of distinct formulations" (*op. cit.*, p. 20) of the naturalistic fallacy.

23. Especially Frankena's paper has had a tremendous influence on this issue. He argues, among other things, that the alleged naturalistic fallacy in fact is a species of a genus, which he coins as the "Definist fallacy". This consists of identifying two properties with each other – a mistake that naturalists, according to Frankena, generally have not committed.

of moral properties and natural properties. It is not entirely clear whether Moore himself aimed at securing the non-identity of properties[24] with this argument. Whatever Moore's view was on this matter, it is safe to say that the test of synonymity in no obvious way establishes the non-identity of properties.

Consider, for instance, the following two predicates: 'Being red' and 'being the colour of hearts in cards'. These are clearly not the same predicates. From the fact that I can wonder whether being red is identical with being the colour of hearts we can reasonably conclude that my notion of being red is different from my notion of the colour of hearts. But from the fact that there are two different notions involved here, it does not follow that being red is not being the colour of hearts in cards.

Proper names and definite descriptions (Russell 1905) are not the only expressions that create problems with regard to this issue. In a much discussed example, involving natural kind terms, Gilbert Harman (1977) objects to the open-question argument that it is in fact logically invalid (Harman 1977; cf. Putnam 1978 and 1981). Harman argues that

> as it stands the open-question argument is invalid. An analogous argument could be used on someone who was ignorant of the chemical composition of water to 'prove' to him that water is not H_2O. This person will agree that it is not an open-question whether water is water but it is an open-question, at least for him, whether water is H_2O. Since this argument would not show that water is not H_2O, the open-question argument in ethics cannot be used as it stands to show that for an act to be an act that ought to be done is not for it to have some natural characteristic C (Harman 1977, p. 19. Cf. Putnam 1981, pp. 46–47, 207–208).[25]

Thus, Harman confronts, in his turn, defenders of the open-question argument with a *reductio ad absurdum* to the effect that one could with a parallel argument refute established results of science.

In the recent literature various naturalists of the objectivist kind have

---

24. Brink (1989, p. 162) is a recent example of a writer who takes Moore as wanting to establish the non-identity of properties: "Moore's open-question argument is designed to show not just failure of synonymy between moral and natural (indeed, nonmoral) terms but failure of identity between moral and natural (indeed, nonmoral) properties". Ibberson (1986) expresses a similar view. On this issue, contrast Ball (1988).

25. Cf. Putnam (1988, pp. 46–47, 207–208). Ball (1988, p. 198) suggests that Moore could deny that 'water = H_2O' is analytic, and claim that it is a synthetic statement which science has proven to be correct. Moore does not deny "the truth of all statements about what activities or experiences *are* good: the claim is only that all such statements are 'synthetic and never analytic'" (cf. Ibberson 1986, chapter IV).

applied the insights of philosophers such as Putnam and Kripke[26] and taken the argument a step further. Thus, it has been argued that proper names, definite descriptions and natural kind terms furnish objectivist naturalism (or a theory akin to it) with an analogical model, by which it could claim that, say, 'Being a good X and being an X which is C is the same property' is true, although it is non-tautological.[27] Such identity-statements express, it is claimed, synthetic propositions which assert necessary identity. However, since Hare's argument, as we shall see, is not concerned with property-identity, there is no need here to follow this discussion here. Let us therefore turn directly to Hare's views.

1.7 Hare's Objection I

The first of Hare's two objections concerning naturalistic definitions of value-terms gets its fullest presentation in *LM*. However, traces of it, and direct references to it are found in several of his later works (e.g. *FR*, p. 84 and *MT*, p. 71). Hare summarizes his objection in the following passage, where later on it becomes clear that 'C' stands for a "conjunction of descriptive predicates" (*LM*, p. 92):

> Now our attack upon naturalistic definitions of 'good' was based upon the fact that if it were true that 'a good A' meant the same as 'an A which is C', then it would be impossible to use the sentence 'An A which is C is good' in order to commend A's which are C; for this sentence would be analytic and equivalent to 'An A which is C is C'. Now it seems clear that we do use

26. Kripke, for instance, claims that 'Water' and 'H_2O' are both rigid designators (Kripke 1980), and that therefore 'Water = H_2O' is an example of "the necessary a posteriori" (*op. cit.* p. 140). For an interesting attempt to show that water is not necessarily H_2O in all logical possible worlds, see Steward (1990). She questions the claim that 'H_2O' is a rigid designator. Hare addresses the "Necessary (Water = H_2O)" example in Hare (1984b). His treatment of this matter seems to me basically sound: In saying that water is H_2O we are either expressing a conceptual truth or we are stating a physical law which is part of physical and chemical theory.

27. For an attempt along these lines see Brink (1989). He says, for instance, "There is nothing metaphysically queer about an ethical naturalist's identification of moral and natural properties. So construed, ethical naturalism requires theoretical identifications of moral and natural properties, similar to other familiar theoretical identifications, for instance, water = H_2O, temperature = mean kinetic molecular energy . . . Hedonistic utilitarianism, for example, can be construed as making such a naturalistic claim; it claims that rightness = the maximization of pleasure" (*op. cit.*, p. 176). Cf. Jarvis Thomson (1990). Relevant is also Ibberson (1987), who argues that naturalists cannot rest content with it being a valid argument in "Chemistry and Optics" that there are necessarily true synthetic statements. Naturalism must show that there is an analogy involved between moral terms and natural kind terms like 'water' or 'temperature'. Ibberson argues from a prescriptivist position that there is no such analogy.

> sentences of the form 'An A which is C is good' in order to commend A's which are C; and that when we do so, we are not doing the same sort of thing as when we say 'A puppy is a young dog'; that is to say, commending is not the same sort of linguistic activity as defining (*LM*, pp. 90–91).

Hare's argument may indeed be regarded as a refurbishment of Moore's open-question argument. At least it appears, in contrast to Moore's argument, to point more directly to why we cannot give a naturalistic definition: The reason is that we then deprive ourselves of the possibility of performing a certain linguistic activity, namely commending (cf. Williams 1985, p. 124).

Now this argument has been examined by various writers, who have not been slow to point out its problematic sides. One complication, which has been focused on by writers such as Blegvad (1959, p. 172), Lange (1966, p. 245) and Sumner (1967), concerns 'C' above. If it stands for any descriptive term(s) it would include, it seems, definitions that could hardly be called naturalistic. Somehow 'C' must be limited to the effect that the definition is about natural properties. However, as Sumner notes:

> Such a restriction is exceedingly hard to come by without circularity, i.e., without defining 'natural characteristic' as one such that a theory that defines 'good A' as 'A having that characteristic' commits the naturalistic fallacy (Sumner 1967, p. 782).

However, Sumner's point will not affect Hare's objection.[28] The reason is that on the most reasonable interpretation of Hare's argument, it will rule out, if valid, not only naturalism but any theory that claims that we are, in issuing moral judgements, performing the kind of speech acts which we perform in issuing descriptive judgements. The argument is directed rather at what Hare later on (as from Hare 1963) calls descriptivism, of which naturalism as well as non-naturalism are examples.

Now, the quote above suggests that what Hare means by naturalism is a second-order theory that contains the claim that moral judgements are

28. Sumner does have a point in that Hare "is not at all *careful* in the way he defines the theory he is attacking" (Sumner 1967, p. 781). Thus, he singles out several passages from *LM* which he thinks could be treated as definitions of naturalism. However, strangely enough he does not mention the passage at the end (*LM*, pp. 92, 93) where it becomes clear that Hare is criticizing attempts to equate value-terms with descriptive predicates. Moreover, it should be pointed out that Hare himself in fact comments that 'naturalist' is an "unfortunate" term (*LM*, p. 82) in that other ethical theories commit the same fallacy as "the group of ethical theories which Professor Moore called 'naturalist'".

equivalent in meaning to descriptive statements. A further quote from *FR* will be helpful here:

> for naturalists hold that the rules which determine to what we can apply value-words are simply descriptive meaning-rules, and that these rules determine the meaning of these words completely, just as in the case of descriptive expressions. For [a naturalist], a value-word is just one kind of descriptive expression (*FR*, p. 16).

Whether anyone actually has endorsed such a theory is a matter that will not occupy me here.[29]

Hare's argument involves the following steps:

*A*1: If 'a good A' means the same as 'an A which is C', then the sentence 'An A which is C is good' is equivalent in meaning to 'An A which is C is C'.

*A*2: If *A*1 is correct, then it is impossible to commend A's which are C by using the sentence 'An A which is C is good'.

*A*3: We use the sentence 'An A which is C is good' to commend A's which are C.

B: 'A good A' does not mean the same as 'An A which is C'

Thus, this argument is directed against a claim about what one linguistic expression *means*.[30] Accordingly, it avoids the criticism put forward against Moore's open question - correctly or not - that this objection cannot show that certain moral properties are not identical with natural or non-moral

---

29. Cf. Kurtz (1970) who wonders whether Hare's naturalist opponents are not in effect "straw men". See also Lange (1966, p. 245).

30. Daniels (1970) objects that from the fact that we cannot perform an act with a sentence which we can perform with another sentence, nothing follows with regard to whether the two sentences are synonymous. We can do the act of defining with "A puppy is a young dog" which we cannot do with "A puppy is a puppy". But, Daniels argues, these sentences mean the same. However, I suspect that Hare would say "defining" and whatever act we do with the latter are species of the speech act of asserting. To have a point Daniels would have to show that two sentences belonging to two different genera of speech acts, would still mean the same thing. See also Rynin (1962) who tries to outline synonymity in such a way that it would cover such cases. However, his attempt rests on a view of "emotive and evaluative meaning" which is at variance with Hare's. See e.g. p. 240, where it becomes clear that an evaluative word is a word which produces persuasive-emotive effects. This goes against Hare's view that the evaluative meaning of a word is located in what we *do* in uttering a sentence containing it. More about this in chapter II.

properties. As the passages presented above evince there is no talk on Hare's part of property identity. The issue is clearly whether a certain sentence can be equivalent in meaning with another sentence.

As it stands, Hare's argument requires at least one further additional premiss. Sumner suggests the following

> For any two sentences, p and q, p means the same as q only if any linguistic function performable by using p is performable by using q, and vice versa (Sumner 1967, p. 785).

The expression "linguistic function" is problematic. It is questionable, for instance, whether Sumner's use of it is consistent with Hare's way of understanding it.[31] However, we have not yet examined Hare's speech act theory of meaning - a subject matter which will concern us in chapter II. I will therefore here grant Hare something to the effect of Sumner's premiss.[32]

Having granted this, the argument seems cogent enough. There is a ring of conclusiveness about it that reminds one of Moore's open-question argument. The argument appears to be a direct *modus tollens*; if $A1$ is true, we cannot commend by using a certain sentence; but, as a matter of fact, we do commend by using this sentence, and hence the falsity of $A1$ follows.

The argument, as may be recalled, is said to "restate Moore's argument in a way which makes it clear why 'naturalism' is untenable" (*LM* p. 84). Hare also maintains that naturalism involves a "fallacy" (*LM* p. 92). What he means by this, however, is not entirely clear, but it seems reasonable to take him as maintaining nothing more than that naturalists, by making moral judgements synonymous with descriptive judgements, commit the "fallacy" of overlooking that we then are no longer able to commend or prescribe by issuing moral judgements. Germane to the argument is also Hare's contention that similar arguments could be put forward to show that

---

31. Sumner erroneously thinks that Hare is using a special sense of 'commending' which departs from what Austin called "illocutionary acts" (Sumner 1967, p. 789). Sumner does not give any explanation for why he thinks so. However, he does offer himself an example that I think suggests that he has overlooked a distinction recognized by Hare, Austin and several other speech act theorists, viz., the one between a judgement's perlocutionary and illocutionary force (see p. 788).

32. A more accurate statement of the missing premiss would replace the expression "linguistic function" with 'only if p and q have the same generical illocutionary act potential'. However, to do so here presupposes an acquaintance with the special speech act terminology, the presentation of which will be postponed until chapter II.

naturalistic definitions of other moral terms, such as 'ought' and 'right' involve the same fallacy (*LM*, p. 171).

On closer examination the argument is less convincing. What does Hare mean by "commending"? And why is it that 'An A which is C is C' cannot be used to commend?

The explanation which Hare gives of "commending" corresponds to the one he gives of "prescribing", with the exception that he prefers to characterize what we are doing in issuing evaluative 'good'-judgements as commending, whereas he speaks of prescribing with regard to 'ought'. Hare has over the years referred to the action-guiding genus of speech acts in different ways. Most often he calls it the prescriptive genus (e.g. Hare 1968a). However, in *LM* (see e.g. p. 4) he uses a somewhat different terminology. There he differentiates between two kinds of sentences, viz., imperatives and indicatives. Moreover, he then distinguishes between statements, which is what we express by using an indicative sentence, and commands, which is what we express by using sentences in the imperative mood.[33]

Now Hare explains commending in the following way:

> When we commend or condemn anything, it is always in order, at least indirectly, to guide choices, our own or other people's, now or in the future (*LM*, p. 127).

We may therefore reasonably replace *A*3 with the following:

*A*3' We use the sentence 'An A which is C is good' in order to guide the actions/choice of the listener.

And hence, *A*2 may be reformulated without, I believe, any substantial violations, to the following effect:

*A*2' If *A*1 is correct, then it is impossible to use the sentence 'An A which is C is good' to guide the actions/choices of the listener.

Why does Hare then support (A2/A2')? His argument for it is, as the following passages will reveal, intimately connected to the argument I examined above in sections 1.3-1.5:

---

33. In Hare (1968b) he gives the action-guiding use a third name, viz., "Imperation". Although it is not all together clear from *LM*, I think it is reasonable to say that commending is a species belonging to the genus of action-guiding speech acts.

> what is wrong with naturalist theories is that they leave out the prescriptive or commendatory element in value-judgements, by seeking to make them derivable from *statements* of fact (*LM*, p. 82. My italics).

> For all the words discussed [i.e. 'good', 'right' and 'ought'] have it as their distinctive function either to commend or in some other way to guide choices or actions; and it is this essential feature which defies any analysis in purely factual terms. But to guide choices or actions, a moral judgement has to be such that if a person assents to it, he must assent to some imperative sentence derivable from it (*LM*, p. 171).

Another passage reveals the underlying reason. Speaking of the sentence 'I ought to do X' that is used in a prescriptive way, and which is referred to as sentence "(3)" in the quotation below, he maintains the following:

> The fact that it is not possible [to effect a completely naturalistic analysis of all uses of 'ought', and thus of 'good'] is entirely due to the intractably evaluative character of (3). It is due ultimately to the impossibility [. . .] of deriving imperatives from indicatives; for (3), by definition, entails at least one imperative; but if (3) were analysable naturalistically, this would mean that it was equivalent to a series of indicative sentences; and this would constitute a breach of the principle established. Thus it is this fact, that in some of its uses 'ought' is used evaluatively (i.e. as entailing at least one imperative) that makes a naturalistic analysis impossible (*LM* p. 171).

These passages suggest how Hare would justify A2'. What is wrong with naturalistic definitions is that they make moral judgements equivalent to a series of indicative sentences, and this is a fallacy, since there are uses of moral judgements that must be analysed as entailing imperatives.

We may restate the premisses of the argument in the following way:

*A*4: If 'a good A' means 'an A which is C' then 'An A which is C is good' is equivalent in meaning to the indicative sentence 'An A which is C is C'.

*A*5: No indicative sentence can be used to guide the actions/choice of the listener.

*A*3': The sentence 'An A which is C is good' is used to guide the actions/choice of the listener.

From *A*5 and *A*3', Hare then concludes that the sentence 'An A which is C is good' cannot be equivalent in meaning to an indicative sentence.

The argument could be cast in more ways. Thus *A*4 as well as *A*5 could be formulated so as to focus on the descriptive meaning of the sentence 'An A which is C is C' or on what we do in using this sentence, namely to state;

and *A*3' could include a reference to the fact that Hare thinks that action-guiding uses entail imperatives (cf. McGuire 1961). However, to do so would not add anything pivotal to the argument. The presence of (*A*3/3') (or some variant of it in terms of the above mentioned features) in the above arguments brings out what must be, for Hare, a disturbing feature, namely that the argument seems plainly question begging (cf. McGuire 1961; Sumner 1967; Walker 1973; Franklin 1973).[34] The naturalist's prompt reply would be to deny that to utter 'An A which is C is good' is to commend, guide actions. Hare, the naturalist may claim, assumes as a premiss the very conclusion he tries to establish, viz., that moral judgements are not equal in meaning to descriptive statements.

Noteworthy is also another feature of the argument, namely, that this objection to naturalism equally applies to any theory that claims that what we are doing in issuing moral judgements does not require that we analyse these as entailing imperatives. Thus, if what is wrong with naturalistic definitions is that we are stating something when we are saying 'This A which is C is good', then if we think, as Hare does, that using the sentence 'This A is C' or 'A which is C' is to state something, we will not be able to commend A's by using these latter sentences either.[35]

Now, as was argued in section 1.6, it is possible to detect two reasons in Hare's work for his claim in (*A*2/*A*2'). In this chapter I have focused on his internalist argument:[36] To be action-guiding "a moral judgement has to be such that if a person assents to it, he must assent to some imperative sentence derivable from it" (*LM*, p. 171). This claim, I maintained, relies to begin with on the linguistic intuition that assent to a moral judgement commits one logically, and not merely causally, to action. Relying on his

34. Walker's objections to Hare's view rest in part on interpretations which I think it is unclear whether Hare endorses. Walker takes it for granted, for instance, that Hare would assent to 'This is a good F' having "truth-conditions" (*loc. cit.*, p. 48). Franklin's article contains a good survey of the debate between naturalists and non-naturalists.

35. Sumner notices this (1967, p. 791). However, he construes the underlying principle in a way that I think Hare would not accept: "For any two words, a and b, a means the same as b only if any linguistic function performable by using a is performable by using b, and vice versa".

36. That this is the major reason why he criticizes naturalism, gets some support from a passage in *MT*: "There is, however, another reason why even somebody who has confined his moral thinking to the intuitive level would reject naturalism [. . .]. This is that moral words have, even at that level, a commendatory or condemnatory or in general prescriptive force which ordinary descriptive words lack (*LM*, 5.4). The person who thinks that the fact that an act would be wrong is no reason for not doing it shows thereby that he has not fully grasped the meaning of the word" (*MT*, p. 71).

linguistic intuition, Hare then proceeds, as shown above, to argue that to account for this internalist feature we must understand moral judgements as entailing at least one imperative. However, this argument was questionable, I argued, since there seem to be purely descriptive judgements, the assent to which would - given that the judgement was correct - logically and not merely causally commit us to action. Only those descriptivist theories would run into difficulties that combine an internalist view with the claim that moral judgements are equivalent in meaning to descriptive statements of states of affairs *other* than states of mind, dispositions - what Hare called objectivist descriptivism/naturalism (see section 1.1). Such a combination would be, to draw on Mackie's expression, queer. On this matter I am in no disagreement with Hare. However, I suspect that such forms of descriptivism are seldom internalist.[37]

Hare's prescriptivism need not rely on it being false that the internalist feature of moral language can be made consistent with a purely descriptive analysis of moral judgements. There is another side to his prescriptivism, namely, the claim that the speaker is telling the listener to make something the case, instead of telling him that something is the case, when he issues a moral judgement. I will round off this chapter by examining an argument that Hare thinks supports his claim that such uses exist.

1.8 Hare's Objection II

The second argument of Hare's that I wish to call attention to is found in his article "Descriptivism" (Hare 1963b). Here he sets out to criticize theories that commit what Hare calls, following the late philosopher of language J. L. Austin, the "Descriptive Fallacy".[38] This fallacy consists in

---

37. Bergström (1990, p. 83) is right in maintaining that "A value-realist does not have to be an internalist. On the contrary, the most natural thing for a value-realist is surely to be an externalist". I have translated the text which is in Swedish: "En värderealist behöver inte vara internalist. Tvärtom är det väl mest naturligt för en värderealist att vara externalist". A caveat: value-realists although not necessarily claiming anything about the meaning of words, will be subject to Hare's criticism in so far they make moral judgements, which they reasonably must conceive of as statements.

38. Hare's criticism of naturalism/descriptivism has given rise to an impressive amount of articles that try to meet the objections. A good survey of the debate is found in Hudson (1970 and more recently in 1988). Beardsmore (1969) outlines the general arguments between Hare and one of his stronger opponents, Mrs. Foot, the relevant articles of whom are found in Foot (1977); Carter (1982) contains a useful bibliography (p. 107, note 20). See also Franklin (1973).

"supposing that some utterance is descriptive when it is not" (*MC*, p. 55).[39] The argument which Hare unfolds in this article aims at securing the distinction between the evaluative and descriptive meaning of a term. Thus, he says

> We can show that such a distinction exists, at any rate, if we can isolate one of these two sorts of meaning in a given context, and show that it does not exhaust the meaning of the term in that context. Suppose, for example, that we can show that in a certain context a term has descriptive meaning; and suppose that we can isolate this descriptive meaning by producing another term which could be used in the same context with the same descriptive meaning, but such that the two terms differ in that one has evaluative meaning and the other not; then we shall have established that there can be these two different components in a term's meaning (*MC*, p. 56).

The example by which Hare hopes to show that there is such a distinction to draw is the following:

> Let us suppose that somebody says that a certain wine (let us call it 'Colombey-les-deux-églises 1972) is a good wine. I think it will be obvious that he says that it is a good wine *because* it has a certain taste, bouquet, body, strength, etc. (I shall say 'taste' for short) (*MC*, p. 56/57).

Hare then lets 'ø' stand for this complex quality, which makes us call it a good wine. Hare further assumes that we can or one day will be able to isolate and manufacture the taste at issue to the effect that by pouring ø into wine lacking it, people would no longer be able to distinguish such a wine from wines having ø by natural procedures. Having set the example in such a way, Hare maintains that he is now in a position to teach a person what ø stands for. He would, he says, proceed by

> lining up samples of liquids tasting ø, and others having different tastes, and getting him to taste them, telling him in each case whether the sample tasted ø or not. It is worth noticing that I could do this whether or not he was himself disposed to think that these liquids tasted good, or that, if they were wines, they were good wines. He could, that is to say, learn the meaning of 'ø' quite independently of his own estimation of the merit of wines having that taste (*MC*, p. 58).

A passage further on, Hare maintains

39. In Hare (*ET*, p. 120, 1986) he is more explicit about the nature of this fallacy: "Descriptivism (or the descriptive fallacy as Austin called it . . .) is the belief that all words get their meaning in the same way as descriptive words and statements do, by having application-conditions or truth-conditions".

> if the explanation [of 'ø'] works, then it is possible to explain the meaning of 'ø' ostensively as a descriptive expression; and when this has been done, we can separate out the descriptive from the evaluative meaning of the expression 'good wine' in the sentence 'Colombey 1972 is a good wine'. For it will be possible for two people to agree that Colombey 1972 tastes ø, but disagree about whether it is a good wine; and this shows (which is all that I am trying to show) that there is more to the statement that Colombey 1972 is a good wine than to the statement that Colombey 1972 is a wine which tastes ø. The more is, of course, the commendation; but I shall not in this lecture try to explain what this is, since I have done my best elsewhere (*ibid.*, pp. 58–59).

The line above stating "The more is, of course, the commendation" is puzzling. Are we to understand Hare as implying that the argument somehow makes it self-evident that this "more" is commendation? Or does Hare only wish to say that it is clear from his other writings what exactly his position is on this issue, namely that we (for reasons given elsewhere) are commending when we say that a wine is good? The latter interpretation seems the more reasonable. It can hardly be said to be clear how the above argument shows that we are commending.

Let us grant Hare that it is possible for us to pick out what the descriptivist means by 'ø'. We may then restate the argument in the following way:

*D*1: If 'X is a good wine' means 'X is a wine which has ø' then (two) persons who agree that X has ø are logically committed to say that X is a good wine.

*D*2: But (two) persons can agree that X has ø and still disagree whether X is good or bad.

The conclusion which Hare draws is that

*D*3: 'X is a good wine' does not mean 'X is a wine which has ø'.

Since *D*2 entails the falsity of the consequence of the conditional in *D*1, *D*3 is arrived at by *modus tollens*. From the former of the above passages, it is clear that Hare wants to show something stronger than this, namely, that there exists the distinction between two kinds of meaning – the descriptive and the evaluative. That is, he wants us to conclude at least the following:[40]

40. We must here assume that Hare has in mind the primary meaning of 'good wine' and not any inverted commas use.

*D*4: The descriptive meaning of 'good wine' can be separated from its evaluative meaning.

Even if we grant Hare that an expression such as 'good wine' may only have two kinds of meaning - descriptive and/or prescriptive - it is unclear just how he thinks *D*4 is arrived at.

The argument above carries considerable strength against at least one kind of descriptivism, namely the kind of theory which Hare in more recent works has called objectivist descriptivism. What diminishes the strength of it is, of course, the fact that 'ø' refers to objective properties. Hare has not here considered the possibility that it has to do with the attitudes of the speaker. The argument above provides no conclusive refutation of descriptivism; it gives us some, I would say, strong evidence that we have not made all there is concerning the meaning of 'This wine is good' explicit by making its descriptive meaning clear.

However, it is not clear just what further conclusions Hare wishes to derive from the above example. The following argument immediately suggests itself, where 'C', as before, stands for some descriptive predicate, and where 'X' refers to an act, thing, or person:

*D*1': If 'This A is good' means 'This A is C', then two persons must agree that A is good, if they agree that A is C'.

*D*2': Two persons can agree that A is C and still disagree whether X is good.

And hence, by *modus tollens*, follows

*D*3': 'This A is good' does not mean 'This A is C'.

But in order to validate the more general conclusion that any naturalistic definition will be mistaken, it is not sufficient to show that some naturalistic definition is mistaken. However, it is unclear whether Hare's argument actually is designed to have this function. It may be best understood as being a potent tool, with which we may examine proposed naturalist definitions. As such it serves (as does Moore's open-question) its purposes.

What should we then say about *D*4? Is this the conclusion we must arrive at in in order to account for the fact that two persons can agree that some wine has ø and still, without committing a logical mistake, disagree about whether it is good or bad?

Descriptivists in general, and naturalists (of a Harean cut) above all,

argue that words such as 'good' and 'ought' have at least some fixed descriptive meaning, whereas prescriptivists maintain that the descriptive meaning can vary from time to time, from person to person. I believe there is strong evidence against the former position, and that there is much truth in Hare's position. However, even supposing that there was stronger evidence telling against the naturalists on this horn of the dilemma, such a belief does not commit me to *D*4, and even less to the conclusion that Hare wants to draw in the prolongation of it, namely that the two kinds of meaning are the descriptive and prescriptive meaning.

There is an alternative way of understanding 'X is a good wine' that bears in mind that people, at least seemingly, often do disagree about what is good or morally obligatory, although they agree on what are the facts. Consider once more the above example:

(s) 'Colombey 1972 is a good wine'

Thus, suppose we do believe that two persons can agree on what are the intrinsic properties of the wine, and still disagree whether the wine is good or not. In such a case we have some reason for considering (s) as being elliptic for:

(s') Colombey 1972 is a good wine for me.

(s') opens for the possibility of being reduced to a descriptive statement, without losing sight of the fact that two persons can agree about what are the factual properties of the wine, and still not both be committed to calling it good. Consider for instance

(s") Colombey 1972 is a wine desired by me.

I am not suggesting that (s) and (s") are synonyms, nor am I here taking a stand as to whether (s") is a plausible descriptive reduction of (s). However, (s"), or at least (s') deserves to be mentioned since it is an example of a descriptive analysis that stresses that people evaluate factual properties differently without committing any logical mistakes, which at the same time avoids Hare's objection above.

I suspect that some readers, including Hare, will take up a sceptical attitude towards what will be considered as a straightforward relativistic analysis of value-terms. However, it must be kept in mind that relativism in

second-order moral philosophy can refer to various theories (cf. Bergström 1990, which contains a list of different types).[41] From the claim that value - words implicitly are connected to the speaker's motivational state it does not follow, for instance, that moral judgements cannot be open to rational discussion. As we shall see in chapter IV, Hare's own view on this matter strongly suggests that we may speak of rational preferences and desires. However, this is not the place to continue this discussion. I limit myself here to having offered an example of a descriptive analysis of a judgement that goes free of the argument, by which Hare wanted to show that there exists two kinds of meaning.

My own limited experience of what people mean by 'good', 'ought' and 'right' has led me to believe that there is not one meaning but several involved, and that any attempt to catch *the* meaning in a single *definiens* is going to fail. Perhaps we should lower our ambitions and search instead for "the lowest common denominator". Hare has found such a denominator in what he considers is the typical speech act involved in issuing evaluative judgements. There are perhaps additional reasons for why we should choose the prescriptive speech act as being logically tied to evaluative judgements?

41. Thus, Hare (1986) objects to MacIntyre's relativism in a convincing way. However, it must be noticed that the kind of relativism which he attacks there is precisely that, viz., one kind of relativism, to which someone who maintains that value-judgements are elliptical in the suggested way is not committed. See Bergström (1990), especially what he calls "semantical relativism". Relevant here is also *MT* (pp. 76, 77), where Hare states his objection against what he calls "old-fashioned subjectivism", namely, that such a view makes moral disagreements impossible.

II. Prescriptivism Developed and Challenged

2.0 Introduction

Having focused in chapter I on what I called the internalist side of Hare's prescriptivism, I will in this chapter follow up the discussion by saying more about what, according to Hare, the speech act of prescribing is. I believe there are difficulties with such a view, which make it implausible as a view of the meaning of moral terms. However, to argue for this requires that we deepen our understanding of what Hare means by saying that moral terms typically have prescriptive meaning. We need, in other words, to examine Hare's underlying theory of meaning.

The philosophical outlook in which Hare develops his view on the meaning of moral terms, falls under what is often referred to as the speech-act theory of meaning. The foundation of this branch of philosophy of language was laid by the late J. L. Austin, whose many ideas and distinctions, found notably in the posthumous work *How to do things with words* (1962), have later on been further developed by several other writers. Since Hare in his account of the meaning of value-terms to a great extent relies on Austin's work, I will begin by outlining in section 2.1 those of Austin's distinctions which Hare has either applied or discussed. In section 2.2 I present an interpretation of Hare's use of 'sentence', 'utterance', and 'judgement'. I examine an argument in 2.3 put forward most recently by the speech-act philsopher John R. Searle against Hare's claim that moral words have prescriptive meaning. Searle also directs another argument against Hare's views. In section 2.4 I look into this objection, arguing that neither of Searle's objections establish that value-terms do not have prescriptive meaning in Hare's sense. In section 2.5 I comment briefly on Hare's claim that value-terms can be used in a non-prescriptive way. A more detailed account is given in 2.6 of what, according to Hare, is involved in performing a prescriptive speech act. To obtain a proper picture of what Hare means by prescribing requires an understanding of what he means by desire and intention. In 2.7 I present my interpretation of his notion of desire and intention. Section 2.8 is mainly earmarked for an argument by notably John Ibberson against Hare's views on past imperatives and past prescriptions.

Moreover, I suggest that there are examples other than past 'ought'-judgements that are hard to understand as prescriptions. Focusing mainly on moral 'ought'-judgements, I then go on to argue in 2.9 that prescriptivism - *qua* speech-act thesis about the meaning of the moral term 'ought' - is untenable. My reason will be that first- and third-person 'ought'-judgements cannot reasonably be considered as prescriptions. I believe this will be evident once we distinguish between two separate, albeit closely related, theses about what is involved, according to Hare, in performing a sincere prescriptive speech act. The first of these claims concerns the speech act as an intentional act; the second is a thesis about the volitional state in which a speaker must be, in order to have performed the speech act *sincerely*. In 2.9.1 I say a few words about a possible way for Hare to meet the objections which have been presented. Finally, I round off the chapter by briefly commenting in 2.9.2 on Hare's work concerning imperative inferences which I have chosen not to examine further in this essay.

2.1 J. L. Austin on Speech Acts

Speech-act philosophers attempt to explain, at least partly, what bits of language mean in terms of what we use them to do.[1] Austin (1962) focused especially on three different kinds of speech acts, which he called the *locutionary*, *illocutionary* and *perlocutionary* act. The locutionary act is the act we perform, he explains, when uttering a sentence "with a certain sense and reference, which again is roughly equivalent to 'meaning' in the traditional sense" (Austin 1962, p. 109). Austin further distinguishes within the locutionary act three other kinds of acts. There is the "phonetic" act of uttering some noises, the "phatic" act of uttering noises which are words, and are produced as words belonging to a certain language, with a certain grammar and a certain intonation, etc., and, finally, there is the "rhetic" act of using words with a certain "more-or-less definite sense and reference" (Austin 1962, p. 95).

The locutionary act must in its turn be separated from the illocutionary

1. Hare has never held that speech-act theories contain all the answers to the entangled matter of meaning, but only that a study of what we do in uttering linguistic elements *contributes* to our understanding of these. I share this belief. Still awaiting us is, I think, a proper evaluation of the pros and cons of such theories in comparison with other theories of language - especially the Fregean tradition that understands meaning in terms of truth conditions. Eventually philosophy of language will witness the rise of a third alternative or the creation of a common ground between these traditions. Until then we unfortunately face a cleavage of opinion.

act, which is the act performed *in* saying something. Consider the following sample of sentences:

(a) You are going to shut the door.

(b) Shut the door!

(c) Are you going to shut the door?

(d) I promise to shut the door.

Intuitively we seem to be doing different things in uttering these sentences, namely (a') stating, (b') commanding, (c') asking a question, and (d') promising. These are all examples of what Austin, and following him, Hare, call *illocutionary acts*. There are other kinds of illocutionary acts, and there are other things we do in performing speech acts that are not characterizable as illocutionary acts.

To the division between locutionary and illocutionary acts, Austin added the *perlocutionary act*. Normally, saying something will also, Austin argued, "produce certain consequential effects upon the feelings, thoughts, or actions of the audience, or of the speaker, or of other persons; and it may be done with the design, intention, or purpose of producing them" (Austin 1962, p. 101). Suppose a chemist, for instance, wants to inform a colleague about the content of a certain glass. He may utter, say, 'This glass contains HCl'. In saying this, the chemist has, *ex hypothesi*, performed the illocutionary act of telling that something is the case. Suppose now that the situation is different. Our chemist is a villain who is interrogating a prisoner. Holding a glass over the head of the prisoner, he utters 'This glass contains HCl'. If by saying this he made the prisoner afraid, terrified, etc., he will have performed the intentional perlocutionary act of frightening, alarming, etc., the prisoner. Hare's distinction between "telling someone to[2] make something the case" and "getting him to do it"corresponds to Austin's distinction between illocutionary acts and perlocutionary acts.[3]

2. Recently, I suspect that William Charlton fails to pay attention to the distinction at issue, when he apparantly brings "to guide" into line with perlocutionary acts such as "to rouse or quiet an emotion, to make something clear or to confuse someone" (Charlton 1988, p. 84; cf. also pp. 85–87).

3. Hare says "On one of the very few occasions on which [Austin] read a paper in public about this distinction. . ., Austin was kind enough to say that he was saying the same sort of thing as I had been saying" (*PI*, p. 53). A recent argument for the separation at issue is found in Hare (1987b, p. 74). For a list of writers who have overlooked this distinction when criticizing Hare's views, see Carter (1982, pp. 35, 36)

Austin's work gives rise to several interesting issues, and various other writers, such as Furberg (1963), Strawson (1964), Searle (1969, 1979), Schiffer (1972) and Wetterström (1977) have tried their hands at carrying Austin's analyses further. One such issue that has attracted much attention concerns Austin's claim that the *meaning* of an utterance is distinct from its *force*. Austin never specified exactly what he had in mind by 'meaning'. However, some passages strongly suggest that he was thinking of the Fregean distinction between sense and reference (see especially p. 100). By the "force" of an utterance, on the other hand, Austin referred to the illocutionary function of a word or utterance. An utterance such as 'Open the door' would on such an account have the illocutionary force of a command, request, order, etc. What various writers objected to was that Austin maintained that it is the locutionary act that has a meaning, by virtue of being the act of uttering a sentence with a meaning. Writers such as Cohen (1964), Searle (1968), and Hare (1971c), for instance, have argued that this division of force and meaning is untenable or at least not as sharp as Austin appears to have thought. Cohen rejected Austin's notion of illocutionary act, while Searle and Hare gave up the idea of a separate rhetic/locutionary act.[4]

Hare's reasons for abandoning a sharp line between the locutionary and illocutionary act is found notably in "Austin's Distinction between Locutionary and Illocutionary Acts" (Hare 1971c). Thus, he says:

> If to perform a locutionary act is to utter words with a certain sense and reference, but not, *qua* locutionary act, with any illocutionary force, what is going to count as a locutionary act? Suppose I say 'The cat is on the mat'. Here the expression 'the cat' and 'the mat' have presumably, a sense and a reference. What about the word 'on'? That, we may allow, has a sense but no reference. But what about the word 'is'? Here we seem to be faced with a dilemma. It certainly has no reference, except, of course, the temporal reference to the present time. Has it a sense? Well, of course it has a meaning; but has it a sense in the narrow interpretation of that word as, apparently, used by Austin, which excludes illocutionary force? (*PI*, pp. 106–107, 1971c).

Hare suggests that the meaning of 'is' is complex. It contains as part of its sense the notion of predication - "the notion of something *being* something". However, in addition to this,

---

4. Both Cohen's and Searle's treatment of this issue have been criticized by Thau (1972). For a recent contribution to this matter, see Hornsby (1988 p. 286, n. 17) who suggests that Austin and Searle may have used "locutionary" ambiguously.

> it also has, as part of its sense, the notion of assertion – what tells us that [the utterance of 'The cat is on the mat'] is a statement and not a command or a question. The dilemma is this: Is this part of the sense, as Austin was using the word? If it is, then Austin will have to admit – as I suspect that in fact he would have admitted – that the sense, at least sometimes, and indeed so far as I can see, always, includes part, at any rate, of the illocutionary force (*PI*, p. 107).

Hare's point is that the locutionary act would have to include, in the sense, a specification of whether the utterance is a statement or command. But these are expressions referring to illocutionary acts. And hence, the performance of the locutionary act will specify at least partly what illocutionary force the act carries:

> It is not merely that all locutionary acts are also illocutionary acts; it is that being *that* locutionary act – which involves being a use of words with *that* sense – makes the act, already, into an act of carrying a certain illocutionary force (*PI*, p. 108).

Hare dismisses the alternative - that it is not part at all of the sense of any of the words uttered in a locutionary act to specify whether the utterance is, e.g., a question or statement. What would saying 'The cat is on the mat' amount to if what we performed was only the act of uttering words with a certain sense and reference, which did not tell us whether what we said was a statement, request, etc.? It would not, Hare maintains, make any sense (*PI*, pp. 108–109).

Closely related to the matter of how to understand Austin's dichotomy between locutionary and illocutionary acts lies a fundamental issue that has divided "speech act" philosophers of language into two broadly camps, namely, whether or not illocutionary acts are conventional acts. Austin himself maintained that such acts as opposed to perlocutionary acts are conventional acts. Strawson and Schiffer, for instance, maintain on the other hand that apart from some peripheral cases, illocutionary acts are not conventional, while, e.g., Searle, Hare, and Wetterström have suggested that all or at least nearly all of them are conventional. Searle has offered an explanation of what he means by illocutionary acts being conventional in terms of the distinction between institutional facts and brute facts:

> [institutional facts] are indeed facts; but their existence, unlike the existence of brute facts, presupposes the existence of certain human institutions. It is only given the institution of marriage that certain forms of behavior constitute Mr Smith's marrying Miss Jones. Similarly, it is only given the institution of baseball that certain movements by certain men constitute the

> Dodger's beating the Giants 3 to 2 in eleven innings. And, at an even simpler level, it is only given the institution of money that I now have a five dollar bill in my hand. Take away the institution and all I have is a piece of paper with various gray and green markings (Searle 1969, p. 51. Cf. Wetterström 1977, ff. 8).

It would be a brute fact that I weigh so-and-so many pounds, whereas it would be an institutional fact that by uttering 'I weigh so-and-so many pounds' I stated[5] something.

Writers who deny that illocutionary acts are conventional have often followed another influential writer in this field, viz., H. P. Grice, who has developed an account of meaning in terms of the effect that the speaker intends the utterance to have on the listener. Grice's analysis has grown very complex, and there is no need here to go further into detail. Searle (1969) and Wetterström (1977) have produced convincing counter examples to Gricean analyses in terms of the perlocutionary effects utterances have on the audience (see also Carr 1977).

2.2 The 'Sentence'/'Utterance'/'Judgement' Distinction

Let us now look more in detail into Hare's view on the meaning of evaluative words. Hare has referred to meaning as a property not only of words and sentences, but also of utterances and judgements. He does not provide us with a systematic account of how he uses these terms. This is unfortunate, since common usage is not very clear.[6] Should we, for instance, limit 'utterance' only to spoken language or should we, as I think is reasonable, include even written occurrences among utterances? Are bodily movements excluded from being utterances? Suppose we nod our heads instead of uttering 'yes' when asked if we want a cup of coffee. Is our head

---

5. Searle's much discussed attempt to logically bridge the so called "is/ought" gap, in Searle (1964), centres to a great extent on the idea that stating about a man that he is promising, is stating an institutional fact, which in effect is, Searle claims, to invoke certain constitutive rules of the instituion. Those rules, in their turn, which give the meaning to 'promise', are such that to "commit myself to the view that Jones made a promise involves committing myself to what he ought to do (other things being equal)" (quote from Hudson 1969, p. 133). Hare's criticism of Searle is found in "The Promising Game" republished in *ET*, for instance. See also *HC*, p. 215, note 19/–8. Relevant is also McNeilly (1972) who examines both Searle's and Hare's views on whether we are under an obligation when having said 'I promise'.

6. Although the notions of speech act and illocutionary force play an important role in Hare's metaethical theory, he has chosen to a large extent to keep out the idiom of speech act theory from his major works *LM*, *FR*, and *MT*. *PI* contains a collection of his most important articles on these issues.

movement not an utterance? Again, is it not too restrictive to apply 'sentence' only to written linguistic entities? Since I have found no passages that indicate the contrary, I will interpret Hare as maintaining that the dividing line between utterances and sentences is not the one we find between language expressions realized as sounds and marks on a paper.

Additional problems arise when we take into consideration that 'sentence' and 'utterance' suffer from a type/token ambiguity.[7] Most often this ambiguity is no source of confusion. This also holds for another possible ambiguity that is perhaps especially associated with 'utterance', viz., that we can use 'utterance' in an "act"-sense as well as in an "object"-sense. Austin, for instance, distinguished between "*utteratum*" and "*utteratio*", using 'utterance' as equivalent to the former, whereas "the issuing of an utterance" carried the sense of the latter (Austin 1962, p. 92; cf. Wetterström 1977, pp. 36-39). Thus we may in speaking of someone's utterance refer to what he was doing, e.g. stating, requesting, commanding, etc. But 'utterance' may be tantamount to the result of a speech act or utterance act (that which is produced in stating or requesting, namely, a statement or a request).

Let us focus attention on Hare's use of 'sentence'. He seems to have somewhat modified his view with regard to this word. Thus, in *LM* and *FR*, he explains sentence-meaning, at least in part, in terms of actual prescriptions, that is, with reference to the speech act performed in using the sentence. In *LM*, for instance, Hare says the following:

> 'Sentence' here is an abbreviation for 'sentence as used by a particular speaker on a particular occasion'; for speakers may on different occasions use words with different meanings, and this means that what is entailed by what they say will also differ (*LM*, p. 25.).[8]

In a later article, "Meaning and Speech Acts", Hare follows William P. Alston in explaining sentence-meaning in terms of illocutionary act *potential*. Alston does not specify in detail just what he means by a sentence

---

7. The pupil who has written one hundred times 'It was wrong of me to use my teacher's car as a canvas for my graffiti painting', will not be pleased to hear that he now must write a second sentence, unless perhaps he realizes that his teacher is speaking of sentence types and not of sentence tokens.

8. Hare has made the following comment about *LM*: "In *LM* I used the slovenly pre-Strawsonian terminology (still surprisingly current) which did not distinguish between sentences and statements or commands. Strictly sentences do not entail one another" (Hare 1979a, p. 162).

having an illocutionary act potential. He seems to have used it in the following way: To say that a certain sentence has the illocutionary act potential *x*, is to say that it is commonly used to perform the illocutionary act *x* (Alston 1964, see e.g. p. 36). Hare's use of the expression is brought out in the following passage:

> the meaning of a certain word can be explained, or partly explained, by saying that, when incorporated in an appropriate sentence in an appropriate place, it gives to that whole sentence the property that an utterance of it would be, in the appropriate context, a performance of a certain kind of speech act. This is the same as to say that an utterance of it would have, in Austin's term, a certain illocutionary force; and it is the same as to say, with Professor Alston, that the sentence has a certain illocutionary-act potential (*PI*, p. 75, 1970).

In "Meaning and Speech Acts" he also often speaks of the *use of sentences*, as for instance when he says "When a sentence occurs categorically and is used to make an assertion, it has the assertion sign in front of it" (*PI*, p. 90). However, if he appears to use the notion of illocutionary act potential in "Meaning and Speech Acts" as a salient feature of sentences (and as opposed to the illocutionary force of utterances), this does not seem to be the case in the article "Appendix: A Reply to Mr G. J. Warnock" (Hare 1971b), which follows immediately after "Meaning and Speech acts" in *PI*. Here he maintains that:

> The meaning of a fully explicit sentence (one in which the speaker has said in full what he means) determines what he must mean if he is using the words in it correctly (as he must be, if he has said what he means) (*PI* p. 96).

He also maintains that:

> it is always open to us to study the speech act by formulating it explicitly and studying the meaning of the *resulting sentence* (*PI*, p. 96. My italics).

From these and the former passages we may single out at least two possible ways of speaking of sentences. Thus, Hare speaks of:

(**B1**) The utterance of a sentence or that the speaker uses the sentence to make an assertion, order, request, etc.,

(**B2**) The speech (or utterance) act results in a sentence.

The expressions occurring in (**B1**) are often used among speech-act theorists. They may be interpreted in several ways, which partly derives

from the fact the word 'use' is, as Austin pointed out, "a hopelessly ambiguous or wide word" (Austin 1962, p. 100), partly because it is unclear whether 'sentence' should be taken as referring to a token or a type.

I suggest as a reasonable interpretation that we understand Hare in the following way: When a speaker performs an illocutionary act, this act results in a sentence-token (or utterance-object). The speaker can now be said to have used a sentence-type in producing this sentence-token. The sentence-token (or utterance-object) will have an illocutionary force, whereas the sentence-type will have an illocutionary act potential.

Interpreting 'sentence' as 'sentence-token' in both of the two expressions of (**B1**) has the disadvantage that it becomes unclear how Hare could sustain the claim that the sentence has an illocutionary act potential. *Qua* token, the sentence will qualify in the same way as the utterance (-object) as having an illocutionary force. The only language-tokens that could be said in a reasonable sense to have an illocutionary act potential, are those that are part of or belong to another language-token that has illocutionary-force.

Suppose a person says

(u) 'Guillem said *obre la finestra*, which is Catalan for open the window'.

Let us say that (u) is an utterance, which has the illocutionary force of a statement. Now the utterance token (u) could reasonably be said to contain two more utterance(/sentence)-tokens, viz., (u') 'Obre la finestra', and the English equivalent to (u'). However, let us confine our attention to only (u'). When Guillem uttered 'Obre la finestre' his utterance had the illocutionary force of a request or an order ('open the window', whereas (u') has only an illocutionary act potential when it occurs in (u). In such a case it is possible to speak of a language token as having an illocutionay act potential.

Are then the only language tokens of which we can speak meaningfully as having an illocutionary act potential, those which are part of utterance-objects? With the possible exception of unintended tokens that are similar to language tokens, I shall, on the strength of the above example, risk advancing an affirmative answer to this question. I do not see how

sentence-tokens could be said to have an illocutionary act potential other than in the sense of being parts of utterance-objects.[9]

A word must also be said about Hare's use of 'utterance'. Above I have most of the time been speaking of utterance-objects. Hare, in "Meaning and Speech Acts", for instance, seems often to have the act-sense in mind when he speaks of utterances. Again, other times it seem equally clear that he has an object-sense in mind:

> the fact that, in an utterance like 'The cat is on the mat', the verb is indicative tells us, in the absence of certain contextual counter-indications, that the speaker is performing one of the genus of speech acts which we may call 'assertions' (*PI*, p. 77).
>
> For my own part I would say that 'Go East' and 'Go West' expressed different locutionary acts but the same illocutionary act (*PI*, p. 95).

As I see it, nothing hinges on whether we, in saying that an utterance has an illocutionary force, are using 'utterance' in its act or object sense. Moreover, as was seen in (**B2**), Hare sometimes speaks of that which results from a speech act as a sentence, and sometimes, given a reasonable interpretation of the two above passages, as an utterance-object. This discrepancy reflects what I believe is a sound idea of Hare's, viz., that there is no sharp line to be drawn between sentences-tokens and utterance-tokens, and sentence-types and utterance-types. Whether we should call the resultant of a speech act a sentence or utterance is in the end a matter of stipulation (cf. Austin 1962, p. 6, note 2; Wetterström 1977).

In his review of *FR*, Bergström (1964) draws attention to, among other things, Hare's use of 'judgement' which he thinks Hare sometimes uses ambiguously to denote sentences, and at other times to denote that "which is - or can be - expressed by some sentence or sentences" (*op. cit.* p. 42. cf. p. 46). He does not further specify what he means by 'sentence' other than what is implied in the following: "for it seems that it is a sentence rather than a judgement (in my sense) which can contain certain terms (terms

9. What should we for instance say of a token that is similar to an English language token but which has been produced by mere accident? Suppose weather conditions by a queer coincidence have engraved in a rock what seems to be a perfect English expression. Ought we to say that it is a language token? That it has meaning? Wetterström (1977, pp. 44–45), who has given an account of many of the central notions involved in this section, with which I am almost entirely in agreement, thinks that it would not be an arbitrary decision to call such a token an expression token. However, it would not deserve to be called an "utterance-object" (*op. cit.* p. 43), since such an expression token does not belong to an utterance-act, nor does it result from an utterance-act. This point seems reasonable.

being linguistic entities) and which can be indicative" (*loc. cit.*). Nor is it clear how we are to understand Bergström's 'judgement'. Since it is often said that sentences express "propositions", perhaps it is to this he alludes. However, as Wetterström (1977, p. 39) has pointed out, "the traditional notion of a proposition does not cover much of the illocutionary force of an utterance". The two "sentences" 'It is raining' and 'Il pleut', for instance, are often said to express the same proposition, where the proposition is intuitively understood as an abstract, non-linguistic entity, viz., *that* which these sentences express. But what are we to say, for instance, about sentences such as 'Let it rain', 'Will it rain?' and 'I promise it will rain'? What propositions do such sentences express? Hare, as well as other writers, has attempted to find a replacement for the traditional notion of proposition inside the frames of speech-act theory.[10] Thus, Hare distinguishes between a sentence's "neustic", "phrastic" and "tropic". The first two were used already in *LM*, and the tropic was introduced in Hare (1970; cf. 1989b), where he narrows the meaning of "neustic". The meanings he gives to these terms are the following:

> **Phrastic**: The part of sentences which is governed by the tropic and is common to sentences with different tropics (*PI*, p. 90. In Hare 1989b, he calls it 'clistic').
>
> **Neustic**: the sign of a subscription to an assertion or other speech act (*loc. cit*).
>
> **Tropic**: the sign of mood (*loc. cit*).[11]

For instance, the sentences:

(a) You are going to shut the door.
(b) Shut the door!

would in Hare's account have the phrastic/clistic "Your shutting the door in

---

10. However, Hare sometimes does use 'proposition'. See for instance *FR*, p. 11.

11. Hare's transformation of sentences into these "subatomic particles" has been the subject of much interest and criticism. For a recent criticism, see Hamblin (1987, p. 108). See also Ryding (1975), who mistakenly, though, associates threats with the prescriptive speech act. An early examination of Hare's views on this matter is found in Braithwaite (1954), who in his turn enriches Hare's neustic with a ninefold classification. Furberg (1971) has drawn attention to what he thinks are similarities between Hare's neustic and Husserl's "Materie", and Austin's notion of "locutionary". See here also Nordenstam (1966). Similar notions to those suggested by Hare are, for instance, developed by Stenius (1960).

the immediate future" (*LM*, p. 18). The idea, which Hare shares with several other speech-act writers, is that on some level of linguistic analysis these sentences will have some semantic element in common.

The phrastic is what shows most affinity with the notion of proposition (cf. Searle 1969, for instance, where "propositional content" clearly refers to that which the sentences above have in common). However, it seems highly implausible that it would be merely a phrastic which Hare has in mind when he speaks of judgements. Judgements will at least sometimes carry truth values, namely, when they are assertive judgements. However, Hare's phrastic/clistic denotes participial phrases that lack truth-value. Whether or not interrogative, assertive, and prescriptive sentences have a common core or not, is a complicated issue that I will not go into here (a good critical examination of Hare's, Searle's, and Stenius' ideas is found in Huntley, 1980; see also Peetz, 1972).

Admittedly, Hare has not made it easy for the reader, and like Bergström I have found it obscure when Hare speaks of use or misuse of judgements. However, Hare does make it clear in *FR* that he uses 'judgement' in a special sense. Thus, he says in a note:

> I have used the term 'judgement', here and in *LM*, in an artificially general sense, in order to avoid subscribing to the fiction that all indicative sentences express statements, and in order to leave *some* questions unbegged (*FR*, p. 10).

Keeping in mind that Hare uses 'judgement' in an "artificially general sense", a plausible interpretation is found if we understand 'judgement' in the light of what has been said till now about sentences and utterances. Thus, a judgement (-token) would, alongside sentence/utterance-tokens, form a class of utterance-objects: They are all, *qua* intended language tokens, the resultant of an illocutionary act. Not all utterances would be sentence-tokens or judgement-tokens. I may, for instance, say or write or think 'Blurp, goes have tingling', which would not amount to a sentence-token or judgement-token. Grammarians and linguists should be able to tell what counts as a correctly formed sentence or utterance for each natural language. We may also agree with what I take to be Bergström's underlying assumption that there are judgements that would not involve language tokens at all, in which case not all judgements would be utterance-objects. However, I do not see what is lost here by allowing 'judgement' in Hare's case to denote utterance-objects, i.e. the resultant of an illocutionary act which is in the indicative mood. The proviso to the effect that 'judgement' is

restricted to utterance objects in the indicative mood, is a reasonable interpretation (see e.g. *FR*, p. 10). Although he does not explicitly say so, I suspect Hare does not want his use of 'judgement' to include imperatives, for instance.

Henceforth I will use 'utterance', 'sentence', and 'judgement' without specifying whether they are used in an act/object or type/token sense. Having realized that there are these ambiguities involved here, it would be unnecessarily cumbrous to pay attention to these distinctions.

2.3 John R. Searle's First Objection

Let us now return to the quote from "Meaning and Speech Acts" above, p. 59. It was said here that word-meaning, at least in part, can be explained in terms of its contribution to sentence-meaning, which in turn was to be explained in terms of illocutionary force (or act potential).[12] Here it is important to recall Hare's distinction between species and genus of speech acts. Hare has not presented any taxonomy of speech acts, but he has, as may be recalled from chapter I, concentrated mainly on two genera, viz. what he sometimes refers to as "prescribing" and "asserting" speech acts. Species of the former group are all ways of *telling someone to make something the case*. To perform a speech act belonging to the latter genus is on the other hand to *tell someone that something is the case*. He maintains that "Species of this genus are: statements, declarations, guesses, and so forth" (*PI*, p. 77, 1970). Examples of prescriptive species of speech acts which Hare has mentioned are: advising, instructing, commanding, ordering, requesting, commending, and entreating.

Hare's claim about word- and sentence-meaning must be understood in the light of this division between genus and species of speech acts. With the exception of one group of sentences, he has never maintained that knowing the meaning of a sentence is knowing that it is commonly used to perform a certain *species* of a speech act. When a person says, for instance, 'You ought to go to bed' or 'Shut the door', we may be in doubt about whether what is said is a piece of advice, a request, or simply an order. The sentence in itself will not reveal which species of speech act is involved. Thus, Hare says in one place:

12. I am unaware whether anyone ever has endorsed the view that words taken in isolation can have their meanings explained in terms of speech acts. Notwithstanding such a theory has been criticized by David Holdcroft, "Meaning and Illocutionary Acts", in Parkinsson (1982). Cf. Wellman (1961 pp. 131–33)

> suppose that I say 'Shut the door'; and suppose that I am a colonel, and the addressee, let us say, a captain. It might be unclear, in such a case, whether this was a request or an order. Worse, I might not myself have specified, even to myself, which it was; if the captain then shut the door, neither of us would need to enquire which it was. . . .When I say 'Shut the door', and do not specify whether it is an order, I may be simply giving the utterance the *generic* illocutionary force of an imperative, and just not caring or specifying whether it is an order or a request or a piece of advice. But suppose that the captain says, unexpectedly, 'I won't'. *Then* I shall have to decide: Do I intend it as an order or just a request. I shall have to say, either 'Oh, never mind if you do not want to' or 'I order you to shut it.' (*PI*, p. 112).

Hare's claim is rather that whatever illocutionary act the speaker will perform, given that he is not using the words idiosyncratically or in an inverted commas way, these acts will have in common that they are all instances of telling someone to make something the case. The exception referred to above concerns a kind of sentence for which Austin coined the term "explicit performatives", of which 'I order you to shut it' is a good example (contrast Jost 1976, p. 261). Other examples of such explicit performative expressions are: 'I promise that', 'I bet', 'I request', 'I state'. Austin pointed out that such expressions have in common that they are commonly "used in naming the act which, in making such an utterance, I am performing" (Austin 1962, p. 32. Cf. *PI*, pp. 102–105).

Hare has little to say about how the different species of prescribing differ from each other (cf. Carter 1982, p. 223).[13] However, although such an account would have been befitting, I do not see how it would accomplish more than that we could exclude or include certain borderline cases.

Various writers have failed to notice that Hare is speaking of genus and not species of speech acts. For instance, G. J. Warnock (1967), in his attempt to criticize the view that moral discourse "must consist essentially and always in the performance of any *single* speech act", contends that:

> there are at any rate dozens of things which those who employ moral words may therein be doing. They may be prescribing, certainly; but also they may

13. However, see *PI*, pp. 54–55, where Hare outlines the distinctions between the illocutionary acts 'ordering' and 'commanding', and the perlocutionary act 'exhorting'. For an analysis of the (perlocutionary) act of alluding which is inspired by speech act analyses, see Hermerén (1992).

be advising, exhorting, imploring; commanding, condemning, deploring; resolving, confessing, undertaking; and so on, and so on (*op. cit.* p. 35.).[14]

But, as he makes clear in Hare 1970, 1971b, the fact that we can perform these and other illocutionary acts, is in itself no objection to his claim that moral terms in their primary sense have prescriptive meaning. The reason is that Hare is talking about a certain genus and not about a certain species of illocutionary acts.[15]

Hare's prescriptivity-claim has been questioned from a different angle, which at first sight seems more serious than the above one. The alleged fact that we are prescribing when we issue an utterance such as 'I ought to help my neighbours', should be understood in the light of the fact that 'ought' can figure in several other kinds of sentences, e.g. 'Ought I help my neighbours?'. When Hare speaks of explaining word-meaning in terms of how words contribute to the illocutionary act potential of sentences, he appears to consider as appropriate the simple present tense indicative sentence-form.[16] *Prima facie* such an idea seems to be vulnerable to criticism. The very fact that 'ought' and 'good' can be incorporated in non-indicative sentences such as hypothetical and interrogative sentences, the utterances of which clearly do not belong to prescriptive discourse, seems to refute Hare's thesis. The classic statement of this argument was made by P. T. Geach (1960) (cf. also Castañeda 1963, pp. 235-234). More recently John R. Searle has maintained a variant of Geach's criticism, and I will therefore mainly focus on his line of reasoning (for a critical examination of Geach's

14. A similar objection is found in Edel (1963, p. 152). Warnock's view is criticized in Hare (1968a): "'Prescribing' is a convenient name for a genus of speech acts distinguished by being 'intimately related to conduct' and by the possibility of deeds being 'consonant or dissonant with words'. . . .Although the limits of this genus can be disputed, certain typical and central uses of 'ought' are clearly species of it" (Hare 1968, p. 438). See also (*LM*, p. 2, *FR*, p. 27), for instance. Hare sometimes uses the term "Command" (e.g. in *LM*) as well as "Imperation" in order to refer to this genus (Hare 1971, p. 56). For a criticism of Hare's distinction between commands (which require authority) and orders (which only require power), see Broadie (1972).

15. Hare has sometimes been mistakenly interpreted as holding the view that to issue a moral judgement is equal to the issuing of an imperative (e.g. Broadie 1972, pp. 181). Hare has made it clear that his prescriptivism should not be confused with such, as he calls it, "imperativism" (Hare 1976b, p. 23).

16. See e.g. Hare (1971, p. 75) in which the sentence "That is a good movie" is considered appropriate.

article, see Zimmerman, 1980;[17] other writers who endorse this argument are Urmson 1968; Franklin 1973; Forrester 1974; Blose, 1974 and von Wright, 1977).[18]

Searle claims that Hare commits what he has termed the "Speech Act Fallacy".

> *Calling* something good is characteristically praising or commending or recommending it, etc. But it is a fallacy to infer from this that the meaning of 'good' is explained by saying it is used to perform the act of commendation (Searle 1969, p. 139).

And some lines further on he purports to demonstrate this by showing that there is an

> indefinite number of counter-examples of sentences where 'good' has a literal occurrence yet where the literal utterances of the sentences are not performances of the speech act of commendation (Searle 1969, p. 139).

On the other hand, Hare has countered by maintaining, and here he has a point, that Searle himself explains the meaning of 'I promise' in terms of the illocutionary act of promising, although he admits that we are not promising when we say, for instance, 'Do you promise to do so and so?' or 'If you promise to do so, I will feel better'. Hare believes that in a way similar to how Searle undertakes to explain 'promise' we can explain the meaning of 'good' in a categorical affirmative such as 'That is a good movie', to the effect that its meaning can be :

> extended to cover the meaning of utterances in which the speech act in question is not performed, provided that the utterance is generated by a transformation, whose form we understand, of the original speech act (*PI*, pp. 84–85, 1970).

Searle maintains that although there is a group of words, namely, the

17. Zimmerman (1980, pp. 220–221) puts forward an objection on his own against the claim that the meaning of a word is partly explained by determining how it contributes to the illocutionary act potential of the sentence in which it occurs. However, I miss in Zimmerman's criticism a reference to Hare's claim that the prescriptive meaning of moral terms is logically prior to the non-prescriptive uses.

18. Forrester endorses an argument that appears to be stronger than Searle's, and would for that reason deserve a special treatment. However, her views on illocutionary force are, I suspect, controversial. She claims, for instance, without hesitation that "a conditional sentence has the illocutionary nature of its conditioned clause. 'If it is nine o'clock, go to bed' has an imperative for its consequent, and it is itself an imperative" (Forrester 1974, p. 768). This claim clearly seems too categorical.

performative verbs, of which the above holds, this cannot be said about terms such as 'good' and 'ought'. Thus Searle asks us to:

> Consider the following examples: 'If this is good, then we ought to buy it' is not equivalent to 'If I commend this, then we ought to buy it'. 'This used to be good' is not equivalent to 'I used to commend this'. 'I wonder whether this is good' is not equivalent to 'I wonder whether I commend this', etc (Searle 1969, pp. 138–139).

Searle's point is that since 'good' obviously means the same in all these sentences, including in an indicative such as 'This is good', then claiming that 'good' is used to perform a commendative speech act cannot be an explanation of the meaning of 'good'. Thus, Searle (as well as Geach; cf. Zimmerman 1980) appears to take the occurrence of moral terms in embedded clauses to constitute a *reductio* of the claim that 'good' and 'ought' have prescriptive meaning. The fact that we are not prescribing when we utter a conditional such as 'If X is good, then Y is the case' shows, it is argued, that the meaning of 'good' does not depend on the prescriptive illocutionary force.

Searle accounts for those cases when 'promise' occurs in sentences that are not used to make promises, by explaining them as being cases of the speech act being *reported* or *predicated*. Moreover, he seems to think that such an analysis can only be applied to performative verbs and not to a word such as 'good'.

Hare claims, in his turn, that such a pattern can be extended to cover words other than performative verbs. He denies that:

> the sentence 'That is a good movie' means the same as the explicit performative 'I (hereby) commend that movie' (*PI*, pp. 75–76, 1970).

However, he maintains that, just as Searle explains, for instance, 'Do you promise?', in a way that is consistent with the claim that 'I promise', has a performative meaning, some occurrences of 'good' can be accounted for in a similar way. In other words, such occurrences are explained as cases where the speech act of commending is reported or predicated.[19]

Aaron Sloman (1969) has pointed out that Searle has not realized the following: of the occurrences of 'promise' that are not cases of promising,

19. The fact that Hare agrees with Searle that prescriptive words that occur in the conditioning part of a hypothetical sentence will not make the utterance of the sentence into a prescriptive sentence, will much later, in chapter V, raise an interesting problem for what Hare has to say about the meaning of 'I'.

there are cases which cannot be explained as reports and predications. The person who says 'I do not promise to come', for instance, is not promising, but nor is he reporting or predicating the speech act of promising. Hare, who is also alert to this problem, believes that such examples where the speech act is not being performed are a further example of the speech act being "in the offing":

> [Searle] unfortunately does not sufficiently distinguish between the different ways in which a speech act can be 'in the offing', and confines his attention to cases in which the act is reported, and the like (*PI*, p. 79, 1970).

Some lines further on he is concerned with those other cases in which the speech act can be in the offing, namely those:

> in which what is negated is not a report of a promise but a promise, and those in which the question [Do you promise to pay me $5] expects as answer not a report, but either a promise or the negation of a promise. In these, the speech act is in the offing in a different way; but all the same it is possible . . .to explain the meaning of the word 'promise', in this syntactical context, in terms of it, as we shall see. The trouble for Searle is that the type of explanation, once understood, can be extended to cover 'good' as well as promise, as we shall also see (*PI*, p. 80).

Hare's idea seems to be the following. We are not commending when, for instance, we ask 'Is that a good movie?'. However, if we understand the meaning of 'good' as used in the categorical affirmative, 'That is a good movie', and we understand the meaning of the interrogative sentence-form, then we can "put the two together" (*op. cit.*, p. 82) and understand the meaning of 'Is this a good movie?'.[20]

The analysis of the indicative present tense sentence should in other words account for why the explanation of what we are doing in uttering such a sentence should enter the analysis of sentences in which 'good' is embedded, whereas the opposite does not hold.

Hare's explanation seems intuitively plausible. To that effect it can be said to meet Searle's criticism. Notwithstanding, there is on Hare's as well as on Searle's part, as long as a further explanation is not given of how "put the two together" should be understood, unfinished business to attend to.[21]

---

20. Cf. Dunn (1986) who believes that Hare answers Searle's criticism. Another way of stating Hare's defense would be to say that when moral terms occur in embedded clauses, they function as tropics – identifying a certain (genus of) illocutionary force; the neustic (subscription indicating particle) is not, however, embedded.

21. Both writers have continued their work on speech acts. See Hare (1989b), for instance. However, this article does not say anything further about how we should

At this point it will be worth quoting a passage at length, which reveals the link which holds, according to Hare, between meaning and intentions:

> The meaning of a fully explicit sentence (one in which the speaker has said in full what he means) determines what he must mean if he is using the words in it correctly (as he must be, if he has said what he means), and determines, similarly, to what he must be referring if he is following the 'demonstrative conventions' (Austin's term from *Philosophical Papers*, p. 90) correctly. If he says 'my left hand', he must be referring to his left hand, or he is misusing the words. It thus determines what speech act he is performing. What Warnock calls 'the circumstances of the utterance and the intentions of the speaker' are therefore not, as he seems to think 'something additional' to what is thus determined by the meanings of the words, so that enquiry into this 'something additional' makes the study of speech acts into something more than the study of meanings. For the meaning, if (as Searle proposes) it is made explicit, will determine what the intentions and the circumstances have to be, if the speech act is to be validly performed (*PI*, pp. 96–97, 1971b).[22]

This passage contains an explicit statement of how Hare conceives of the relation between meaning and intentions: Sentence-meaning determines the intention which the speaker must have in order to have performed a valid illocutionary act. Moreover, to understand the meaning of an utterance involves understanding with which intention the speaker issued his utterance. Hare as well as Searle and Wetterström sometimes speak of illocutionary act devices, e.g. word-order, punctuation, the mood of the verb, the presence of performative verbs, tense and the person-denotation of the verb (Hare 1970; Searle 1969). Such conventional devices determine how we should understand which illocutionary act is performed, and ultimately, which illocutionary intention the speaker had in issuing the utterance.

Suppose the speaker wants to utter, say, 'Do X!', and use this sentence with the meaning it has as a result of language conventions. In such a case the person must utter it with the intention of 'telling someone to make something the case', i.e. he must, according to Hare, be prescribing (Searle

---

understand "put the two together".

22. For an interpretation that is based on the *LM*-account, see Castañeda's "Imperatives, Decisions and 'Oughts'" in Castañeda & Nakhnikian (1963). Alston's account of illocutions differs, I believe, from most recent writers in that he seems to claim that for two speech acts to have the same illocutionary force, they must express the same proposition (*op. cit.*, pp. 36–39). Cf. Cohen (1970, p. 556), who claims that "There is an infinite variety of speech acts in a language, so far as there is an infinite variety of different sentences", which indicates, I think, that he agrees here with Alston.

and Dunn use the term 'Directives').[23] Otherwise it will be a sign that he has not understood what this sentence means or that he is using it idiosyncratically.

Now with regard to 'good' and 'ought' we may reasonably ascribe the following view to Hare (e.g. *PI*, p. 75): In conjunction with *certain* illocutionary act devices 'good' and 'ought' determine that if a person utters, say, the present tense indicative 'That is a good movie' (which includes these illocutionary act devices) or 'You ought to see that movie', the person must have in uttering this sentence the intention to perform some such illocutionary act such as commendation or advice, etc. that belongs to the genus of prescribing.[24] Otherwise it will be a sign that he has misunderstood the meaning of the sentence or because he wants to mislead the hearer or that he uses 'good' or 'ought' in what Hare calls an inverted-commas way (see chapter I; *LM*, p. 125, 164). The latter possibility complicates the matter, and I will return to this issue below. On the other hand, when 'good' or 'ought' figure in sentences with certain other illocutionary act devices - notably the ones found in hypothetic or interrogative sentences - the speech act required will be of a different kind. The prescriptive speech act will be in the offing.

2.4 Searle's Second Objection

Searle raises another objection against the thesis that 'good' or 'ought' have prescriptive meaning. He does not explicitly mention Hare, but it is reasonable to understand him as directing his criticism against Hare, who, after all, is the foremost exponent of the theory that Searle criticizes:

---

23. Searle (1979) presents the following taxonomy of speech acts: Assertives (we tell people how things are), Directives (we try to get them to do things), Commissives (we commit ourselves to doing things), Expressives (we express our feelings and attitudes), and finally Declarations (we bring about changes in the world through our utterances). What Searle means by Directives comes close to what Hare calls the genus of prescribing - although I suspect that Hare would object to the characterization of this group as case of "getting" hearers to do things. Shwayder in "Uses of Language and Uses of Words: With Application to a Problem of Frege", in Parkinsson (1982), expresses doubt about the possibility of presenting a taxonomy of speech acts. Such scepticism is sound, if we consider species of speech acts. However, I am less sure whether such an attitude should be extended to cover a listing of genera.

24. Wetterström expresses what may amount to a similar view: "I think that words of the kind exemplified by 'good', 'evil', 'ought', 'duty', characteristically works as i-devices [(illocutionary devices] which give utterances that employ them various sorts of evaluative or normative forces, notably ethical ones" (Wetterström 1971, p. 172).

> In ethics it has commonly been supposed that 'good', 'right', 'ought', etc. somehow have an imperative or 'action-guiding' meaning. This view derives from the fact that sentences such as 'You ought to do it' are often uttered by way of telling the hearer to do something. But from the fact that such sentences can be uttered as directives it no more follows that 'ought' has an imperative meaning than from the fact that 'Can you reach the salt?' can be uttered as a request to pass the salt it follows that *can* has an imperative meaning. Many confusions in recent moral philososphy rest on a failure to understand the nature of such indirect speech acts (Searle 1979, p. 32).

The failure which Searle speaks of, is one of not distinguishing between *sentence* meaning and *speaker* meaning. We can, according to Searle, prescribe (or according to his taxonomy, perform a directive speech act) when we utter a sentence such as 'You ought to do it', but not because the meaning of the sentence is prescriptive, but because we perform what Searle calls an *indirect* speech act: we perform one speech act by performing another. For instance, if I say, to take another example, 'You are standing on my foot', I would be stating something. However, sometimes by stating this, we request our hearer to remove his or her foot. On such occasions the best explanation according to Searle is to say that what has been added to the utterance is not a new sentence meaning but a new speaker meaning (Cf. Hamblin 1987, pp. 7, 8. See also Falk 1953).[25]

Now Searle's point seems to be - at least in the case of the word 'ought' - that those who claim that 'ought' has prescriptive meaning, have overlooked that when 'ought' has been used in a prescriptive way, it is because it has been used in an indirect speech act - in much the same way as 'you are standing on my foot' is sometimes used in order to make a request.

Whether or not Searle's explanation of such so called indirect speech acts is felicitous, it is clear, I believe, that it does not follow from his "indirect speech act theory" that 'ought' *cannot* have prescriptive meaning in the sense in which Hare thinks it has. The reason is that Searle's criticism is based on the assumption that a sentence such as 'you ought to do it' has an assertive illocutionary act potential (cf. Cohen, 1970). But this is exactly what Hare denies. This claim is simplified and needs to be qualified. That is, what Hare maintains is that a sentence such as 'You ought to do X' has both a prescriptive and an assertive act potential, where the former, he

25. In his paper "Goading and Guiding" Falk outlines a view that - other differences apart - reminds one of Searle's indirect speech-act theory. Thus, he maintains that normative judgements can be used prescriptively in much the same way as statements of fact can be used prescriptively (Falk 1953, especially pp. 150–1).

argues, is logically prior to the latter (see here the next section).[26] The discussion at issue concerns basically whether the illocutionary act devices of 'ought-sentences' place them in the group of prescriptive language, as Hare claims, or among some other genus of speech act such as the group of assertives (which seems to be Searle's view).

In other words, the actual issue between Hare and Searle concerns the fundamental question whether these words are used to prescribe with by virtue of their cognitive meaning and circumstances (such as having a certain desire, for instance) - which at least in the case of 'good' is Searle's view - or because the conventions of language associate a prescriptive act potential with these words (Cf. Wetterström, 1977, pp. 170–171).[27] Only after having decided this issue, can we reasonably agree or disagree with Searle about whether Hare commits a speech-act fallacy or confuses speaker meaning with sentence meaning.

Searle considers 'good', for instance, to be similar to a word such as 'praiseworthy', which is used to praise as a

> *consequence* of the fact that 'praiseworthy' means what it does, i.e., 'worthy of praise' (Searle 1969, p. 153).[28]

Hare on the other hand, denies, as was seen in chapter I, that it is by virtue of some cognitive or descriptive meaning of 'good' that this word is characteristically used to commend. The term 'good' (and *pari passu* 'ought') has primarily this function in language that when the term occurs in a sentence containing certain illocutionary act devices, 'good' acts as a sign that the speaker is performing the illocutionary act of commending. Hence, if the speaker knows the meaning of the word and he wants to use it

---

26 During a discussion in Lund, and later in Ewelme, May 1989, Hare expressed doubts concerning Searle's explanation of indirect speech acts. Such situations could according to Hare be explained in terms of perlocutionary acts and mutual knowledge between speaker and hearer – for instance, that no one wants to get his feet stepped on. Dunn (1986), on the other hand, appears to accept Searle's explanation. See p. 79.

27. A discussion between Searle and Hare would not only, I believe, turn around the question of whether 'good' has prescriptive meaning. Equally important is the question of whether prescribing is, as Hare maintains, logically prior to other speech acts performed in using 'good'. For an argument to this effect see *LM*, pp. 118

28. Drawing upon Wittgenstein's analysis of 'game', Searle contends that the meaning of 'good' is best understood as a family meaning, in which the following is prominent: "Meets the criteria or standards of assessment or evaluation" (*op. cit.*, p. 152).

correctly and not in order to, say, mislead the hearer, he must utter the sentence with the intention to commend or to guide action.

2.5 Inverted-Commas Uses

Hare's thesis that 'good' and 'ought' have prescriptive meaning is made somewhat complicated when we take into account what he has called the inverted-commas use of these terms.[29] For instance, in *LM* he says:

> We noticed that it is possible for people who have acquired very stable standards of values to come to treat value-judgements more and more as purely descriptive, and to let their evaluative force get weaker. The limit of this process is reached when, . . .the value-judgement 'gets into inverted commas', and the standard becomes completely 'ossified'. Thus it is possible to say 'You ought to go and call on the So-and-sos' meaning by it no value-judgement at all, but simply the descriptive judgement that such an action is required in order to conform to a standard which people in general, or a certain kind of people not specified but well understood, accept. And certainly, if this is the way in which an 'ought'-sentence is being used, it does not entail an imperative (*LM*, p. 164).

It is unclear just how we should understand the process of going from prescriptive to inverted-commas uses. The expression "more and more" lends some support to the thought that there may exist intermediate steps, when a judgement might be more or less prescriptive, or more or less descriptive. But when this is translated into the idiom of speech-act theory I have difficulties in seeing how this could be the case.[30] That is, given the two kinds of meaning, a person may either be performing a speech act belonging to the genus of prescribing or one belonging to the genus of asserting. Perhaps all that Hare wishes to say is that the desire or preference involved in making a prescriptive judgement may be of varying strength, and that at a certain time it will vanish and be replaced by a belief of which the descriptive judgement is an expression. However, the very process which leads from a prescriptive judgement to an "ossified" inverted-commas judgement gives rise to more questions. How should we understand the relationship between "acquiring very stable standards" and

29. To this kind of use should be added what he calls the conventional use: "the speaker is merely paying lipservice to a convention, by commending, or saying commendatory things about, an object just because everyone else does" (*Ibid.*, p. 125).

30. See Jost (1976) who considers whether Urmson's distinction between central and peripheral illocutionary forces (Urmson, 1969) could replace Hare's distinction between primary and secondary meaning. Jost's conclusion is negative. I share Jost's criticism of Urmson's distinction, but for different reasons than Jost.

the acquisition of inverted-commas moral judgements? Could we, for instance, issue an inverted-commas judgement, such as the one mentioned above, and have in addition a desire that the person should go and call on the So-and-sos? It seems unlikely that the acquisition of standards in itself would account for why a person would lose the desire involved. However, I will not pursue these questions, but concentrate on what Hare understands by the prescriptive uses of moral judgements (for a critical article of Hare's account of this issue, see Knox 1970).[31]

I have interpreted Hare here as implying that when we utter, say, 'You ought to do it' in an inverted-commas way, we utter it with the intention to perform a speech act belonging to the assertive genus. Although Hare never explicitly maintains this, I believe this to be a reasonable interpretation. On this interpretation, 'ought' will, on the one hand, act as a sign in conjunction with certain other illocutionary act devices that the speaker is prescribing; on the other hand it will act as a sign that the person is asserting something or doing some other speech act belonging to the assertive genus. Another way of putting it would be to say that the term 'ought' in one and the same sentence (s) determines that (s) has a prescriptive and an assertive illocution act potential (here Searle's account of indirect speech acts would have the advantage over Hare's theory in offering a simpler account).[32]

Now, do 'good' and 'ought' act as a sign that an assertive and/or prescriptive speech act has been performed? One way of testing Hare's claim that 'ought' has (partly) prescriptive meaning is to examine whether there are 'ought'-utterances - other than the inverted-commas uses (including the ironic and conventional ones) - that are difficult to combine with the idea that we are prescribing when we utter them. In order to answer this question we need to further deepen our understanding of Hare's view on prescribing.

---

31. Knox's article raises some interesting questions. However, the criticism put forward rests unfortunately, to a great extent, on assumptions that Hare denies, and which Knox does not argue for.

32. Should we then say that 'ought' is ambiguous? Hare denies this. Following Nowell-Smith (1954), he describes moral words as "Janus-words" which "have two or more aspects to their meaning, one of which may on occasion be emphasized to the neglect of the others. We cannot say that such a word is ambiguous; it is indeed an inseparable element in its meaning that it can shift in this way" (*FR*, p. 75). I am inclined to disagree with Hare over this issue. Compare, for instance, what would be the case if an expression such as "Do so-and-so" acted as a sign that the speaker was either issuing an imperative or an assertion that something is the case. On my part, I would call such an expression ambiguous.

2.6 Prescribing

The question that will concern me in the sections to come is whether the logic of prescriptions differs from the logic of 'ought'. To begin with we need therefore to get a more detailed account of what Hare considers to be a prescriptive speech act. We need in other words to achieve clarity about the following question.

(*) How are we to understand what it is to tell someone to make something the case?

In particular there is a specific matter that is important to be clear about, namely:

(**) Does the speech act of telling someone to make something the case involve the speaker's desire that the listener should make something the case?

Now, we saw earlier that Hare maintains that the meaning of a sentence, among other things, determines what the *intentions* of the speaker must be, "if the speech act is to be validly performed" (*PI*, p. 97, 1971b). Thus, it seems reasonable to ascribe the following thesis to Hare - where the sentence (s) has a prescriptive illocutionary act potential:

(P) A speaker *a* performs a prescriptive speech act in uttering the sentence (s), only if *a* by uttering (s) intends to tell the addressee(s) to make something the case – where this something is determined by (s) and where the term 'tell' should be understood in its generical sense.

Hare must reasonably endorse something like **(P)**. In order to prescribe we must surely have issued the judgement with the intention to tell the hearer to make something the case rather than, for instance, to tell him that something is the case. By adding the proviso "where this something is determined by (s)", I hope to link together the conventional aspect of language with its intentional aspects. However, for my purposes nothing much hinges on this point. If the link is considered too weak or general, **(P)** may be regarded as being about intended and not actually performed speech acts. **(P)** would still, I believe, be a necessary condition for the performance of an illocutionary prescriptive speech act.

Unfortunately Hare has not offered any explicit explanation of what he thinks is involved in "telling someone to make something the case". However, he has, as shown earlier, argued that there is a logical connection

between prescribing and being disposed to act, which the following passage reveals:

> We say something prescriptive if and only if, for some act A, some situation S, and some person P, if P were to assent (orally) to what we say, and not, in S, do A, he logically must be assenting insincerely (*MT*, p. 21).[33]

Moreover, Hare has made it clear (*LM*, pp. 14–15) that prescribing is not an act of persuasion or influence. To prescribe is to give an answer to a question such as "What shall I do?" Thus, he maintains that:

> Telling someone to do something, or that something is the case, is answering the question 'What shall I do?' or 'What are the facts?' When we have answered these questions the hearer knows what to do or what the facts are – if what we have told him is right. He is not necessarily thereby *influenced* one way or the other, nor have we failed if he is not; for he may decide to disbelieve or disobey us, and the mere telling him does nothing – and seeks to do nothing – to prevent him doing this (*LM* p. 15).

Although this sheds some light upon the prescriptive speech act, the question remains what conditions a speaker must meet in order to perform such an act.

Let us now leave **P** and turn to the more specific question (**) above. As may be recalled, this question concerned whether the speaker, in addition to the intention referred to in **P**, also should be understood as having a desire or intention that the listener should do that which he is told to make the case. It can be seen from the following passage that Hare believes that there is a link between prescribing and this other volitional state:

> It is indeed true of imperative sentences that if anyone, in using them, is being sincere or honest, he intends that the person referred to should do something (namely what is commanded). This is indeed a test of sincerity in the case of commands, just as a statement is held to be sincere only if the speaker believes it (*LM*, p. 13, see also *FR*, p. 170 where he says that value-judgements have the characteristic 'of being dispositions to action').[34]

---

33. Cf. *LM*, sections 1.7 and 2.2.

34. Taylor (1980) claims "that it isn't altogether clear whether Hare means that any evaluation *whatever* incorporates some desire, or whether the point is restricted to evaluations by an agent of courses of action etc. to be undertaken by himself" (p. 513). In contrast to Ibberson (1986) and myself, Taylor chooses the latter interpretation (he does, however, recognize in a footnote that Hare in "Wanting: Some Pitfalls" implicitly maintains the former). Thus, Taylor claims that according to Hare, a third-person judgement such as "'Fisher should move his bishop' does not incorporate the desire on my part [the speaker's part] that Fisher should move his bishop" (p. 514). Such an interpretation seems to run counter to the quote presented. Besides, if we follow Taylor it becomes unclear how Hare would distinguish sincere prescribing from insincere

Moreover, in the passage quoted earlier, in which Hare gave an explanation of what is involved in "telling someone to do something", he made it clear that the listener

> may decide to disbelieve or disobey us, and the mere telling him does nothing – *and seeks to do nothing* – to prevent him doing this (*LM*, p. 15; my italics.).

These passages indicate that, according to Hare, we can perform a prescriptive speech act without having the intention that the addressee of our prescriptions should do the act prescribed. We can do this when we are insincere, just as we can state or tell that something is the case, without actually believing what we are stating. That this is the case can be seen *a fortiori* if we turn to his essay "Wanting: Some Pitfalls" (Hare 1968b). To begin with he recognizes that:

> There is, of course, a tiresome problem about the relations between the desire that p, which is not a speech-act, and the expressed command or request that p, which is. However, I do not feel called upon to say something about this problem, because it is a quite general problem about the relations between thought and speech, affecting assertions and commands alike (*PI*, p. 52).

Still, two pages later on he gives an example that clearly suggests that he thinks that one can execute a prescriptive speech act without having the intention/desire that the addressee should do the act prescribed:

> The sadistic schoolmaster, who commands his boys to keep silent in the hope that this will cause them to talk so that he can beat them, is still commanding or telling them to keep quiet (*PI*, p. 54.).[35]

Suppose prescribing that *a* should do X entails intending or desiring that *a* (should) do X. In such a case we would have to say about the schoolmaster that he has two inconsistent intentions, viz., the intention that the boys (should) be quiet as well as the intention that they should not be quiet. A more reasonable interpretation is to take this example as

---

prescribing.

35. In Hare (1987b, p. 74) Hare uses once more the example of the sadistic schoolmaster in order to argue for the separation of illocutionary and perlocutionary forces. However, he purports here to explain the prescriptivity of "Keep quiet" in terms of "what compliance with it would consist in".

expressing that prescribing does not necessarily involve, according to Hare, intentions that someone should make something the case.

Moreover, Hare has recently addressed this topic in "Some Sub-Atomic Particles of Logic" (Hare 1989b). He discusses here different ways in which subscription is withheld in the performance of a speech act. One such example of non-subscribing - what he calls a *mimetic* use of language - is expressed in the following passage:

> The proscenium arch which protects actors is an obvious example of such a subscription cancelling device. It is a convention that things said behind it are not being subscribed to. That is why, as I have been informed by Cristopher Taylor, actors have a rule that if a *real* fire breaks out backstage, the person who discovers it has to shout, not 'Fire!' (for fear that it might be thought to be part of the play) but some other expression earmarked for this purpose (Hare 1989b, p. 28).

We do not consider that the actor who says in a play 'open the window' or 'it is raining outside', actually is having an intention/desire that the addressee should open the window, or that he actually holds the belief that it is raining. But nevertheless, we do understand what the actor is doing in uttering these sentences, viz., telling someone to make something the case and telling him that something is the case respectively.

These passages strongly suggest that Hare regards the speech act of prescribing to be explainable (at least in principle) without invoking the speaker's desire that the addressee should do the act prescribed. Prescribing involves such desires only when the speaker is *sincere or subscribing* to the speech act. Hence, I believe it is reasonable to add to **P** the following sincerity condition (**SC**, for short):[36]

> **(P)** A speaker *a* performs a prescriptive speech act in uttering the sentence (s), only if *a* by uttering (s) intends to tell the addressee(s) to make something the case - where this something is determined by (s) and where the term 'tell' should be understood in its generical sense.
>
> **(SC)** A speaker *a* prescribes sincerely only if (**P**) holds, and only if he intends that the addressee(s) should do what is prescribed.

Both of these theses emphasize the need for answering a further

---

36. Searle (1969) has a similar way of analysing sincere speech acts. He says, for instance, "The distinction between sincere and insincere promises is that in the case of sincere promises, the speaker intends to do the act promised; in the case of insincere promises, he does not intend to do the act" (p. 60).

question, viz., what does it mean, according to Hare, to intend or desire something?

2.7 Hare's View on Desires

Hare's notion of desire is a general one, to the effect that he does not differentiate between wants, desires and, it seems, intentions. In Hare (1972d, p. 98), for instance, he makes it clear that he intends his notion to be the "equivalent of Aristotle's most general notion in this area, *oregestai*, to be motivated towards the doing or having of". Using the noun 'orexis' to this verb, he defines "to have an *orexis*, to want something" in the following way:

> a man has an *orexis* to do or get or retain a certain thing if and only if, other things being equal, he will seek to do or get or retain it (*MC*, p. 98, 1972d; cf. *FR*, p. 170).

I will not in this section attempt to outline all aspects of Hare's notion of wanting, but confine my attention to two important claims that he makes. The first matter concerns his thesis that there is a conceptual link between desiring and the assent to prescriptions. The second matter is his idea that desires are dispositions to act.

The first idea that there is a conceptual link between having a desire and assenting to a prescription, namely an imperative, is perhaps best expressed in "Wanting: Some Pitfalls" (Hare 1968b). Hare follows in this essay to some extent Kenny's "Imperative Theory of the Will" that to have a desire or an intention is to say "in one's heart" (*PI*, p. 51) a first-person imperative (cf. *MC*, p. 98). For instance, he suggests that:

> the man who intends to do A is the man who in his heart *subscribes* (not necessarily subvocally and not necessarily occurrently) to the command 'let me do A' (*PI*, p. 52) .

Now, as has been argued by Huntley (1980), for instance, Kenny's theory appears clearly to be circular, *qua* account of wanting.[37] However, it is

37. Huntley (1980) says correctly about Kenny's theory: "In order to avoid circularity it must provide an account of the notion of saying (and, in particular, of saying-in-one's-heart that which does not involve essential reference to the psychological notions, such as belief and desire, which are to be explicated in terms of it". This will not, he thinks, be possible. The same objection could be raised against the quoted passage.

imperative to realize that Hare's point of bringing in this idea of "subscribing in one's heart" is rather a different one than to give an *analysis* of intending/desiring.[38] In much the same way as philosophers often speak of 'that' clauses as the linguistic expressions of the content of beliefs, Hare aims with the above theory to establish a similar link between words or expressions that express the conceptual content of desires. "Unless we make some such move as Kenny's", he argues, "we shall not be able to display the logical relations between desires and other thoughts or expressions" (*PI*, p. 47).[39] Hare's idea is that it is the imperative which is the appropriate expression in language for what a person is thinking or experiencing when having a desire.

The second matter which must be noticed, with regard to Hare's notion of desire, is that he holds what is often called a dispotional view of desires. To have a desire is to be disposed to act. Unfortunately, Hare does not further specify how he conceives of dispositions. A matter, for instance, which I would like to have seen clarified by Hare, is how he regards the relationship between beliefs and desires/dispositions. There is at least one view in the literature which I find unreasonable to ascribe to Hare. Let me therefore briefly comment on this issue.

In the literature there are various ways of regarding desires as dispositions.[40] On what has been called the *phenomenalist view* of

---

38. Crowell has argued that Hare operates with two conflicting senses of 'desire', viz., one in which desires are understood in terms of assent to imperatives, and another where desires are "any felt disposition to action" (Crowell 1974, pp 160–161). However, it is unclear to me in what sense they conflict.

39. Cf. Hare (1987b, p. 73): "To want something to happen is to be in a state of mind which, if it had to be put into words, could be expressed by saying that one accepts the prescription that it happen. The acceptance itself does not have to be verbalized (dumb animals can want things); but if it *were* verbalized, that would be a way of doing it" See also Smart, who objects to Hare's suggestion, saying that "Certainly not all wantings or desirings are cases of accepting prescriptions" (in Pettit, Sylvan, and Norman 1987, p. 194). But as the above passage shows, Hare is not arguing that desiring *is* a case of accepting a prescription. Relevant here is also Locke (1981, pp. 537–538), who objects to Hare's view on desires, the main point being that "desires can conflict without involving us in self-contradicition or committing us to logically incompatible imperatives". However, the reason for this which he suggests is one that Hare would not have to deny, namely that "a person can genuinely want to do something and yet not do it, even though he is able to, because he has a stronger desire or more reason to do something else instead."

40. Various writers have pointed out that the term "dispositon" admits of at least two interpretations (e.g. Armstrong 1968; Wiggins 1978/79; Persson 1981).

dispositions, for instance, to be disposed in a certain way is to be prone to be in a particular state when a particular condition would be realized. Thus, to say of an object *a* - a crystal vase, for instance - that it possesses a dispositional property, say (to take an often used example), brittleness, amounts to saying that *a* is liable or prone to break, whenever a particular condition obtains. In our example, such a condition would be realized if a hammer, for instance, fell on *a*.

In contrast, in what we may call the *realist* conception, to speak of *a* being brittle, entails that *a* is in a non-dispositional, internal state that possesses some characteristic *C* (e.g. a certain molecular structure) to the effect that a full sufficient causal condition for *a* breaking at a time *t* will be the facts that *a* at time *t* possesses *C* and that a certain particular condition obtained, such as it being hit by a hammer, etc.[41]

In both of these views having beliefs is not sufficient for the acquisition of a disposition to act. Whether I acquire a disposition or not, given that I have a belief, will be a contingent matter. What is needed in addition is, in the latter view, that I am in a certain internal state. What is required, according to the phenomenalist view, is on the other hand left unspecified.

Phenomenalism outlined as above is extensionally equivalent to realism. However, this does not hold for a third kind of view that should best be understood as being a special version of phenomenalism. With regard to the ascription of desires, it does not leave it unspecified what is required in addition to beliefs. Rather, it claims that beliefs are sufficient for the acquisition of desires. According to this view - which has been called the "desire, *qua* belief" view - to have a desire is to hold a belief that results in action.

The idea that beliefs are sufficient for the ascription of desire is unfortunate. The main reason is that it precludes the possibility of regarding beliefs as causes for action. The view at issue effectively cuts the ground from under any attempt to offer a causal explanation of why representing something gives rise to desires, and in ultimately action. For

41. Ryle's (1949) view on dispostions has been interpreted and convincingly criticized as expressing a phenomenalist view, by Armstrong (1968) and Persson (1981), who both argue for a realist view. Whether Wiggins (1978/79) should be understood as taking a stand for a realist view is an open matter. However, he is aware of the problems related to the phenomenalist view. Wiggins also refer to the work of D'Alessio (1967) and his D.Phil. (1968) (Oxford) on dispositions. Unfortunately, I have not read D'Alessio's work.

this reason, I will assume in what follows that Hare is not endorsing something like a "desire *qua* belief" view.[42]

Now, consider the following passage from *LM*, hich explains what is involved in assenting to a first-person command:

> we are said to be sincere in our assent if and only if we do or resolve to do what the speaker has told us to do; if we do not do it but only resolve to do it later, then if, when the occasion arises for doing it, we do not do it, we are said to have changed our mind. . . .It is a tautology to say that we cannot sincerely assent to a second-person command addressed to ourselves, and at the same time not perform it, if now is the occasion for performing it and it is in our (physical or psychological) power to do so (*LM*, p. 20. Cf. *FR*, p. 79).

Some lines further down he turns to third-person commands, the assent to which he characterizes as to "join in affirming". Unfortunately he does not make it clear here or elsewhere what exactly he has in mind with this.

In the light of what Hare says about what is involved in assenting to a first-person command such as 'Let me do X!', it seems reasonable to interpret Hare as endorsing the following:

To intend that a person *b* should do a certain act X, is to have an intention which, if expressed in words, naturally would take the form of the first-person imperative 'Let me bring about that *b* does X' (cf. Hare 1987b, p. 73). And an assent to such an imperative would in turn involve doing something which brings about that *b* does X, if, as Hare himself puts it, "now is the occasion for performing it" and it is in the prescriber's (physical and psychological) power to do so.

Let us add this interpretation of what Hare appears to mean by "person *a* intends that the addressee should do the act prescribed" to the two former theses, and let us call it (**I**):

> **(P)** A speaker *a* performs a prescriptive speech act in uttering the sentence (s), only if *a* by uttering (s) intends to tell the addressee(s) to make something the case – where this something is determined by (s) and where the term 'tell' should be understood in its generical sense.

---

42. There is the possibility that we may have a notion of desire that admits of more than one interpretation. The scope of the "desire, *qua* belief" view should perhaps not be taken to cover all desires. Such a possibility has been seriously considered by e.g. McNaughton (1988). However, a hybrid theory to that effect seems theoretically unsatisfactory. Anyone endorsing such a mixed account would have to explain why a desire *qua* belief should not be considered an anomaly in a theory that also claimed that beliefs are not sufficient for the acquisition of desires. This would, of course, work the other way around too.

(SC) A speaker *a* prescribes sincerely only if *P* holds, and only if he intends that the addressee(s) should do what is prescribed.

(I) For *a* to intend that *b* should do what is prescribed (where *a* and *b* can be the same person), is for *a* to do something which *a* believes will bring it about that *b* does X (if it is in *a*'s physical and psychological power to do so).

The above division distinguishes between the intention to tell someone to make something the case, and the intention to do something which will bring about what is prescribed. The fulfilment of the former intention takes place when the speaker in uttering a certain sentence (viz., one which has a prescriptive illocutionary act potential) tells someone to make something the case. This much we may, I believe, safely say of this intention. Thus, according to Hare, to issue a sentence such as 'You ought to do X', counts by virtue of its illocutionary act potential as a way of telling people to make something the case. If you have the intention to tell the person to do X (which you can have without necessarily wanting him to do X), then one way of fulfilling this intention is to say 'You ought to do X' (Whether you choose an 'ought'-sentence or an imperative sentence form depends, to simplify the matter somewhat, on whether you want to make a universalizable judgement or not).

Given Hare's view of having a desire and assenting to an imperative, something more could perhaps be said about the intention to tell someone to make something the case. By extrapolating, this intention could be understood as an intention to make the hearer of the utterance *recognize* that if he assents to the judgement, say, 'You ought to do X', he (i.e. the addressee) must also assent to the prescription 'let me do X' (where 'me' refers to the person assenting to the prescription). An insincere person, for instance, may intend to make the hearer recognize this without wanting the person to comply with the act. Perhaps he expects the hearer to act contrary to the prescription, which is what he wants. Still, this is extrapolating, and the criticism that I believe can be directed against (*P*) does not rest on it being interpreted in the above way. Grice (1968), Schiffer (1972), Searle (1969, 1979) and other writers have attempted to provide an adequate basis for an analysis of illocutionary acts. Austin, for instance, maintained that in order to be performed an illocutionary act must "secure uptake" (Austin 1962, p. 116. Cf. Cohen 1964; Wetterström 1977, p. 21). A speaker cannot be said to be warning, for instance, unless his audience hears what he has said. For my present purposes it is not necessary to examine in detail to

what extent Hare would agree with this or whether he thinks some other condition to be necessary.

To fulfil the latter intention, viz., the intention that the addressee(s) should do what is prescribed, it is sufficient that the speaker does something, once the occasion arises, which he or she thinks will bring about what is prescribed. An insincere speaker or an actor can utter the words 'You, bring me a glass of water', and if the speaker knows the meaning of this sentence (knows, among other things, that it has a prescriptive illocutionary act potential), he must utter this sentence with the intention of telling the person to do something. However, the speaker, being insincere or acting, does not have the intention to do something that brings about what is prescribed.

Thus, to sincerely utter the judgement, say, 'You ought to keep your promise', as an evaluative judgement, is to (intend to) tell to the addressee to keep his promise (prescribing), and to perform such a speech act sincerely involves an intention that he keeps his promise, and again, this is a disposition to do things to this effect when the occasion arises. Moreover, to utter it in an *inverted-commas* way is to issue it with the intention to tell that something is the case. And a condition for sincerly uttering such a thing is to have a belief that that something is the case.[43]

2.8 Past Imperatives and Impossible Intentions

In the two ensuing sections of this chapter I wish to address some objections that can be directed against Hare's claim that value-judgements primarily are, in the sense outlined above, prescriptions. Moreover, let me pay attention especially to 'ought'-sentences. There are mainly two general objections that interest me, namely:

---

43. Cf. Ibberson's interpretation of "sincere prescribing", with which I am very much in agreement. Consider the following, for instance: "But what is it to intend that another person or an inanimate object should do something? If I am playing golf and slice the ball into the trees, I might naturally say that I didn't intend the ball to land in the rough, I intended it to land on the green. But surely this can only mean that I intended *to do something* which in the circumstances would result in the ball landing on the green. In this case my intention was unfullfilled because my action did not have the intended result. Now on the strength of this single paradigm I shall risk advancing the following generalization: to intend that some event occur is to intend to do something which in the circumstances will result in that event occurring (unless the event is a 'basic action', such as raising one's arm usually is, for which there is no other action we perform in order to bring it about" (Ibberson 1986, pp. 35–36).

(i) Hare's analysis of what is involved in sincerely prescribing requires us to have intentions that we cannot possibly have.

(ii) 'Telling someone to make something the case' requires an audience, which is absent in some cases of 'ought'-judgements – notably in third-person 'ought'-judgements.

Let me begin by considering (i). It follows from Hare's universalizability-thesis that anyone who issues a singular moral judgement in its typical sense, must also assent to the universalization of this moral judgement. From the universal principle we can derive imperative sentences "in all persons, as well as in all the tenses" (*LM*, p. 188). In other words, if I hold, say, 'You ought to keep your promise', I am committed to hold that any person X who is similar or relevantly similar to the addressee (including being in similar circumstances), also ought to keep his promise. To prescribe sincerely that X ought to keep his promises involves in its turn to have an intention (not necessarily occurrent) to do something to bring it about that anyone to whom the principle applies, now or in the future or in the past keeps his/her promises. But what are we to make of past prescriptions?

Consider the following judgement: (s) 'You ought to have done X'. *Prima facie* such a judgement seems to create problems for Hare's analysis in two ways. First, it appears to undermine Hare's **P**-thesis. Is there any plausible sense of 'telling someone to make something the case', such that we could be said in uttering (s) to be telling the addressee to make something the case? Secondly, it challenges the sincerity condition **SC**, in that it seems implausible to require of the speaker that he must have, in sincerely issuing (s), the desire or intention to make the addressee change the past. Since I will focus attention in the next section on cases that pose a problem for the **P**-thesis, I will here mainly be concerned with the latter kind of objection.

John Ibberson is perhaps the one who has argued most strongly that Hare's prescriptivism requires too much of us. He finds rightly the idea of such past prescriptions obscure (cf. Braithwaite 1954, p. 258; Wellman 1961; Cooper 1973; Hamblin 1987, p. 32; contrast Bosque 1980):

> To say that it is morally wrong to commit suicide is, according to Hare, to prescribe universally that *no one* at *any* time (past, present, or future) commit suicide. To prescribe this sincerely is to *intend* that no one in the

> past, present or future commit suicide. But it is logically inconsistent to intend that no one ever has committed suicide when you know perfectly well that people already have (Ibberson 1986, p. 36; cf. Najder 1975).[44]

Hare agrees with Ibberson that we cannot obey past imperatives (Hare 1979a, p. 164) but that we still should "put up with them" (*LM*, p. 188). Moreover, he has expressed the view that we should regard past imperatives as a "by product" (Hare 1979a, p. 164).[45] Unfortunately Hare does not say anything further about the kind of "cannot" which he thinks is involved, other than it is not:

> necessary here to go into the vexed question of whether the impossibility of altering the past, which is the reason why past-tense imperatives lack a use, is a logical or conceptual impossibility or a causal or physical impossibility (Hare 1979a, p. 164).

Can I have a desire to do something which is logically impossible? Could we, for instance, have a desire to build a square circle, if we had knowledge of the properties involved? Could I - to consider the second alternative that Hare mentions - have a desire actually to jump to the moon, although I was perfectly aware that it was a physical impossibility for me? In order to

---

44. Najder's book contains serious misunderstandings. He overlooks the fact that Hare uses 'command' in a generical sense (pp. 99–100), and that Hare distinguishes between telling and getting someone to do something (which enables him to criticize Hare for presenting the "world as filled with the clamour of importunate despots and herdsmen, who thirst to influence actively the behaviour of their neighbours", p. 100). Notwithstanding, Najder correctly sees that past actions are a problem for someone like Hare (p. 110). Wellman (1961), on his part, claims that "since it is impossible to change the past, it makes no sense to tell someone to have done something. Meaningful direction is limited to actions which lie in the future at the time when the directive is uttered" (p. 243). Cooper in his turn says "Retrospective moral judgements are first and foremost judgements. They have affinity not so much to the issuing of commands as to the making of statements and the application of criteria" (p. 232). Finally, Braithwaite (1954) says about past imperatives that they are "logically impossible" (p. 258). The idea that imperatives are future-bound has been questioned by Bosque (1980), who argues that in Spanish it is possible to construct perfect-tense grammatical imperatives. These constructions behave like imperatives (with the exception that they do not accept "por favor"), although they refer to the past. Bosque's examples are, however, all sentences that equally well could have been phrased as 'ought'/'should'-sentences (the English "You should have come yesterday" can take the spanish form "Deberias haber venido ayer" as well as the grammatical imperative "Haber venido ayer"). But, as Hamblin (1987, p. 50) points out, this suggests that those alleged past imperatives rather should be considered as carrying deontic than imperative force.

45. C. L. Hamblin is another writer who considers past imperatives to drive a wedge between the logic of imperatives and the logic of ought/should: "Retrospective *should*-statements and strictly present ones occur regularly where the corresponding imperatives would be impossible" (Hamblin 1987, p. 125). For a review of Hamblin's impressive work, see Ryding 1988.

answer these questions we need a much more detailed account than the *I*-thesis gives us of what a desire/intention is. Personally, I would say that a way of describing the above examples is to say that rather than involving intentions to do the logical or causally impossible, I am *wishing* to do the impossible. Moreover, I would say that what differentiates a wish from an intention/desire is that the latter but not the former conceptually involves a belief to the effect that it is possible to do that which you are disposed to do. Wishes are to be distinguished from intentions/desires in that the person who is wishing something believes that this something cannot obtain - where 'cannot' here refers to both logical as well as physical impossibility. It seems, for instance, perfectly correct to say, 'It is impossible for me to jump to the moon, but I sure wish I could'. However, it would sound odd to say 'I intend to go to the moon on a bicycle, although I know that I cannot'.

I find, in other words, that the idea that I can be said to intend/desire what I believe that I cannot do, runs counter to what it means to have a desire/intention to do something. Therefore, as long as we are required to have intentions to change the past - regardless of whether it is logically or causally impossible - I do not see how we can put up with them. I agree with Ibberson[46] that given the above interpretation, past imperatives are indeed an obstacle to prescriptivism.

There is a possible way by which Hare could meet the above objection. From the passages quoted earlier, we have seen that Hare indeed speaks of having "intentions" as a sincerity condition. As was outlined in section 2.7, Hare's view of desires, intentions, and wants is not sufficiently detailed to make it possible to separate these motivational states from each other. Hare might, for instance, meet the objection from past prescriptions by claiming that all that is required of the person who utters, say, 'You ought to have helped him' is that he *wishes* that the addressee had helped him. However we want to analyse our notion of "wishing", it must clearly be consistent with the idea that we can wish that the past was different. There is nothing odd about an expression such as 'I wish I had met you yesterday'. To consider past 'ought'-judgements as expressing wishes would enable

46. In *LM* Hare made an analogy between imaginary numbers and past imperatives, which later was defended in Hare 1979. In Ibberson 1986 (p. 37) he argues, in my view successfully, that imaginary numbers are not to be understood as numbers. He says about "j" – which he uses as a symbol of an imaginary number – that it "does not denote any number at all. To say that j is not a real number is simply to say that it is not really a number". See also Hare 1987 in Pettit *et al* 1987, p. 77.

Hare to retain the idea that such judgements are not merely the expression of our beliefs.

Hare might in addition even want to argue that wishes involve dispositions. There is a sense, it might be argued, in which not only desires and intentions but also wishes involve dispositions to act. A person could, for instance, in a reasonable enough sense be said to be disposed to do some such act as changing the past or, generally, to do what he believes is impossible. This would be the case when the person has a disposition to want something, if he could do this something. An analysis of wishing in these terms claims, in other words, that wishing is having a second-order disposition to want something, *if* some non-actual condition obtain (such as the fact that I believe that I could do that which I am disposed to want to do).

Whether or not such an analysis captures anything essential about wishing, Hare could maintain that in order to comply with the sincerity condition for prescriptive speech acts, the speaker needs at least to wish that what is prescribed also should obtain.

It is unclear what Hare's view is on these matters. He has in a different context distinguished, as will be seen in a moment, between wants and idle wishing. However, what he says there does not suggest to what extent he is ready to endorse the above defence.

Now having said this I venture that in the absence of a more detailed account of desires/intentions on Hare's part, it is reasonable to take sides with Ibberson. Past imperatives are indeed an obstacle to prescriptivism, and Hare's views on this issue need a revision. I do not see how we could be said to have intentions to act - given the fact that we knew that it was impossible to for us to do the act. Moreover, if it is unreasonable to require of us to have such intentions for changing the past, as a sincerity condition, we have reason to question also that the illocutionary act performed in issuing past 'ought'-judgements belongs to the genus of 'telling someone to make something the case'.

We do not have to turn to past imperatives to find examples that pose difficulties with regard to the sincerity condition **SC**. For instance, consider the following sentence:

(a) You ought to hand over her wallet of your own accord.

The proviso here that the person should act of his/her own accord clearly suggests, I think, that it would be wrong to say that the speaker has an

intention that requires him logically to assent to the imperative 'Let me do something that brings about the voluntary act of person so-and-so'. Moreover, the idea that we actually are telling someone to act of his own accord seems highly questionable. The reason is that if the addressee performs the act prescribed and believes he does so because he is being ordered, requested, etc., he is no longer doing what was prescribed, viz., acting of his own accord.

It is not difficult to find cases that *prima facie* run counter to a prescriptivist interpretation. Such cases could perhaps be considered as prescriptions, but the question is whether this is the most reasonable analysis. There are, for instance, several states of a person, the presence of which the person cannot deliberatedly arouse or evoke. Consider, for instance, the following sentences:

(1) You ought to feel ashamed.
(2) He ought to believe her.
(3) Their child has been killed, and they are not sad. I say they ought to be sad.
(4) You have everything that anyone reasonably can wish for, and still you are not satisfied. You ought to be glad.

There is, of course, a way in which these sentences could be said to be prescriptions in Hare's sense. Take example (3), for instance: It could be understood as advice, a request, or even an order that the persons should show *signs* of sadness. (2), in its turn, could be issued as a request that the person, for example, ought to consider and thoroughly examine her arguments. However, let us close the door for such uses that it, admittedly, could be argued exist, and consider the possibility of (2) and (3) being requests, orders, etc., that the persons (should) be in the state of sadness and belief respectively. Such an interpretation faces difficulties for two reasons. First, these sentences contain predicates that are not easily associated with the imperative sentence form. Although there is nothing wrong grammatically with 'Be sad' or 'Believe her', they seem odd. The reason for this is that, as Hamblin points out, these *stative* imperatives "enjoin states to get into", rather than "actions to perform" (Hamblin 1987, p. 54). Although the addressee assented to the request, order, etc., there is no guarantee that the person actually would be sad or would believe. Although he does not develop the idea, Taylor (1965) is another writer who notices that there is a problem here. Thus he says:

> the object of the prescription, that f should be brought about, may not be the sort of thing which could be the intentional object of any of the above mentioned speech acts [order, request, entreaty, piece of advice, etc.]; one's reason for going to the play might be that one would enjoy it, *but one could not prescribe in any ordinary sense of that term that one should enjoy it* (Taylor 1965, in Wallace & Walker 1970, p. 63; my italics).

Secondly, we have reasons, I think, for questioning that a sincere utterance of any of these sentences involves, on the part of the speaker, intentions to do something - especially if we take them to be naturally expressed by an imperative such as: 'Let me do something which brings about state X' - where X refers to sadness, etc. An assent to the latter appears to presuppose a belief that one can bring about X. However, such a belief is simply not sound. That is to say, one can, of course, do something that one believes will have as a result that the person becomes sad. Consider, for instance, (3). By robbing the parents of something they liked, for instance, we could make them sad. However, it seems clearly far-fetched to maintain that the desire involved in issuing (3) is a desire that the parents should be sad *simpliciter*. After all, what (3) is about is that the parents should be sad *because* of what has happened to them, and not merely that they should be sad.

If the above examples of judgements enjoining states rather than actions are not easily analysed as prescriptions, what kind of illocutions are they examples of?

Some of them, I venture, may be regarded as expressions of wishes. As mentioned earlier, Hare himself has distinguished between:

> two sorts of wishes, which do need distinguishing, namely those which are naturally expressed in the imperative, and those which are naturally expressed in various 'optative' constructions (e.g. 'Would that I were a bird'). This supports Miss Anscombe's distinction between the kind of wanting of which 'the primitive sign is trying to get' and 'Idle wishing'. We might say that the latter is idle, unlike the former, because its expression, unlike an imperative, does not command any action (*PI*, p. 51, 1968b; cf. *PI*, p. 20).

Thus, in the light of this distinction retrospective ought-judgements, for instance, could be considered as being naturally expressed in the optative, say, 'Would that so-and-so had not happened'. Ibberson (1986), for instance, sees in this a way out of the difficulties that surround past imperatives. Thus, he maintains that:

> The most, I think, that Hare can reasonably require is that from moral

> judgements we should be able to derive imperative sentences in *some* of the persons and *some* of the tenses, and optative sentences in the rest. Otherwise no one could sincerely assent to any moral judgement (Ibberson 1986, p. 40).

The "optative" approach seems most reasonable in those cases where the intentions are, or at least seem to be for logical or causal reasons, of an impossible nature. However, not all problems for a prescriptive analysis arise, as we shall see, because of this. Let us therefore turn to the second main objection.

2.9 First- and Third-Person 'Ought'-Judgements

As may be recalled the second objection was the following:

> (ii) 'Telling someone to make something the case' requires an audience, which is absent in some cases of 'ought'-judgements - notably third-person 'ought'-judgements.

Intuitively, second-person 'ought'-utterances do seem to be examples of someone telling someone to do something. In other words, *P* seems to be a possible expression of what we do in issuing such a second-person 'ought'-judgement. However, whether or not one shares this intuition, there does not seem to exist such a similarity with regard to notably third-person 'ought'-utterances. The second objection centres, in other words, on the claim that Hare has not given us an explanation of what "telling someone to make something the case" is in those cases where our moral judgement is directed not to the listener but to some third-person. First-person 'ought'-utterances, on the other hand, pose problems on their own, and I will have occasion to comment also on these later on.

Consider the speaker *a*, who utters to an addressee, *b*, say, '*c* ought to pay a visit to his dying mother'. The oddity involved in this kind of utterance is that it is far from clear how we could be said to be telling somebody absent to make something the case. Moreover, it seems plainly wrong - with the possible exception of when *b* is among the addressees (as for instance in 'Everyone ought to do X') - to characterize the speech act as one of *a* telling *b* to make something the case. Naturally *a* might have the desire or intention to do something that made *b* see to it that *c* visited his mother, but given the above interpretation (**P**, **SC**, **I**) this is not sufficient for the

speech act involved to be characterized as belonging to the genus of prescribing.

Ibberson (1986), Hamblin (1987) as well as Clarke (1986) have recently focused on this matter.[47] Clarke, for instance, says:

> We can concede at least a partial similarity between the second person normative 'You ought to do A' and the imperative 'I advise you to do A'. Much less obvious, however, is any similarity for first and third person normatives (*op. cit.*, p. 89).[48]

When Hare enriches the imperative mood in *LM* with regard to tense and person (*LM*, chap. 12),[49] he is aware that first- and third-person imperatives are somewhat strange. Nevertheless, that we in our "ordinary grammar" do not include imperatives of this kind is, he thinks, a contingent fact (contrast Sellars 1963).[50] If we were omnipotent we would "make great use of third-person commands" (*PI*, p. 7, 1949).

Still, even if we got used to the idea of third-person commands, *qua* expressions of our intentions, there remains the problem of how to

---

47. See here also Braithwaite (1954) and MacIntyre (1957), who discuss the differences between first-, second-, and third-person uses of 'ought'. Other more recent writers that consider first- and third-person 'ought' utterances to pose problems for prescriptivism are Wetterström (1986, pp. 91–92) and Fumerton (1990, pp. 28–29).

48. Clarke wrongly ascribes to Hare the thesis that ought-judgements are to be assimilated to (what he calls) imperatives, which enables him to forward the following objection: "But the crucial contrast between normatives and imperatives lies in their susceptibility to justification. It is always appropriate for a normative to ask for its justification. . . .For imperatives, in contrast, justification is sometimes inappropriate." (Clarke 1985, pp. 89–90). That there is this difference has been argued by Hare in several places. See here, for instance *FR*, pp. 36–37, where Hare traces the difference between 'ought'-judgements and imperatives to the fact that the former but not the latter are universalizable. Finally, notice that Clarke in the quote refers to 'I advise. . .' as an imperative. I think this is an unfortunate use.

49. With regard to person, he says thus: "All we have to do is to take the phrastic of an indicative sentence in that person, and put after it the imperative neustic" (*LM*, p. 189). In Hare, 1949 he says: "The restriction as to person, which is in any case much less hard and fast, could be removed entirely if circumstance so required. We do not in fact use the first person singular, because we do not need to tell ourselves to do things, we just do them. . . .On the other hand, if we were omnipotent, and could command the obedience of all persons and all things, we should no doubt make great use of third-person imperatives" (Hare 1971, p. 7).

50. Sellars believes that "It is no accidental feature of imperatives that one can only tell people to do things in the future. And it is no accidental feature of imperatives that they are formulated by the use of tensed verbs in which the tense has its full and ordinary force" (Sellars 1963 in Castañeda & Nakhnikian 1963, p. 180). Here I part company with Hare. Sellar's article is one of the earliest critiques of Hare. In the article mentioned he launches an attack on Hare's examples of imperative inferences.

characterize the speech act involved in issuing an 'ought'-judgement that entails such a command. One possibility that was mentioned earlier would be to consider third-person 'ought'-judgements as being expressions of wishes. The speech act involved here would then not be one of telling someone to make something the case, but would rather be similar to those performed in uttering various 'optative' constructions such as 'Would that *c* visited his dying mother'. However, since it is unclear whether Hare would endorse such an interpretation, let me here omit a further examination of this possibility.

Hamblin (1987), who has done much to elucidate grammatical as well as functional imperatives, considers the idea that third-person commands

> are to be looked at for what they require of their addressees rather than for what they require of their subjects (Hamblin 1987, p. 58).

If *a* told *b* '*c* do so and so', we could, Hamblin maintains, sometimes understand it as a "elliptical plain-predicate imperative", viz., '*b* bring it about that *c* does so and so'. However, even if he thinks there would be something to be said for such an idea, he believes that it

> would be literally untrue of even some quite simple imperatives, such as those in which the subject is among the addresses and needs to be mentioned only in order to be picked out (Hamblin 1987, p. 53).

But perhaps Hare thinks that to issue a third-person 'ought'-judgement, say, 'He ought to return to his children', is among other things to tell the addressee to bring it about that the father returns to his children.[51] In other words, if I were to say 'He ought to return to his children' to *b*, I would actually be prescribing to *b* to bring it about that the father returns. However, such an idea seems far-fetched for several reasons. First, if you wanted *b* to bring it about that the father returns, why go to the trouble and issue a third-person 'ought'-judgement? Why not say simply 'See to it that the father returns'? After all, *b*'s action does not necessarily, it seems, have to be morally required. Or, in case *b*'s action also was considered morally obligatory, such an analysis would in effect analyse the above third-person

51. There is a further possibility. Consider the following: "I say that moral judgements are prescriptive because in their typical uses they are intended to guide our conduct; to accept one is to be committed to a certain line of action or to prescribing it to somebody else" (*ET*, p. 44). I suspect that the latter part concerns assent to third-person judgments. However, since Hare does not further elucidate this idea I will not concern myself with it.

judgement as being equivalent to the following '*b* ought to make the father return to his children'. It seems clearly unreasonable to say that in issuing the third-person judgement, we are actually issuing the second-person judgement. Secondly, and more important, it just does not seem to be case that we require an action of the addressee when we issue a third person 'ought'-judgement to him/her. For instance, it seems to make perfect sense to say:

> 'Tom, I know that you do not want any help, and I accept that. But still, Charlie ought at least to ask you if you need a hand. After all, you have done so much for him over the years.'

Is the speaker here telling Tom to do anything? Does his utterance logically require assent to the imperative 'Tom bring it about that Charlie offers you his help'? But why should he prescribe such a thing if he respects Tom's desire to do it on his own. It does not make sense. But can he then hold the judgement that 'Charlie ought . . .'? Yes, since *ex hypothesi* what he wants is that Charlie on his own initiative offers Tom his help - even if he knows that Tom will not accept it. Take another example:

> The only flowers you get, Mary, from your husband are the ones you ask him to buy to you. He ought to *surprise* you, and one day bring you some flowers on his own initiative.

Surely the speaker is not here telling Mary to do something which in the end will surprise her. How does one go about doing such a thing?

A further consideration should make it clear that such an analysis of third-person 'ought'-judgements is implausible. Often third-person 'ought'-judgements are issued as answers to questions of the form 'What ought *c* to do?' - where *c* is a third-person. It does seem strange to think that the answer (at least part of it) to such a question would consist in a prescription to the effect that the questioner should do something.

The above considerations support my conclusion that there are strong reasons for saying that the speech act involved in issuing third-person 'ought'-judgements cannot be said to belong to the genus of prescriptive language. The idea that we are actually telling the addressee to make something the case when we sincerely utter a third-person 'ought'-judgement, seems in any reasonable sense of "telling someone to make something the case" untenable.

What should we then say about first-person 'ought'-judgements? Here I

think it is important to distinguish, albeit in a very general way, between at least two different kinds of context in which a person may use a first-person 'ought'-sentence. First, there is the person who speaks to himself, being aware that he does so. Secondly, there is the case in which the first-person 'ought'-judgement occurs in a conversation between two persons.

Dunn, for instance, who calls attention to the former case, considers first-person prescriptions/directives to be not "*bona fide* or serious" prescriptions/directives but rather:

> self-dissociative exercises of the imagination in which individuals imagine that they are issuing directives to someone *else* (or that one part of themselves is issuing directives to a *different* part of themselves, or some such; cf. Hare *FR* 5.8, p. 81) (Dunn 1987, p. 61. A similar view is expressed by Braithwaite 1954, p. 257).[52]

Carl Wellman, however, believes that "the director and the directee might be the same person", but adds some lines later on that:

> Such reflexive directing seems to be a degenerate case, however, for it is not clear how I could command, request or advice myself to do something (Wellman 1961, p. 234).

In contrast, Neil Cooper (1973) does not seem to find such reflexive prescribing/directing too strange. Thus, he says:

> In telling myself that I ought to behave in a certain way I am trying to get myself to do what in a sense I want to do. Of course, what I am telling myself that I ought to do is almost what I don't want to do, otherwise there

52. Dunn (1987, p. 61) also argues that Hare misses that directives only express the intention that someone else should do something: "Thus, although it is correct that imperatives are suitable sentences in which to express certain intentions or desires, those intentions or desires do not include intentions or desires to do something *oneself*: they include, rather, only intentions or desires that someone else should do something". Unfortunately Dunn does not specify what he means here by "something *oneself*". On one interpretation, Hare could reasonably say that issuing, for instance, 'You, do x' requires at least some kind of disposition to act, viz., being disposed in a wide sense, to do something when the occasion arises in order to make the person do the act required. However, Dunn appears erroneously to think that the speaker, according to Hare, in issuing 'You, do X' expresses not only the intention that the other person should do X but also the speaker's intention to do, himself, X. Such an intention will in fact be logically required in the case of, say, 'you ought to do x'. But the reason for this derives from the fact that 'ought', according to Hare, is universalizable. This apart, Dunn gives good reasons for why commending is not a prescriptive speech act. However, since Hare has made it reasonably clear that he regards commending as a species of "telling someone to make something the case", I doubt whether it will have much impact on Hare. Dunn's point will at the most lead to a verbal manoeuvre (see Dunn 1986, pp. 41–43; cf. Hartland-Swann 1960, p. 112).

> would be no necessity for me to tell myself that I ought to do it. This is why it is the morally weak man in particular who needs to tell himself what he ought to do (Cooper 1973, pp. 231–232).

Cooper seems, however, indirectly to endorse the view that we are prescribing to some other part of ourselves, namely the one who does not want to do the act which we prescribe. Whether or not this is the case (Cooper's view, I must confess, puzzles me), I agree with the former writers that first-person prescriptions/directives are not *bona fide* prescriptions. Be that as it may, the problem of how to account for "inner conversations" is not exclusively a matter that troubles prescriptivism (cf. *PI*, p. 85).

We get a more interesting perspective if we focus on the latter kind of context. A first-person 'ought'-judgement can, of course, pop up in a conversation for various reasons. The man who is asked 'Why are you going to visit Fred?', and answers 'Because I ought to visit him', is, it seems, giving the questioner (the first part of) an explanation, a reason for why he is visiting Fred. Another situation would be the following:

> Person A: 'I have promised to be back at 5 p.m., and I ought to leave now.'
>
> Person B: 'It is funny, I also have an engagement at 5 p.m., and I ought also to leave now.'

What kind of speech acts are the above examples of? Rather than being cases of someone telling oneself to make something the case, they seem to be examples of someone passing on information to a hearer, as a reason for something in a way that is natural to the assertive, but not prescriptive mood.

By stressing that the speech act of prescribing is one hallmark of evaluative utterances, Hare exposes himself to the above objections. How much weight do these objections carry against a prescriptive analysis of moral 'ought'-judgements? Let me now address this question in somewhat greater detail.

2.9.1 A Counter-Objection

To try to answer the above question let me concentrate on the objection that third-person 'ought'-judgements constitute an obstacle to a meaning-analysis of 'ought' in terms of **P** and **SC**.

Hare maintains, as has been shown, that when we issue a fully evaluative moral judgement, we are prescribing. I have tried to show that especially

third person judgements are not cases of prescribing. If I am right about this then one consequence for Hare, who explains meaning in terms of illocutionary speech acts, would appear to be that he has to say the following: 'ought' in 'You ought to do X' has a different meaning from 'ought' in 'He ought to do X'.

A feature of the criticism put forward above is that I have assumed all the time that Hare uses "telling someone to make something the case" in an ordinary sense in which it refers to that (genus of) illocutionary act which we perform in issuing imperatives. Nowhere have I found any explicit trace to the contrary. However, the possibility cannot be excluded that Hare may have a wider notion of "telling someome to make something the case" than I have here assumed. Hare might, for instance, agree that it sounds odd to say that one can tell somebody who is absent to make something the case. Still, the speech act can be one which is similar to the one of imperatives, in that it is a sincerity requirement in such cases that the speaker has a desire that the subject of the prescription performs the action prescribed. Hare could even say that *prescribing is precisely that speech act, the definition of which requires such a desire.*[53] Let us call this the *wide account of prescribing*. Such a move would seem to meet the objection that we cannot be said to be prescribing when we issue a third-person 'ought'-judgement.

I grant Hare this possibility. I want to stress I do not think that the objection from third-person 'ought'-judgements conclusively refutes the idea that we are logically required to have certain desires (viz., those specified by *SC*), when issuing 'ought'-judgements. Hare can, to put it differently, retain the idea that the speech act we perform in issuing moral judgements involves desires on the part of the speaker that the addressee should do something. Notwithstanding, I suspect that such a wider interpretation will constitute a drawback for Hare's case that it is part of the meaning of 'ought' that the issuer of an 'ought'-judgement is required to have a desire that the addressee should do what is prescribed. The wider notion of "telling someone to make something the case" will no longer in any obvious way lend plausibility to the claim that issuing an 'ought'-judgement sincerely requires of us to desire that the addressee should do something. It is reasonable to say that the sincere issuing of an imperative requires of the speaker a desire that the addressee should do something.

53. I owe this point to Ingmar Persson. Cf. Gowans (1989), who by "prescriptivism" understands a theory that "asserts that there is an important connection" between accepting a practical judgement and acting on it (p. 189).

But if "telling someone to make something the case" is not to be understood as that speech act which we perform in issuing a second-person imperative, invoking the wider notion will not make it obvious why such a desire must be present. A similar point can be made with regard to the claim that telling someone to do something "is answering the question "What shall I do?" (*LM*, p. 15).

To find a remedy for these (and I suspect other) drawbacks will require, if undertaken within a speech-act theory of meaning, another formulation of *P*. Hare must in other words give us a characterization of what we are doing in issuing a third-person 'ought'-judgement. Whether such an alternative formulation will handle the problems which third-person 'ought'-judgements constitute, and at the same time carry explanatory value with regard to the two issues above, remains to be shown.[54]

2.9.2 Imperative Inferences

Let me finally round off this chapter by commenting briefly on an important issue that I somewhat reluctanly sidestep, namely Hare's view on imperative inferences. That there can be such inferences was argued in his early work "Imperative Sentences" (*PI*, 1949), where he concentrated on purely imperative inferences, and in *LM*, where he in addition turned his attention to mixed inferences, i.e., where the premisses are in the imperative as well as in the indicative mood. Whether there can be such inferences is in itself an interesting question. However, in the case of Hare, it is safe to say that his primary interest in this matter derives from his claim that value-judgements entail imperatives. We do reason and infer conclusions in terms of moral terms such as 'ought'. Therefore, it becomes necessary for Hare to show, given his imperative analysis of value-judgements, that there can also be imperative inferences. This gives rise to a question concerning what should count as valid imperative inferences. Imperatives do not admit of the ordinary truth values 'true' and 'false'. Just

54. Could we not say that to issue, say, "Tim ought to ..." the speaker had the *intention to express his resolution or desire that* Tim should do something? Braithwaite (1954), for instance, has attempted such an interpretation of Hare (cf. Schüller 1986, p. 83). The problem with such an analysis is that it is questionable whether "express" refers to an illocutionary act. For instance, I may express an opinion, belief. But, it seems, I will do this by performing notably speech acts belonging to the genus of assertives. Moreover, I can express my resolution to do a certain thing by, say, promising to do it. To reformulate **P** in terms of "Expressing" would not solve the problems I have mentioned. Relevant here is Searle (1979, p. 10). See also *MP* (p. 34) which contains an interesting passage where Hare connects the term 'express' to the Emotivists.

how we should analyse 'validity' with regard to imperative inferences, is still under much discussion.[55] It has been argued, for instance, that the attempt to apply classic logic of propositions to imperatives runs into serious problems. Some writers (e.g. Warnock 1976), have gone as far as to say that imperatives have no logic at all. The attempt to incorporate imperatives into the domain of logic faces more or less serious obstacles (Hare's attempt has been questioned by several writers, e.g. Bergström 1970; Stalley 1972; Carter 1982, and Hamblin 1987).[56]

Consider, for instance, the logical rule, which is accepted by most logicians, that we may infer 'p or q' from 'p'. Such a rule from the propositional calculus applied to imperatives seems to many counter-intuitive. Should we really, to take a much discussed example from the late Danish philosopher Alf Ross, accept that from 'Post the letter' we can infer 'Post the letter or burn it'. Such an inference appears to square badly with our linguistic intuitions.[57] Hare, who accepts the rule, does not deny this. However, he claims that the seemingly paradoxical nature of 'Post the letter or burn it' arises, not because the rule at issue does not apply to imperative logic, but because of certain general conventions of language. Hare claims that we generally presume that a person makes as strong a statement as he is in position to make. If we were to say 'Post the letter or burn it' we imply[58] that the listener may refrain from posting the letter, under the condition that he burns it. However, when we issue a command such as 'Post the letter' we do not imply the weaker command that he may refrain from posting the letter so long as he burns it. If the listener infers the latter from 'Post the letter' he has erred, not against the logic of imperatives but against the convention that a speaker makes as strong a statement as he is

---

55. Cf. Hare 1969 and his discussion of Kenny's "Logic of satisfaction". More recently Hare has expressed that "Much darkness has been shed by looking for surrogates for truth-conditions, in the case of prescriptions. To understand the meaning of the imperative mood-sign is, rather to understand what differences the use of an imperative makes to the communicative situation, and in particular what requirements are thereby incurred by the speaker and others" (Hare 1989, p. 36). For an attempt to outline a semantic theory for imperatives in terms of context of their utterance, see Chellas 1971.

56. See also Toulmin 1950, Turnbull 1954, Castañeda 1955, 1963, and Peetz 1981.

57. Hare's attempt to solve what has been called Ross's paradox is found in Hare 1967. See also Hare 1969. An examination of the discussion of Ross's paradox is found in Wedar 1985 (chapter XI) and in Johanson 1988. Relevant is also Sosa 1966.

58. Following H. P. Grice, Hare uses "conversationally implicate" (*PI*, p. 29) rather than 'imply'.

in position to do (explanations similar to Hare's are found in Hamblin 1987, p. 77; Føllesdal & Hilpinen 1971, p. 22).[59]

Another related topic of some controversy concerns the proper analysis of what is the contradiction of an imperative. Should we say, with Hare, that 'Do x' is the contradiction of 'Do not do x'? This does not seem counter-intuitive - at least it corresponds well with my own linguistic intuitions. Moreover, Hare maintains that in this regard imperatives behave differently from 'ought'-judgements. The contradictory of '*a* ought to do x' is not '*a* ought not to do x', but rather 'It is not the case that *a* ought to do x'. There is always this latter, third alternative, namely, to make a judgement of indifference, whereas such a possibility is excluded with regard to imperatives. Other writers, notably Bergström (1970), have suggested that in some cases the contradiction of 'Do x' is 'You need not do x' rather than 'Do not do x'.[60]

It appears clear that in comparison to the propositional calculus, there is much less consensus with regard to imperative logic. To condemn the attempt to lay down the rules of this logic as an impossible mission or a cul-de-sac, on the grounds that logicians do not seem to agree on what a valid imperative inference amounts to, seems to be a premature conclusion.[61] More light will hopefully be thrown on the issue of imperative inferences.

---

59. Hare has recenctly, in *PM* (p. 138) returned to this example of a conversational implicature (see note 58)

60. Ibberson in his turn claims that Bergström has a point, but only if by 'command' we understand order. If we, as Hare does, let it range over advice and requests Bergström is mistaken (Ibberson 1986, pp. 52–53).

61. Hamblin's book (1987) contains an extensive bibliography on the subject of imperatives.

III. Universalizability

3.0 Introduction

For nearly four decades, Hare's claim that moral judgements are universalizable has been under more or less heavy discussion. In part the discussion has concerned the nature of this claim. Hare's critics have either argued that it is a substantial moral principle (at least it contains some non-logical part) or that it is indeed a logical thesis, as Hare himself maintains, but therefore also trivial or at least of little importance to moral issues.

With regard to the universalizability-thesis, Hare's later works have clarified his position on at least two major topics. The first concerns whether the combination of prescriptivity and universalizability together serves to distinguish moral judgements from non-moral ones. In *FR* (p. 168), and especially in *MT*, Hare adds a third defining characteristic, viz., that moral judgements have the property of "overridingness" (*MT*, pp. 52-54).[1]

The second matter on which Hare's later views have undergone a change, or at least become more precise, concerns the role which the universalizability-thesis plays in moral reasoning. Hare thus makes it clear in *MT* that his attempt to show "how the requirement to universalize our prescriptions generates utilitarianism" (*MT*, p. 111) must be backed up by a further claim (cf. Allan Gibbard 1988 and Ingmar Persson 1988, 1989). Hare maintains in *HC* (p. 204), for instance, that his "derivation" of utilitarianism from the logical properties of moral terms requires what Gibbard calls the "Conditional reflection principle" (Persson speaks of the "principle of hypothetical self-endorsement").

Problems related to his method for answering moral questions will be discussed in the ensuing chapters. In this chapter I will present an interpretation of his universalizability-thesis. In addition to this I will comment on various objections to Hare's thesis that it is part of the meaning of moral words that they are universalizable.

---

1. Hare's earlier works have been criticized for omitting this feature of moral judgements. See for instance D.H. Monro (1967). Cf. also P. Singer (1974 p. 3).

In section 3.1 I present the two necessary conditions which a principle must meet, according to Hare, to qualify as a universal principle. Hare's explanation of why he thinks we are logically committed to universalize our moral judgements is the topic of section 3.2. This explanation centres on the notion of descriptive meaning, which evaluative judgements are said to carry besides their prescriptive meaning. Section 3.3 focuses on an important difference which Hare thinks exists between descriptive terms and prescriptive ones. Moreover I comment on the claim that all words can be said to be universalizable. In section 3.4 I am concerned with an objection that has been directed against Hare's claim that moral judgements are universalizable in precisely the same sense as descriptive judgements are universalizable. Moreover, I bring out what is the central difference between descriptive and evaluative judgements according to Hare. I discuss briefly in 3.5 some examples of objections directed at Hare's universalizability-requirement, which are based on a misinterpretation of this requirement. I follow up this discussion in the two ensuing sections, in each of which I consider an objection against Hare's claim that we are logically committed to universalize our judgements. Thus, in 3.6 I argue that there can be uses of 'ought' that are not universalizable. In 3.7 I look into an objection from J. L. Mackie against Hare's claim that the requirement to universalize our moral judgements, is a logical requirement. I argue that Mackie's objection is not successful. The remaining sections of this chapter are devoted to Hare's views on the so-called supervenience relation. After a general introductory comment on this topic, in section 3.8, I present in 3.9 his latest account of this relation. Finally, in 3.9.1 I consider an objection to Hare's account of supervenience.

Although implicit in *LM*, the idea that moral judgements are universalizable was originally presented 1955 in a paper entitled "Universalisability".[2] Apparently it was the first time the term was used in connection with ethical discussions (cf. Potter and Timmon 1984, p. IX; Rabinowicz 1979, p. 12).[3] Since then 'universalizability' has been associated with several different ideas such as, for instance, that moral agents should be impartial, or that we should treat our fellow men as persons or as ends and not merely as means, or that we should be fair, or ready to put

2. Notice the spelling. In later works Hare prefers 'universalizability'.

3. During a conversation in Ewelme, may 1990, Hare pointed out that the word 'universalizability' was used among philosophers in Oxford before the publishing of "Universalisability".

ourselves in the shoes of those persons whom our actions will affect. For some writers (e.g M. Singer 1961; K. Bach 1976-77) the requirement to universalize our judgement has been understood as the requirement to imagine what would be the case if *everyone* did the act under consideration. These notions have in their turn deep historical roots in western tradition. Several writers have pointed out that rather than speaking of one concept of ethical universalizability, one should speak of a family of concepts that to a lesser or higher degree resemble each other (e.g. D. H. Monro 1967, Brian E. Schrag 1975, Jan Narveson 1985, Bernard Williams 1985, J. L. Mackie 1985). Many writers agree, though, that there is a distinction to be drawn between *substantive* and *non-substantive* universalizability-principles.[4] Examples of principles that often are cited in the literature as being substantial universalizability principles are the so called Golden Rule:

> Do unto others as you would have others do unto you (One ought to treat others as one would wish them to treat oneself. Cf. *FR*, p. 34),[5]

and Kant's Categorical Imperative (sometimes referred to as 'Formula I' or 'Formula Ia'):

> Act only on that maxim through which you can at the same time will that it should become a universal law (Kant 1785, AK., IV, 421).

Another much discussed example is Marcus Singer's Generalization Argument:

> If everyone were to do that, the consequences would be disastrous (or undesirable); therefore, no one ought to do that (Singer 1961, p. 4).

---

4. For a recent attempt to distinguish between substantial and non-substantial universalizability principles, see Potter and Timmons (1985, pp. XII, XIII). See also Rabinowicz (1979).

5. Whether the Golden Rule is a substantive or non-substantive principle is still under discussion. Apparently Sidgwick took it to be non-substantive (or at least to admit of a non-substantive interpretation) in that he considered his own non-substantive principle to be the "abstract and universal form" of the golden rule. Hare also thinks that by suitable interpretation it could be turned into a formal principle (*FR*, p. 34; cf. Ross 1964, p. 125). Rabinowicz (1979, p. 14), for his part, considers it to be a substantive rule. As Potter and Timmons (1985) point out, there have been attempts to interpret the Golden Rule as being non-substantive (Singer 1963, and Blackstone 1965) and substantive (Gewirth 1978). Finally, reading Mackie (1981) I became familar with Bernard Shaw's paraphrase: "Do not do unto others as you would have that they should do unto you. Their tastes may not be the same".

The moral acceptability of a person's actions was judged by Kant in the light of whether or not its maxim can be willed consistently as a universal law. Singer's argument that if the consequences of everyone's doing some action x would be undesirable, then no one ought to do x, also sets forth, in its turn, a test that forbids certain actions as being immoral.

On the other hand, a utilitarian such as Henry Sidgwick formulated what some writers regard as the first universalizability principle (cf. Rabinowicz 1979, p. 12):

> if a kind of conduct that is right (or wrong) for me is not right (or wrong) for some one else, it must be on the ground of some difference between the two cases, other than the fact that I and he are different persons (Sidgwick 1901, p. 379).

Intuitively, Sidgwick's principle would seem to qualify as a non-substantive principle. This holds also for Rabinowicz' formulation:

> Moral properties of things (persons, actions, states of affairs, situations) are essentially independent of their purely 'individual' or 'numerical' aspects (Rabinowicz 1979, p. 11).[6]

In contrast to the categorical imperative and Singer's principle, the two latter ones are hypothetical in nature: They do not tell us anything about which acts are right or wrong (or which act has a certain moral property) but merely that *if* one judges one act as right (or wrong) then one cannot – on pain of inconsistency – refuse to judge a similar case as right (or wrong).

Each of these principles[7] raises two equally fundamental questions that concern their justification and application respectively. Much of the attention given to Hare's universalizability-thesis derives from the fact that he has, as perhaps no one else, justified his thesis by claiming that it is true by virtue of the meanings of moral terms.

---

6. Another example is Williams' (1985, p. 115) "enough is enough" principle, which he considers to be "the formal and uncontentious principle of universalizability". It says that "if a certain consideration is truly a sufficient reason for a certain action in one case, it is so in another". Williams must here reasonably be understood as saying at least in another similar case.

7. Principles referred to as universalizability principles have, of course, been put forward by many other writers (e.g. by Alan Gewirth, John Rawls). Moreover, it should also be noticed that some writers have attempted a hybrid theory, combining the formal requirement of universalizability with a substantive value-principle (see for instance Lansing Pollock 1973). For a survey of different uses of 'universalizability', see Jan Narveson "The How and Why of Universalizability" in Potter and Timmons (1985). Cf. Rabinowicz (1979).

Substantive principles, whether or not they are consequentialist or non-consequentialist in nature face all the classical questions that have haunted these respective positions. If the universalizability-thesis is a logical thesis it avoids this issue. Whether it does so depends entirely on whether or not we believe that our interest in the conceptual role played by the universalizable 'ought' requires certain dispositions. If we believe this to be true, then whether the universalizability-thesis, *qua* logical thesis, avoids questions of justification will depend on how widely shared these dispositions are. Moreover, if the universalizability-thesis in addition can be shown to be of at least some practical importance with regard to our moral reasoning, the gain would be considerable. Let us therefore now begin to examine what Hare means by universalizability, and how he explains it.

3.1 Hare's Universalizability-Thesis

As a preliminary statement I will understand "Hare's universalizability-thesis" as follows:

(U) An issuer of a singular evaluative judgement is logically committed to universalize his judgement.

Now it is my contention that Hare endorses two major lines of argument in order to establish his universalizability-thesis. The most elaborate of these arguments can be regarded as an attempt to secure consistency in the use of value-terms. The other argument, in its turn, results in a restriction with regard to what is admitted in a universal principle. Thus, in contrast to certain writers (e.g. Cottingham 1983), I regard Hare's universalizability-thesis as not merely a consistency-requirement. These arguments are to be viewed as establishing two restrictions which judgments or principles must meet, according to Hare, in order to be universal. In "Principles" (Hare 1972c), for instance, Hare maintains:

> I think that many people who have done elementary formal logic would call the following proposition universal 'All John's sons went to Harrow School between 1932 and 1942'. They would call it universal because, if formalized, it would start with a universal quantifier. But this, though necessary, is not sufficient to define the stricter sense of 'universal' which is needed in moral philosophy. The proposition contains the expressions 'John', 'Harrow', '1932' and '1942' which either are, or require for their defintion, references to individuals otherwise than by description. In the stricter sense of 'universal' a universal proposition is not allowed to contain such terms (*ET*, p. 51–52, 1972c).

An idea to much the same effect is expressed more recently in *ET* (p. 74, Hare 1984b; cf. *ET*, p. 51-52)[8], where it becomes clear that in order to be universal, a judgement or principle must meet the following two requirements:

(U1) The principle must be governed by a universal quantifier.
(U2) The principle must not contain any ineliminable individual constants.[9]

Thus, **U1** and **U2** are the necessary restrictions which Hare places on what is to count as a universal principle. Intuitively **U2** is to be regarded as securing that an assent to a principle meeting **U2** does not contain ineliminable references to any individual. It must be noticed here, as Hare makes clear in "Universalisability", that token-reflexive expressions such as for instance 'me' and 'this' as well as proper names are not necessarily excluded by **U2**. Such expressions are admitted when they "occur in the context of words attributing similarity" (*MC*, p. 22, 1955).

Thus, if I were to say, for instance, (s) 'Charles ought to help Jim', I am logically committed to assent to the principle

(u) For any person exactly similar to Charles, it is the case that this person ought to help Jim or any other person exactly similar to Jim.

Hare often phrases the universal principle in terms of the relation "relevantly similar". However, since in *MT* and works posterior to it he has made it clear that he does not require anything stronger than "exactly similar", I will in this work entirely focus attention on this latter relation. Thus, he has recently claimed the following:

> I need for my purposes only one kind of universalizability, which I expounded most recently (but without change of view from my earliest treatment of the related notion of supervenience) in [Hare 1984b]. This is the doctrine that moral judgements made about one situation, etc., have, on pain of logical inconsistency, to be made about any situation which is exactly similar in its universal non-moral properties (*HC*, p. 268).

---

8. In "Universalisability" (in *MC*) he makes it analytic by virtue of the meaning of 'moral' that universal moral judgements do not contain individual constants. However, in *FR* this claim is qualified in an important sense, as we shall see later on.

9. For an attempt to argue that universalizability is more than a requirement for consistency, see Ward 1973.

Now, it may be argued that a principle formulated in terms of exact similarity will in effect only refer to an individual. For instance, consider a singular judgement such as (s) 'Charlie ought to satisfy my interest'. In order to hold this as a moral judgement the speaker must assent to some such principle as

(u) For any person exactly similar to Charlie, it is the case that this person ought to satisfy the interest of any person exactly similar to me.

Assume now the truth of what is often referred to in the literature as Leibniz's thesis about the identity of indiscernibles (*principum identitatis indescernibilium*). This principle states that *exact similarity amounts to identity*. Given the truth of this principle, two things which have the same properties will on this claim be identical. (u) would in such a case be trivialized to the effect that it would apply only to two persons, namely, the speaker and Charlie (cf. Nielsen 1985, p. 92).

Hare has recently made it clear how he would answer such an attempt to trivialize the universalizability-thesis. Thus, he maintains the following in *HC*:

> There are various stronger and weaker interpretations of Leibniz's principle of the identity of indiscernibles. I hold that it is true in some of its weaker versions, but false in some of its stronger. For example it is true that no two numerically different things can share all their properties, including those specifiable only by the use of singular terms; but it is false that no two numerically different things can share all their universal properties (*HC*, p. 284).[10]

Hare's position seems reasonable. Suppose we let the similarity relation range over only universal properties – keeping those properties out of it that require reference to an individual. In such a case I do not see why two things which have exactly similar properties must on such a claim be

10. Cf. Rabinowicz (1979), who sets out to show that a version of the universalizability principle can be set up which does not rely on any notion of relevance, but which nevertheless does not become trivialized by Leibniz's thesis. Rabinowicz's work is highly technical. For a short formulation of his argument see Rabinowicz 1985. It should be noticed that Hare's universalizability thesis is not trivialized by what Rabinowicz calls "local Leibnizianism" – the view that "*within a given world*, indiscernibilities amount to identities" (1979, p. 82). However, the stronger version discussed by Rabinowicz, which claims that "indiscernible individuals are never distinct from each other" (p. 81) – regardless of whether the comparison concerned is within a given world or across worlds – would trivialize Hare's thesis. See e.g. *MT*, pp. 112–113.

identical. Interpreted in the above way, I venture to say that the *identitatis indescernibilium* thesis is implausible.

It is important to notice that Hare conceives of the similarity-relation as ranging over only universal properties. Hare has made this point clear also in relation to how we should understand the expression "exactly similar to/like. . .". Such an expression must, he claims, be replaceable by what he calls a "universal term".

> wherever the expression 'like a' is used in a maxim, it must be possible in principle to substitute a universal term for this expression without altering the import of the maxim. This universal term may not already exist in our language; but if necessary we can coin a new universal term for the purpose (*MC*, p. 23).[11]

A universal term, we may presume (Hare does not give any further account here of what he means by a universal term), will then be a term that refers to only universal properties. Such a property is in turn a property "which is describable without reference to individuals" (*FR*, p. 107).[12]

At this juncture, a further idea should be mentioned. As is apparent from many of his examples, Hare lets the similarity relation range over relational as well as non-relational properties. Examples of the latter kind of properties are to have red hair, to be a human being, to be a stone. Relational properties, on the other hand, are, for instance, the property of being a mother, of owing money, of being a husband of a physician.

All of the above properties are examples of universal properties. But consider the following: Suppose you were to make a reference to an individual by saying 'X is the wife of President Clinton'. It seems reasonable to say that in such a case, you have ascribed to X a relational property, the description of which involves a reference to an individual, namely,

---

11. Hare speaks here of "like". However it is reasonable to assume that he would grant the expression 'exactly similar to'. In contrast to my account, Hill (1974) argues that Hare's universalizability principle 'can be interpreted to mean something like the principle of identity or Leibniz's law" (*op. cit* p. 48). He also says that Hare "means more by this principle than the principle of identity or Leibniz's law" (*op. cit.*, p. 49). This claim is puzzling, and I do not know on what ground Hill puts forward this suggestion.

12. Cf. Popper 1980. See, for instance p. 66 and the subsequent discussion: "An individual concept is a concept in the definition of which proper names (or equivalent signs) are indispensable. If any reference to proper names can be completely eliminated, then the concept is a universal concept".

President Clinton. Given the truth of the statement, we can say that X has a non-universal relational property (cf. Wetterström 1986).[13]

U1 and **U2'** are then the necessary conditions which a principle must meet in order to qualify as a universal principle. Now, if by the expression "universalization" in (**U**) below

> (**U**) An issuer of a singular evaluative judgement is logically committed to universalize his judgement.

we understand the process of transforming the singular judgement in the light of **U1** and **U2** (cf. *FR*, p. 219), (**U**) constitutes one way of expressing Hare's universalizability-principle.[14]

A caveat is here in place with regard to the expression "logically committed" above. What exactly is this relationship between the singular judgement and the universal judgement? Hare has throughout his work used different expressions to refer to this relationship. Thus, if we go back to *LM* we are said in issuing moral judgements to "signify" our acceptance of a principle (*LM*, p. 162). Hare has also maintained that in issuing a singular moral judgement we "seem to imply (in a loose sense)" that there is some universal principle. More recently he describes it as a case in which the issuer of the singular value-judgement "subscribes" to a universal premiss (*ET*, p. 71, 1984b). These expressions can be given several interpretations. The choice of expressions lends some support to the idea that Hare is suggesting that before issuing a singular judgement, a universal principle occurs in the content of the deliberation to the effect that it is part of our conscious thoughts. Let us say that a person who issues a singular 'ought'-judgement (s) *phenomenologically* subscribes to a universal principle

---

13. In developing his notion of universal statements and universal properties Wetterström expresses several interesting ideas. See, for instance, his distinction between "intrinsic properties", "qualities", and "relational properties" (p. 24). See also what he has to say about "unique" universal properties (pp. 25–28). Contrast Schrag who appear to deny the existence of non-universal properties: "Any characteristic or property at all is universal, not particular" (1975, p. 62).

14. Wetterström suggests that Hare's universalizability-thesis should be understood as, what he calls, a "linguistic background-convention". He characterizes such conventions as "rules of speech, violations of which constitutes abuses of language in Austin's sense (1962)" (Wetterström 1986, p. 37). However, he formulates Hare's universalizability-principle in a way that is alien to Hare's theory, namely, as an imperative instruction (see p. 38). This tends to blur a distinction which Hare endorses, viz., that between analytic meaning-rules and those universal *prescriptions* to which we are committed in issuing evaluative judgements.

(u) if and only if the inference (s) from (u) and a subsuming premiss(es) of facts (p) is the content of the person's conscious thought.

Such phenomenological subscription should be distinguished from what we may call *tacit* subscription. Often we operate with tacit reasons. We infer our conclusions from premisses of which only some, if any, are at all explicit to ourselves. Our reasoning seems frequently to take on such elliptical forms.

Does it follow from Hare's universalizability-thesis that a person must be tacitly or phenomenologically subscribing to a universal principle? I do not see how it could. We must separate two questions, of which the universalizability-thesis, *qua* logical thesis, is an answer to only one:

Logical question:	What logical conditions must someone meet who has issued in earnest a singular ought-judgement?
Genetical question:	How do people decide or come to hold the singular ought-judgements they in fact hold?

These two questions (here applied to 'ought') are not necessarily related. We may answer either of them without committing ourselves to a standpoint with regard to the other. To say that people who are issuing singular 'ought'-judgements are phenomenologically or tacitly subscribing to universal principles is to make at least in part an empirical claim that can be an answer to the latter but not the former, whereas the universalizability-requirement, *qua* thesis about the meaning of 'ought', is an answer to the former.

Now, the theory of moral reasoning which Hare has developed, as we shall see, does not require that we understand "imply in a loose sense" or "subscribes" as cases of phenomenological or tacit implying or subscribing. All that Hare needs is an answer to the logical question to the following effect: a person who issues a singular evaluative judgement, say, 'I ought to do X' cannot issue such a judgement and refuse to assent to, as outlined above, the universalization of this judgement.

Let me now turn to the explanation which Hare gives for why he thinks we are logically committed to universalize our moral judgements.

3.2 Descriptive Meaning

The second chapter of *FR* - "Descriptive Meaning" - contains the most extensive exposition of what Hare means by his claim that issuers of a

singular moral judgement are logically compelled to universalize their judgements. Calling attention to the relation between descriptive and moral language, he makes it clear that moral judgements are universalizable *"in the same sense"* (*FR*, p. 12) as descriptive judgements are universalizable:[15]

> We must now notice the connexion between the fact that some judgements are descriptive and another feature which it has become the custom to call, when we are speaking of moral judgements, *universalizability*. It is important to emphasize that moral judgements *share* this feature with descriptive judgements, although the differences between them in other respects are, as we shall see, sufficient to make it misleading to say that moral judgements are descriptive. Nevertheless, in so far as moral judgements do have descriptive meaning, in addition to the other kind of meaning which they have, they share this characteristic, which is common to all judgements which carry descriptive meaning (*FR*, p. 10; my italics; see also p. 12).

> it follows from the definition of the expression 'descriptive term' that descriptive judgements are universalizable in just the same way as, according to my view, moral judgements are (*FR*, p. 16).

Now Hare next explains why descriptive judgements are universalizable. As may be recalled from chapter II, one of the ways in which Hare explains what he means by a descriptive judgement is that it is a judgement, the meaning (including its reference) of which "determines uniquely its truth-conditions, and vice versa" (*ET*, p.18, 1976; cf. Ibberson 1986):[16]

> the universalizability of singular descriptive judgements is a consequence of the fact that the meaning-rules for the descriptive terms which they have to contain are universal rules, and universal rules of a certain type (*FR*, p. 13).

Thus, we have the two claims:

(M1) Moral judgements are universalizable in the same sense as descriptive judgements are.

(M2) Descriptive judgements are universalizable by virtue of the special type of meaning-rule that they must contain.

---

15. This explanation of why moral judgements are universalizable is not limited to *FR*. In *MT* he makes explicit reference to it: "The reason is connected with the universalizability of moral judgements; this, as I have shown at length elsewhere, results in their having, like purely factual or descriptive statements, a descriptive meaning (*FR*, 2.1 ff.; *LM* 7.1 ff.)" (*MT*, p. 8). See also *MT*, p. 219.

16. What Hare means by a descriptive predicate is interpreted by Ibberson (1986, p. 24) as a predicate "whose sense lays down the conditions a thing must meet to satisfy it".

Let us begin by examining the latter claim.

3.3 Meaning-Rules

Let us consider some passages that express Hare's view on meaning-rules:

> I may perhaps be allowed to say that meaning of any kind (so far as it is words that are said to have meaning) is or involves the use of an expression in accordance with certain rules; the *kind* of meaning is determined by the *kind* of rules (*FR*, p. 7; cf. *ET*, p. 126).

Stressing that he is not making language more inflexible than it is, he maintains some lines further on that:

> It suffices for our present purpose to say that by 'rules' I do not mean very simple general rules which can be formulated in words (3.4), but, rather, that consistency of practice in the use of an expression which is the condition of its intelligibility (*FR*, p. 7).

Several pages later on he presents what seems to be one distinguishing characteristic of descriptive meaning-rules:

> One of the features of descriptive meaning, as opposed to other sorts of meaning, is that it relies upon the concept of *similarity* (*FR*, p. 13).

A few lines further on he adds that

> It is a *direct consequence* of this that we cannot without inconsistency apply a descriptive term to one thing, and refuse to apply it to another similar thing (either exactly similar or similar in the relevant respects) (*FR*, p. 13; my italics).

Consider the following judgement: (s) 'This piece of wood is rotten'. What Hare apparently is saying is that when a person issues (s) sincerely he must endorse a rule which says that anything like this in these respects is rotten.

The reason why the speaker must assent to such a rule is, it seems, a consistency requirement. It might be objected that this cannot be something special for descriptive words, since we are required to use all our expressions consistently. D. H. Monro (1967) and Brian E. Schrag (1975), for instance, have argued respectively that all words are what they call descriptively universalizable. Schrag maintains, for instance, that:

> part of the function of any rule of applicability is to enable consistency of predication (Schrag 1975, p. 64).

Following Monro (1967), Schrag maintains that all predicates, except proper names and demonstratives, are *descriptively universalizable*. He explains this as follows:

> If a term (correctly) applies to any one subject, then it, by implication, applies to any other thing like the subject in the relevant respects (Schrag 1975, p. 59).

Schrag discusses three words which Hare has claimed are not universalizable, namely "it", "please", and "wanting". Consider for instance what he has to say on the latter.[17] Schrag quotes to begin with the following passage from *FR*:

> In this respect wanting is like assenting to a singular imperative. . . .If I want to do A in these circumstances I am not committed to wanting anyone else placed in exactly or relevantly similar circumstances to do likewise (*FR*, p. 71).

Schrag then goes on to claim:

> wanting is certainly universalizable in our sense of descriptive universalizability. If I am justified in applying the term 'wanting' to some sort of attitude of mine then I am obliged to apply the term to any other attitude of mine relevantly similar to this one, and for that matter, to anyone else's attitude which is relevantly similar to mine (Schrag 1975, p. 67–68).[18]

However, notice that Hare is speaking of the phenomenon of wanting, whereas Schrag is speaking of "applying the term 'wanting' to some sort of attitude of mine". We must distinguish between consistency in wanting from consistency in describing some attitude as wanting. Part of what Hare is doing on the page quoted from *FR*, is to argue that 'I want' is not always an expression used to describe an attitude of the speaker. According to Hare, to say 'I want x' is at least sometimes not to describe an attitude, but should rather be understood as a prescription. Whether this is true or not is an issue we do not need to go into here. The doubts I expressed earlier in

17. For Hare's account of 'it' see *FR*, p. 9, and *MT* p. 70.

18. A similar point is made by Hill (1974, p. 294).

chapter II concerning the prescriptive analysis of first-person judgements extends also, I suspect, to 'I want'.[19]

Consider the following passage from Hare (1986), which suggests how Hare would meet Schrag's contention above about the function of any rule of applicability.

> Descriptivism (or the descriptive fallacy as Austin called it. . .) is the belief that all words get their meaning in the same way as descriptive words and statements do, by having application-conditions or truth-conditions. The importance of imperatives for ethics is not that moral judgements are *imperatives*, but that imperatives are a standing counter-example to refute this mistaken view (*ET*, p. 120).

Some lines further on he makes it clear that imperatives lack application-rules:

> [Descriptivism] obviously does not work for imperatives, which cannot be true (there are no *application*-rules for the imperative verb-form). To know the meaning of 'Shut the door' is not to know under what conditions we can truly say it, nor even under what conditions we can say it at all. The Latin word 'esto' ('let it be so'), which we may take as a pure imperative verb-form, does not get its meaning from rules which tell us when and what we can, when using it, command or request (*ET*, p. 120–121).

The last lines of the above quote are important, since they suggest a way of qualifying the statement from *FR*.[20] That is to say, by extrapolating we may take Hare as implying the following: What is characteristic of descriptive meaning-rules is not only that they rely upon the concept of similarity. To this we must add that the meaning-rule of a descriptive term,

---

19. Dunn (1987, pp. 66 ff.) contains an interesting discussion of how to understand sentences such as 'I want/will do *x*', which is critical to Hare's approach. Ward (1973) claims that 'I like strawberries' is not universalizable, since "if a person says that he likes a thing, is not intentionally trying to mislead his listeners, then it is true that he likes that thing. A person does not need to know why he likes a thing [and because of this] he does not have to say why he likes some things and dislikes other things in order to for it to be true that he likes some things and dislikes other things" (*op. cit.*, p. 290). I agree with Ward that 'I like' is not necessarily universalizable. However, his explanation seems to me incorrect. Since I take it that one can like a certain thing without liking another thing that is exactly similar in universal respects, there is no logical reasons for why 'I like' is universalizable in those cases.

20. Is it really the case that the imperative form does not have *any* rules of application? It seems somewhat too harsh and categorical to claim that this form lacks such rules. Is it not more reasonable to say that the imperative form does not have the same kind of application rule as descriptive terms? However, a discussion of this issue would easily turn into a mere verbal quibble. That the application rule of descriptive terms does not apply to the imperative form seems a reasonable point of view.

say, 'rotten', is such that it tells us when and/or what we can, when using it, describe as rotten.[21]

Let me now turn to the second claim, **M1**, that moral judgements are universalizable in the same sense as descriptive judgements.

3.4 Descriptive and Evaluative Judgements

M1 has been the object of much criticism. Holmes (1966, p. 114), for instance, maintained in an early criticism of Hare's view that "it is misleading to say that moral judgements are universalizable in precisely the same sense as descriptive statements",[22] and several writers (e.g. Monro 1967; Locke 1968 and Schrag 1975), for instance, have found reason to say that Hare uses 'universalizable' in at least two senses, viz., what they call descriptive and consequential universalizability.[23] The gist of their argument can be put as follows: Hare claims that descriptive terms are universalizable in the same sense as evaluative terms (cf. *FR*, p. 11, 16).[24] However, there are passages, it is held, that clearly indicate that Hare, in discussing certain evaluative terms, uses a sense of 'universalizable' that is not applicable to a

---

21. Intuitively it seems reasonable to add the disjunction 'and/or', since it is not obvious that the application rule of any descriptive term must include conditions for when to use it.

22. Holmes' article contains several interesting points. However, his main objection that the universalizability-thesis and the prescriptivity-thesis are in conflict with each other is based on an interpretation that Hare's recent works makes clear is a misinterpretation.

23. Actually Locke does not use the expressions "descriptive" and "consequential" universalizability. However, his distinction between two senses of universalizability coincides with Monro's and Schrag's use of these expressions. Both of the latter writers speak also of what Monro (1967, pp. 164–68) calls "socially universalizable", which they both incorrectly ascribe to Hare. Schrag (1975, pp. 86–87) for instance, explains it in the following way: "If any speaker of a language is justified in applying a certain term in a certain situation, then every user of the language is justified in applying the term in relevantly similar situations. This is because if a person (justifiable) applies a term in a certain situation, it must be in virtue of some features of the situation, and not merely because some particular user of the word is in the situation" (Schrag 1975, pp. 86, 87). It is not all together clear to me what is here meant by "justified" or "applied". As it stands the thesis seems to me much too strong. Be that as it may, "socially universalizable" is not what Hare has in mind when he speaks of the universalizability of moral judgements.

24. Schrag does not formulate it in exactly this way, but I believe what follows accords well with the gist of his objection. See Schrag 1975, p. 93 s.f)

descriptive term such as, for instance, 'red'.[25] Hare must therefore, it is argued, give up the claim that evaluative terms are universalizable in the same sense as descriptive terms are universalizable (see e.g. Schrag 1975, pp. 91–92). Locke puts it in the following way:

> Thus the judgement 'This is red' is universalizable in that we can pass from it to the 'principle' that anything which is like it in being red is red. And if the judgement 'This is good' is universalizable in precisely the same way, as Hare says it is, then this means that we can pass from 'This is good' to the 'principle' that anything which is like it in being good is good (Locke 1968, p. 26).

Locke's point is that when Hare discusses universalizability in relation to moral judgements, he has another sense in mind, namely, that if we claim, say, that some *a* is good there must be something which "as a matter of fact" (*op. cit.*, p. 28) makes *a* good.[26]

It is unclear to me just how the **M1** claim from *FR* should be understood. However, in recent works Hare appears to concede the criticism (or at least to qualifiy his claim). He now makes it reasonably clear that descriptive judgements and moral judgements are universalizable - not for the same reasons - but in the same sense that we cannot logically deny in issuing either kind of judgement that there is some universal rule (or principle) involved (e.g. Hare 1984b).[27]

Hare devotes more space in *FR* to differentiate evaluative judgements

25. Schrag quotes for instance the following passage from *FR* (I include only a part of the passage): "Let us suppose. . .that there are two pictures very like each other. To call one good and the other not, commits the speaker to saying that there must be some differences between them which makes them differ in respect of goodness, and if it be granted that there must be some difference between two pictures, one of which is good and the other not, then it follows that if a man calls a picture good he is committed to calling any other picture good which is exactly similar" (*FR*, pp. 140–141). Schrag notes that what makes us call a thing, say, red, can be that the thing has the quality of redness. In the case of 'red' we would not be entitled to say that the person is committed to saying that there must be some difference between two objects that make them differ in respect of redness, other than this that one is red and the other is not.

26. Hare's later work, as will be seen, makes it clear that he is not saying that there must as a matter of fact be some underlying property that makes a thing good. With regard to 'red', see e.g. the following passage from *FR*, p. 15: "If I call a thing red, I am committed to calling anything else like it red. And if I call a thing good, I am committed to calling any X like it good. But whereas the reason in the former case is that I must be using the word 'red' in accordance with some *meaning*-rule, the reason in the latter case is much more complicated". The complication arises, as will be seen, from the fact that 'good' is a primarily prescriptive term.

27. That this is the case will hopefully become clear when I present his latest views on the supervenience-relation in section 3.8.

from descriptive judgements than he does to make it clear what these kind of judgements have in common. We need therefore to look more into what distinguishes the universal rule in the case of evaluative judgements from the descriptive rule.

Now, by contrast with descriptive terms evaluative terms do not, Hare maintains, have their whole meanings fixed for them by descriptive criteria.[28] Hare does not deny that the meaning of descriptive words cannot alter. There is, as he puts it, "'family resemblance' and 'open texture' and all that" (*FR*, p. 26).

What he is saying is that if two persons are using a descriptive term with different meanings, they could clear up their misunderstanding "by means of an agreement on the use of the word" (*FR*, p. 28). But in the case of moral disputes there could still be disagreement although the persons agreed on what non-moral properties were involved.[29]

The reason why moral judgements do not have their meanings fixed, is that they in addition have prescriptive meaning. We choose ourselves how we want to apply our value-words. For that reason we could be said to make our own meaning-rules in applying our value-words to one object, act, or person rather than to another. In other words, when we make up our moral standards about which objects are to be prescribed as good or obligatory, etc., we are, in so far as we are ready to universalize our judgement, free to choose whatever features or properties we want.[30] Linguistic conventions limit our use of descriptive terms. I cannot, for instance, use the English

---

28. The idea that the descriptive meaning of an evaluative term may vary from person to person has made several writers think it is misleading to use 'descriptive meaning' about such terms. See Holmes (1966, p. 117), and Vesey (1965), for instance.

29. There is a difference between the account in *FR* and *MT* on this issue. Thus whereas Hare in *FR* accepts as a fact that what he calls the fanatic may agree with us about what are the non-moral properties of the situation under consideration, and still make a moral judgement that is inconsistent with our judgement, in *MT* such a possibility will not be real for a perfect critical thinker, as we shall see later on.

30. See for instance *MT*, p. 69: "an important feature of moral language, neglected by naturalists, that we can go on using the moral words with their same meanings to express moral opinions at variance with the received ones, as moral reformers do. This would be impossible if the moral words were tied by virtue of their very meanings to fixed properties of actions, etc" (cf. *MT*, p. 70). In *FR* he claims that when we take into consideration that moral terms are also prescriptive, it will be clear that in enunciating the meaning-rule we shall be doing more than "specifying the meaning of the word . . .it is not mere verbal instruction that we are giving, but something more: moral instruction" (*FR*, p. 2). The claim that there are no limits to what could be a moral rule has been criticized by several writers. See for instance Warnock (1974), Foot (1958), Anscombe (1958). Hudson (1985) contains a good summary of the main objections to Hare's view.

word 'carnivore' to denote *any* kind of animal; and if I did, my disagreement with someone who used 'carnivore' in another way could be settled by making clear what (natural) properties animals that are carnivores have to have. Moral terms are not, according to Hare, restricted in this way; their meaning-rules are not fixed by linguistic conventions (see *MT*, pp. 70, 71; cf. Ibberson 1986).[31] When we take the prescriptive element into consideration,

> the descriptive meaning-rule becomes more than a *mere* meaning-rule (*FR*, p. 23. My italics).

The prescriptive element, in Hare's view, turns the meaning-rule into something synthetic, viz., a moral principle. To take Hare's own example: The man who says that a certain man is good because he "for example... feeds his children, does not beat his wife, &c." will not just be explaining the meaning of the word 'good'. The listener will learn something synthetic.

> It will be synthetic because of the added prescriptiveness of the word 'good'; in learning it, he will be learning, not merely to use a word in a certain way, but to commend, or prescribe for imitation, a certain kind of man. A man who whole-heartedly accepts such a rule is likely to *live*, not merely *talk*, differently from one who does not. Our descriptive meaning-rule has thus turned into a synthetic moral principle (*FR*, p. 23).

Thus, whereas the rule involved in issuing a descriptive judgement will be analytic, the rule of the evaluative judgement will by synthetic, expressing the preferences (as outlined in chapter II) of the speaker. Both kinds of rules will be universal in the **U1** and **U2** sense. However, while it is the meaning of the descriptive term that gives the necessary and sufficient conditions for how to apply the term, in the case of evaluative judgements it will be the speaker's desires that determine what is to be called 'good', 'right', 'morally obligatory', etc.

3.5 In Criticism of (U1) and (U2)

Although there seems to be a general consensus today that moral judgements are universalizable in some sense or other, Hare's universalizability-claim has given rise to a considerable discussion. Several

---

31. There are, as may be recalled, value-words, according to Hare, that are used with a more or less fixed descriptive meaning. He calls these secondary evaluative terms. For instance, 'lazy' or 'industrious' are words that can be used in a non-prescriptive way.

writers produced what they took to be counter-examples to Hare's universalizability-thesis. Alasdair MacIntyre (1957), for instance (see Mackie 1985 for a more recent example), gave what he thought to be cases in which a person could not make himself universalize his judgement, since, among other things, this would mean he would be legislating for someone else, which, in its turn, would be a sign, it was argued, that the person was intolerant or morally arrogant. Other writers (see for instance Cohen 1967, Locke 1968, Laymon & Machamer 1970, Hudson 1988, Mayo 1986) have effectively criticized MacIntyre's examples, and I will therefore not repeat the criticism here.

Part of the reason why Hare's universalizability-thesis was criticized by Macintyre and others concerns the very formulation of the universalizability-criterion. The universalization of a judgement under discussion was often rendered a weak formulation that turned the principle under discussion into a *general* rather than in the strict sense *universal* principle, i.e. formulated in accordance with (**U1**) and (**U2**). Hare's distinction between the terms 'general' and 'universal' is, I believe, sound: "'generality' is the opposite to 'specificity', whereas 'universality' is compatible with 'specificity'" (Hare 1955; *MT*, p. 41; *FR*, p. 39; *ET*, pp. 51–52; *MP*, pp. 16–18).

Peter Winch, for example, regards universalizability as the requirement

> which would have it that a man who thinks that a given action is the right for him to perform in certain circumstances is logically committed to thinking that the same action would be right for anyone else in relevantly similar circumstances (Winch 1972, p. 6).

However, it is not sufficient to qualify that the circumstances must be similar or relevantly similar; the expression "anyone else" must be rendered 'anyone else who is exactly or relevantly similar to him (with regard to universal properties)'. Once it is clear that the similarity-relation ranges over people as well as situations and acts, Winch's (1965, republished in Winch 1972) example of a non-universalizable moral judgement loses much of its intuitive appeal.[32]

In what follows I will continue to avoid formulations of universal principles in terms of words such as 'Everyone ought to. . .' or 'One ought to. . .' (cf. *MT*, p. 62 s.f). Several writers, (at times Hare himself) who discuss

32. For a criticism of Winch's example, see e.g. Kolenda 1975), Levin 1979, Norton 1980, and Maclean 1984. Wiggins (1987, pp. 166–184) has a defence of Winch's claim.

Hare's view avoid the more cumbrous formulation in terms of exact similarity (with regard to universal properties). I think this is unfortunate. If 'everyone ought to. . .' is merely used as shorthand for 'For any person exactly similar to *a* (with regard to universal properties), it is the case that this person ought to. . .', it should be avoided, since it is likely to give rise to misunderstandings. If it is not a substitute, it is a misinterpretation.

Not all writers agree with Hare that only a principle formulated in terms of (**U1**) and (**U2**) qualifies as a universal principle. Bernard Mayo (1986), for instance, has recently expressed the view that universality in morals admits of degrees. However, I suspect that what Mayo has in mind is that we can adhere to principles that vary with regard to generality rather than to universality.

Consider the following singular judgement: (s) 'Kim ought to visit his dying mother'. According to requirements (**U1**) and (**U2**), the speaker must assent to some such principle as:

(U) For any person exactly similar (with regard to universal properties) to Kim, this person ought to visit his dying mother or any person who is exactly similar (with regard to universal properties) to Kim's dying mother.

(U) is a highly specific principle; it concerns only those persons that are exactly similar (with regard to universal properties) to Kim (and the mother). We could make it more general by maintaining that the principle involved would state something such as:

(G) Any person who has a dying mother ought to visit his or her dying mother.

Although there may be various reasons for why someone holding (s) may be expected to hold something like (G), it is hard to see how consistency can require the speaker to hold (G). It may be reasonable or morally demanded that one ought to visit one's mother, but consistency cannot alone require the speaker to endorse (G) because he holds (s). There is no inconsistency involved if the person uses 'ought' in accordance with (U), although he refuses to use it about, say, Charlie who has been maltreated by his mother all his life.

Hare's universalizability-thesis demands of us that if we hold a moral judgement about a particular case, we are logically committed to hold it about any other exactly similar case (with regard to universal properties).

Now, not all writers have interpreted Hare in this strict sense. Locke. (1968), Pollock (1972), and Norton (1980), for instance, have in their criticism of Hare's universalizability-thesis interpreted it in a sense other than that of (**U1**) and (**U2**). Pollock (1972, p. 31), for instance, seems to believe that the scope of Hare's similarity-requirement ranges over acts and circumstances but not over persons. The same idea is present in the other works. Thus, Silverstein considers the example of the creditor who wonders whether he ought to put his debtor in prison. According to him, Hare's universalizability-thesis requires of the creditor to ask whether he is ready to hold the same argument for a hypothetical case that is similar in all respects except that it is now the creditor who is a debtor (Silverstein 1974, p. 62).[33] Norton, in his turn, calls attention to what he thinks is an internal disparity in Hare's universalizability-criterion.

> The action which is in itself or in its principle universalized is the product of the conjunction of two factors, namely persons and circumstances. The disparity appears in the qualification of but one of these factors – the circumstances – by the terms 'like,' or 'similar,' or 'relevantly similar' etc. For parity to obtain, the formulation would have to read either 'binding on anyone in any circumstances,' or 'binding on like persons in like circumstances' (Norton 1980, p. 519).

Further on Norton maintains that Hare warrants the non-qualification of persons in universal principles (Norton 1980, p. 520). Finally, Locke for his part maintains that Hare

> continually states his view in terms of 'same situation' or 'same circumstances', as if he took the thesis to be that if *A* ought to do *X* then anyone, no matter who he is, in the same *objective* state of affairs ought also to do X (Locke 1968, p. 33).

As may be recalled, Locke distinguishes between two kinds of universalizability, viz., that which, according to Locke, Hare argues for, and secondly, the universalizability that, in Locke's view, Hare actually holds –

---

33. Silverstein (1974, p. 62) argues against the "requirement that the hypothetical cases one considers be exactly similar to the actual one". Suppose a person *C* wonders whether he ought to do *y* to a person *D*. Silverstein maintains now that the requirement in terms of exact similarity would mean that the hypothetical case which *C* is required to consider, "must be simply another case in which *C* does *y* to *D*". But this would be, he thinks, "a plainly unacceptable interpretation". Silverstein claims therefore that the hypothetical case which *C* must consider "is not to be exactly similar *simpliciter* to the actual case, but exactly similar except that he, *C*, plays the role of *D*". However, Hare does not maintain that the hypothetical case must be exactly similar in all respects, only exactly similar in universal respects. See also section 5.2.

where the former but not the latter would range over persons (see section 3.6).

In contrast to these writers I do think that the idea that the similarity-relation extends not only over circumstances and acts but also over persons, is present in Hare's earlier works. Already in "Universalisability" (Hare 1955), for instance, this is strongly suggested. Consider, for instance, the following extract from a conversation between what Hare calls a "Kantian" and an "Existentialist" - where he thinks "most of us would be as baffled as the 'Kantian'" (*MC*, p. 22).

> E.: You oughtn't to do that.
>
> K.: So you think that one oughtn't to do that kind of thing?
>
> E.: I think nothing of the kind; I say only that *you* oughtn't to do that.
>
> K.: Don't you even imply that *a person like me* in circumstances of this kind oughtn't to do that kind of thing when the other people involved are the sort of people that they are (*MC*, p. 21, 1955; my italics).

Part of the reason why Hare's universalizability-thesis has been criticized, derives from the fact that he has used as examples of universal principles, principles that on a strict reading cannot be said to be the universalization of the particular singular judgement which he discusses. Notably in his works prior to *MT*, there are various examples of universal principles that all have in common that they are generalizations rather than universalizations of the singular judgement which he examines (contrast Marcus G. Singer in Potter & Timmons 1985, p. 51).[34] Such examples are found *passim*. Thus, in "Universalisability" he maintains that "One ought to keep one's promises" (in *MC*, p. 16, 1955) is "the sort of maxim" a person who says "I promised to" would be invoking (cf. *FR*, p. 91; *MT*, p. 218). The most puzzling example I know of comes from *LM* (p. 158), in which the universalizability-thesis, as may be recalled, is only argued for implicitly.

34. Singer makes a point of the fact that Hare as an example of a universal principle uses "One ought never to make false statements to one's wife". This principle, he says, "is not universal, because it applies only to males and not to females" (in Potter & Timmons 1985, p. 51). I think this is an unreasonable way of using 'universal'. Be that as it may, Hare makes it clear already in "Universalisabilty" that with "universal principle" he covers principles containing "bound variables, that are applicable in one sense to anyone, but in another sense only to one person" (*MC*, p. 16), namely, in this case, to the one who has a wife. An objection similar to Singer's is found in Locke (1981, p. 548, n. 32). Relevant here is also Wetterström (1986), whose idea of a "unique universal property" I find illuminating.

Here Hare maintains that the judgement 'They ought not to build any more by-passes round Oxford' "depends" on such a principle as 'When traffic census figures show that all but a very small proportion of the traffic in a town is terminating traffic which could not use a bypass, large sums of money ought not to be expended on one". Whatever the relation is between this principle and the singular judgement, consistency alone together with the **U2**-thesis cannot hardly commit one logically to assent to this principle.

I believe, however, that a reasonable interpretation of the presence of these principles is that Hare has used these examples for the sake of illustration. Instead of using the cumbrous formulation in terms of exact similarity, he has given example of principles that it is reasonable to expect that people would invoke. However, I will not pursue this point any further, since Hare has made it very clear in *MT* and more recent works that the sense in which moral judgements are universalizable, is the one required by (**U1**) and (**U2**).[35]

I have above outlined Hare's account of his universalizability-thesis. Most of his extensive work on this matter has concerned the (**U1**) requirement. In the remaining two sections of this chapter I will examine two objections to Hare's universalizability-thesis, which I think carry a certain strength. Both arguments centre on the (**U2**) requirement, which in a reasonable enough sense could be said to be the "weak spot" in his account. This should come as no surprise. (**U1**) or something to its effect is required to secure consistency in our use of moral words, and although I have along the way expressed some minor doubts and questions concerning Hare's defence of (**U1**), I share his linguistic intuitions with regard to this requirement. The case of (**U2**) is different. Hare here appeals in a much more direct way to the reader's linguistic intuitions. To begin with, I will consider whether there are non-universalizable 'ought'-judgements. I will argue that there are, and that this fact undermines Hare's claim that his universalizability-thesis is a logical thesis. The objection which I consider in section 3.7 questions the scope of the universalizability-thesis. It concerns whether we should extend our judgements to hold for hypothetical cases. This argument, influential as it is, will not, however, seriously threaten Hare's universalizability-thesis.

---

35. In *MT* he thus says: "if we make different moral judgements about situations which we admit to be identical in their universal descriptive properties, we contradict ourselves" (*MT*, p. 21).

3.6 (U2) and Non-Universalizable Reasons

Suppose a person were to say (s) 'I ought to sacrifice myself but no one else who is exactly similar to me, being in exactly similar circumstances (with regard to universal properties) ought to sacrifice himself'. Hare describes such examples sometimes as cases of contradiction (e.g. *MT*, p. 21) and sometimes more carefully as utterances that have an effect which is "similar to that of a contradiction" (*LM*, p. 134) or "encounters the same kind of incomprehension as is encountered by a logical inconsistency (for example a self-contradiction)" (*MT*, p. 115).

The requirement in **(U2)** that moral principles must not contain ineliminable references to individuals, has been questioned by several writers.

Torbjörn Tännsjö, for instance, puts the matter in the following way.

> Only ethical considerations can prove that certain properties such as nominal, numerical and nonnatural ones, are not relevant to ethics (Tännsjö 1974, p. 143.).[36]

However, of these "properties" it is clearly only what Tännsjö calls the numerical property that is excluded by Hare's universalizability-thesis. The fact that a person invokes non-natural properties as a distinguishing feature, for instance, does not in itself qualify the judgement as being non-universalizable. Only if this property cannot be described other than by reference to an individual would it be ruled out (see for instance *FR*, p. 18 s.f., where he discusses the case of a non-naturalist who accepts the universalizability-requirement; see also *ET*, p. 72). Nor are what Tännsjö calls nominal properties excluded by the universalizability-thesis. Two persons differ, according to Tännsjö, in a 'nominal respect' if they do not bear the same name (Tännsjö 1974, p. 139). However, to carry the name 'Tim' is clearly to have a universal property, i.e., it is a property which in a reasonable enough sense is describable without referring to a certain individual or group of individuals (cf. Rabinowicz 1979, p. 92). To justify

36. Steven Luper-Foy (1990), in an article about aggression, employs a vocabulary in which properties such as "*being unique, or exceptional, in some given respect*, [are] properties that not everyone could possibly have; they are necessarily non-universalizable" (p. 213). Luper-Foy seems to be correct that in one and the same world, everyone cannot be unique in some given respect. Still, such a property would be universal in that we could imagine a possible world in which new roles were imagined. Cf. Wetterström (1986, pp. 25 ff.), who accounts for three different ways in which a universal property can be unique.

why a certain person ought to do a certain act by saying that he was named 'Tim', would in Hare's view not be to misuse the term 'ought'. It would be having an extremely odd reason for making a moral judgement.[37]

We get a different case when we turn to what Tännsjö calls numerical properties, which he explains in the following way:

> The idea about 'numerical' properties is that the fact that a *particular* person is involved, and not another person, must not be taken to be of ethical importance. If something is right for me, then it is also right for anyone exactly like me in all respects except that he is 'numerically' different from me, i.e. except that he is not me but another person (Tännsjö, 1974 p. 139).

There is here an unclarity with regard to the expression "numerical property", that we may put aside.[38] Nor does Tännsjö seem altogether comfortable with the expression. The expression apart, the upshot of Tännsjö's contention seems clear enough: Hare cannot exclude such judgements as the above on purely logical grounds. This is possible only on ethical considerations. Thus, can we hold the judgement, say, 'I ought to do x' and not be logically committed to universalize it?

Hare has addressed this topic in a number of works. In "Universalisability" (*MC*, 1955), for instance, he discusses, following Gellner, what he calls "E-type" and "U-type" valuations. The former involve an ineliminable reference to an individual, whereas the latter are devoid of any reference to individuals. Hare then makes the following claims:

(1) All actions for which there are reasons involve maxims.

(2) These maxims may be either of type E or of type U (*MC*, p. 21, 1955).

To these two claims he then adds the following:

(3) All *moral* evaluations are of type U.

---

37. Cf. Hare's discussion of Foot's example "Suppose that a man says that somebody is a good man because he clasps and unclasps his hands, and never turns NNE after turning SSW" in Hare (1963b). See also Hudson 1989, p. 14.

38. The reason I find an entity's property of being the same as itself problematic, is that it is unclear whether we can meaningfully speak of comparing something to itself. The only case when this is meaningful is when we compare two states of an entity at different times, which is not what is involved here. For an interesting article on this issue, see Durrant 1973.

That is to say, he seems to use 'moral' to exclude type E judgements. When Hare in *FR* returns to this matter, he focuses on 'ought'-judgements and not moral evaluations in general. Moreover, he makes an important (and often overlooked) qualification of his earlier account, claiming now that it is *not* in virtue of the meaning of 'moral' but rather of the meaning of 'ought', that he excludes evaluations that contain ineliminable reference to individuals.[39]

Henceforth I will use "type E 'ought'-judgements" to refer to judgements for which a type E reason is given. Consider the following passage from *FR*:

> Plain imperatives do not *have* to have reasons or grounds, though they normally do have; but 'ought'-judgements, strictly speaking, would be being misused if the demand for reasons or grounds were thought of as out of place – though the reasons need not be ulterior ones; some universal moral judgements already incorporate all the reasons they need or can have . . ..
>
> Nevertheless, it may be that there is a debased use of 'ought' in which it is equivalent to a simple imperative (though I must confess that I have come across such a use only in the writings of philosophers). Just in case, however, there is such a use, it is convenient to put the matter in the following way: in by far the majority of judgements containing the word 'ought', it has the sense that requires them to be universalizable; there *may* be some peripheral cases where it does not have this sense; but at any rate in its *moral* uses (with which we are chiefly concerned) it always does. The word 'moral' plays here a far smaller role than I was at one time tempted to assign to it. It is the logic of the word 'ought' in its typical uses that requires universalizability, not that of the word 'moral'; the word 'moral' needs to be brought in only in order to identify one class of the typical uses, and that with which as moral philosophers we are most concerned. This means that the ambiguity of the word 'moral', which is notorious, need not worry us at this point. For in whichever of its current senses the word is being used, it suffices to exclude those peripheral uses of 'ought' (if they exist), in which it is not universalizable (*FR*, pp. 36–37).

The idea that non-universalizable uses of 'ought' could exist but that these would not be damaging to his theory is also present in *MT*:

> It is not necessary to my argument to insist that 'ought', for example, is *never* used non-prescriptively (it sometimes is) or non-universalizably. Many words have various senses (for example 'all' and 'if'). All that it is necessary is for there to be a recognizable sense in which it has the properties which I have claimed. People who then ask moral questions *in this sense* – as, I think, we do much of the time – are bound by the logical rules which these

39 Contrast e.g. Williams (1985, pp. 5–6), Mayo (1986, pp. 134–147), and Blackburn (1986, p. 220), who recently have criticized Hare for making it true in virtue of 'moral' that moral judgements must meet what I call the U2-requirement. For an objection to arguments based on our ordinary notion of ethics, see Wetterström (1986, p. 32).

properties dictate. If they want to switch to *different* questions, that is up to them (*MT*, p. 22).

If we concentrate on first-person 'ought'-judgements, it seems as if Hare has considered three possible kinds of 'ought'-judgements, viz., when (s) 'I ought to do act *a* in situation S' is issued

(type A) without any reasons at all,

(type E) with reasons which include at least one ineliminable reference to an individual

(type U) with reasons which refer only to universal properties.

I do not find it hard to follow Hare in claiming that (type A) is a clear misuse. The person who maintains that he ought to do something, and, on being asked for the reasons, answers that there is no reason at all (nothing about me or the situation whatsoever which makes him think that I should do the act, or, nothing which makes him want me to do the act) would be using 'ought' idiosyncratically.[40]

What then of the type E judgement? Hare never explicitly maintains that there are 'ought'-judgements of a type E. Thus, when Hare in the above passage from *FR* speaks of non-universalizable uses of 'ought', he has in mind, on the most reasonable reading, not type E judgements but rather what he referred to in "Universalisability" as type A judgements. However, when in *FR* later on he returns to the issue of non-universalizable 'ought'-judgements, he does seem to consider what I have called type E 'ought'-judgements:

> there is the concept *ought*$_3$, which is prescriptive but not universalizable. Since I do not think that the word 'ought' is actually ever used in the sense of 'ought$_3$', I should prefer to use some other word to stand for this concept (e.g., some word defined in terms of 'want' or of the plain singular imperative); but I use the word 'ought$_3$' in order not even to seem to beg any questions (*FR*, p. 165).

Some lines further on Hare maintains that although the questions that are asked in terms of 'ought$_3$' are ones which it is perfectly proper to

40. The idea in *LM* that value-judgements differ from imperatives, in that the latter need not logically be backed up by reasons, is questioned by Monro (1967, p. 172). Relevant here is also Gellner (1956).

discuss, they are quite different from the ones we ask using 'ought' in its prescriptive/universalizable sense.

> 'Ought$_3$' has to do with questions of self-interest which is not universalized – self-interest, and the interest of groups, such as *my* family, and *my* country, which are defined by reference to an individual (*FR*, p. 165).

What Hare here means by 'ought$_3$' appears to be something like what in 'Universalisability' he called type E valuations, rather than type A valuations. The person who issues an 'ought' judgement because of self-interest can hardly be said to be making a judgement without reason, nor is his reason necessarily universalizable.

Hare thinks that people never actually use such type E 'ought'-judgements. My language intuitions are less firm than Hare's on this issue. I do think type E 'ought'-judgements are made. Whether they are peripheral or frequent uses, is a matter that cannot be settled without gathering a large amount of data, which I am not in a position to do.

The interesting point here is that Hare, in contrast to some writers (e.g. Caton 1963, p. 51)[41], clearly appears to recognize the logical possibility of type E 'ought'-judgements. This concession throws light on the status of the universalizability-thesis. If we do accept that there is a sense in which a person may have a type E reason for his 'ought'-judgement, it appears hard to uphold the claim that the universalizabilty-thesis is a purely logical thesis. The point of the criticism, it seems to me, is that, since (moral) 'ought'-judgements sometimes do (or at least can) involve ineliminable reference to individuals, they are not universalizable in Hare's sense. From the fact alone that it is unusual or unlikely that people would express such type E 'ought'-judgements, we cannot conclude that 'ought' carries at least two kinds of meaning, viz., what corresponds to the type E and type U uses respectively.

---

41. Caton claims that all reasons are universalizable. Later on he qualifies this, saying now that this proposition is "analytic of the notion of at least the sort of reasons involved here" (Caton 1963, p. 51). The kind of reasons involved are moral reasons. Cf. also Brunton (1966) who objects to Hare's exclusion of self-interested uses of 'ought'. Brunton correctly sees that it is the fact that they are non-universalizable that is Hare's reason for preferring 'want' rather than 'ought'. However, Brunton argues that "the egoists[. . .] *do* accept universalizability, with the proviso that their attitude towards prescription varies according to the person or group they are considering" (*op. cit.*, p. 123). Brunton's objection must be based on a sense of 'universalizability' that differs from Hare's, since Hare's universalizability-thesis does not admit of any proviso. If the egoist accepts universalizability, he is committed to prescribe similarly for any situation that is exactly similar with regard to universal properties.

Hare is, of course, free to focus his attention on type U judgements. Moreover, he may have a point that it would be desirable to separate type E (and type A) judgements from the former kind. However, it should be clear that the point of such a manoeuvre would be to clarify our language, and not to exclude certain judgements as being non-moral. Hare is right in thinking that considerations that bear on the type E question, say, 'What ought I to do, if I take my self-interest into consideration?' differ from considerations bearing on type U questions, such as 'What ought I to do, if I take the preferences involved in the situation into consideration?'. But from the fact that we review different types of considerations in answering these questions, we cannot conclude that the latter but not the former question is a moral question, let alone that it is logically suspicious (cf. Williams 1985).[42] The considerations which we admit have a bearing on our 'ought'-question may vary from person to person, from occasion to occasion (some people would, for instance, take only the preferences of living beings into account, others would add that the *potential* desires of a foetus and/or certain "death-bed wishes" should be included).

Suppose that the above objection is valid, to the effect that it remains to be shown by Hare that the person who endorses an 'ought'-judgement, the reason for which contains an ineliminable reference to himself, is making a logical mistake. How much a dent will it make? On the issue of whether the universalizability-thesis is a thesis about the meaning of certain words or not, the objection should have an impact. But what about Hare's overall aim?

To answer this, consider the following passages from *MT* in which Hare maintains that it is not necessary for him to show that 'ought' is never used "non-universalizably":

> if we were to alter the meanings of our words, we should be altering the questions we were asking, and perhaps answering, in terms of them. We come into moral philosophy asking certain moral questions, and the questions are posed in terms of certain concepts. If we go on trying to answer *those* questions, we are stuck with *those* concepts (*MT*, p. 18).

> Many words have various senses (for example 'all' and 'if'). All that is necessary is for there to be a recognizable sense in which ['ought', for

41. A somewhat similar argument is endorsed by Williams, who thinks we cannot conclude that there are moral and non-moral meanings of 'ought' (or 'should') because we may "review a particular type of consideration among those that bear on the ['ought'-] question (1985, p. 6).

> example] has the properties which I have claimed [prescriptivity and universalizability]. People who then ask moral questions *in this sense* – as, I think, we do much of the time – are bound by the logical rules which these properties dictate. If they want to switch to *different* questions, that is up to them (*MT*, p. 22).

Now although Hare does not reveal what these questions are, I venture the opinion that what he says here seems *prima facie* plausible enough. A change in concepts will enable us to ask new questions.[43] Nevertheless, the question remains why this alleged fact is supposed to neutralize the seeming threat to his moral theory that derives from the claim that it is a contingent matter which meaning we want to give to our words. The reason seems clearly to be that those questions which Hare speaks of are somehow conceived of as being important or of a special interest to us. But does this not suggest that there is in the background one or more evaluations that these questions are of special weight to us? I believe it does. However, even if this were so it would not necessarily affect the alleged formal status of the universalizability-thesis. We may recognize that a word plays a certain conceptual role. Whether we ought to use this word or what dispositions are required for us to be interested in having such a word is another question. James Griffin makes a similar point. Thus, he says:

> We might, in the course of building a moral theory, revise certain features of those key terms, so that we should no longer be asking exactly the same questions as at the start, nor entirely new ones either. Could there be, initially, any ground for saying that such partial revision will not, or should not, take place? (Griffin 1986, p. 354, note 30).

Griffin agrees with what he takes to be Hare's response to such an objection, viz., that once we complete our theory there will be "no need to revise the universal prescriptivity of 'ought' and 'must'. But this can be claimed only at the close of theory-building, so it cannot be the source of the theory" (*loc. cit*).

Hare realizes, of course, that his theory about moral language cannot avoid the question 'why express universal, overriding prescriptions?' – or to rephrase it in Hare's sense 'Why make moral judgements?'. Why not express singular prescriptions? He has addressed this issue in various

43. In section 3.9.1 I will voice an objection that may seem to question what Hare says here. That is, whereas I take it to be close to a truism to say that if we change the meanings of our words, we would be altering the questions we were asking in terms of them, the contrary seems much more controversial. In other words, asking questions about different aspects of a problem, does not imply that we have changed the meaning of the words in the questions.

places. The most extensive discussion is found in *MT*, where Hare does produce an argument as to why we should use moral language in the sense outlined by himself. Thus, in chapter 11 of *MT* he sets out to give an answer to the question 'Why should I be moral?'. The reason he gives is that it is in our prudential interest. That something is in a person's prudential interest is, if I have understood Hare correctly, to say that it is something that will satisfy the preference which results when the person has weighed his present preferences against his future ones. Hare's view on prudence has been questioned, and I will return to it in chapter V, where a more detailed account will be required.[44]

Now, although Hare thinks there are prudential reasons for why we should be moral, he carefully points out that it is not the case that what we ought to do will always be in our prudential interest (cf. *MT*, p. 91). To claim the contrary would be an extreme view. However, we can, he says

> abandon this extreme view and still achieve something more modest, which nevertheless is adequate for the defence of morality (*MT*, p. 191).[45]

Hare unfolds his argument, by considering what advice we should give to a child that we are bringing up - advice which would have "only the child's interest at heart". Should we, for instance, advise him to follow, in all situations of prudential decision, a policy of doing an egoistic or self-interested cost benefit analysis? This would be, Hare maintains, misguided advice, since on many occasions there would not be time to make such analysis. More practicable advice, according to Hare, would be to inculcate

---

44. See for instance *FR*, where he asks the following: "But if a man wants to escape from my concepts, where is he going to flee to? To singular prescriptions, expressing selfish desires? Or to universal but non-prescriptive judgements? He is at liberty to take either of these courses; but if he does so, he will not disturb us. For then, though we shall still be in dispute with him about *what to do*, or about *what the facts are*, we shall no longer be in dispute with him about *what we ought to do*. We are in a position to say to him, 'If you do not consent to talk in our terms, the remaining points of dispute between us will be such as can be expressed without using any terms that anybody could call moral or even evaluative. We are ready to have disputes with you of all kinds, but let us keep the kinds distinct"(*FR*, p. 201; see also p. 192). It is, however, unclear in *FR* what the non-moral kind of dispute would consist of.

45. Cf. what he says in *HC*, p. 214: "There are prudential reasons for thinking morally; but prudential and moral thinking are distinct". Relevant is here also his comment to Nagel: "I argue in *MT*, ch. 11 that there *are* prudential reasons for becoming a morally motivated person; but to write into the definition of morality a requirement that everyone has to be morally motivated, as Nagel seems to want to do, is to beg the question, in just the way he has accused me of doing" (*HC*, pp. 254–255. See Nagel 1988 in *HC*).

in the child some "prima facie prudential principles" (*MT*, p. 192). To what extent, Hare next asks, will these *prima facie* prudential principles coincide with the moral *prima facie* principles we would inculcate in him? The latter principles would be those that would secure "not just his interest, but the interest of all those affected by his actions" (*MT*, p. 193).

In the following passage Hare, as I see it, summarizes his view on this matter:

> My guess is that the safest and best way of bringing up our child is to implant in him, if one can, a good set of moral principles plus the feelings that go with them, the feelings being strong enough to secure observance of the principles in all ordinary cases, but not, of course, neurotically strong, or stronger than is needed for their purpose (*MT*, p. 198).

The point of implanting not only principles but feelings as well, is that "*some* aversive feeling is required in order to secure consistent obedience to the principles" (*MT*, p. 197). Such feelings as remorse, shame, and repugnance - what he calls *moral* feelings - will, he thinks, be more effective than other feelings. Supposing this is true, Hare argues, the child will have an "extra prudential reason for obeying his principles". If the child by acting out of self-interest breaks a *prima facie* moral principle, he will for prudential reasons have to include the disutility that derives from feeling shame, remorse, etc. for not obeying the principle.

Now, I am in no fundamental disagreement with Hare over this issue. Some of the empirical assumptions on which the argument is based can, as Hare himself more than once points out, be disputed. Nevertheless, the underlying idea that there are prudential reasons why we should be moral, seems to me sound. Living in a world which I have reasons to believe is shared by other beings similar to myself, I will have prudential reasons for not merely taking my own self-interest into account.[46]

The idea that 'ought'-judgements, regardless of whether we group them under headings such as "moral" or "self-directive", can be backed up by type E reasons is sometimes contested on grounds that "moral reasons" are

46. Hare's reason for why we should be morally motivated may seem too weak. Many recent works have questioned, for instance, our common-sense intuitions regarding personal identity, and it has been strongly argued by some writers (e.g. Parfit 1984) that it is not rational to think that personal identity is what matters. Conclusive reasons will perhaps emerge from this discussion, for why the belief that a certain individual is present should not be considered to carry any special weight. However, this is not the place to enter into this complicated matter, on which there is little agreement. Relevant here is also Williams (1985), who has taken up a very pessimistic attitude towards our possibility of finding reasons for being moral.

of a certain special kind, different from the ones involved in reasoning about what will satisfy one's own self-regarding desires and interests. However, this line of reasoning is clearly alien to Hare's prescriptivist position, and I will therefore confine myself to making a brief comment (cf. Wallace 1988).[47]

At the bottom of the view that not just any reason can be a moral reason, lies often the idea that we must distinguish between what has been called "exciting" or "moving reasons" from "justifying reasons" for acting.[48] The former kind of reason would explain why a person holds a certain ought-judgement. Such an explanation would ultimately mention what beliefs and desires of the agent's internal state caused him or her to hold the 'ought'-judgement.

But as various writers over the years have maintained (e.g. Grice 1967), often when we say of someone that he has a reason for doing something, we are not referring to something about his internal state. When we say, for instance, that Jeppe has a reason to quit drinking alcohol, we tend to persist in our opinion regardless of whether Jeppe says he wants to drink or not. The reason Jeppe has for changing his drinking habits, it is said, is such that it justifies why Jeppe ought to stop drinking alcohol irrespective of what he beliefs and desires.

I take it to be uncontroversial that, as a matter of fact, we *speak* of reasons in both of these ways. The real issue, as I see it, is not whether there are different kinds of statements about reasons, but whether we actually are dealing with two kinds of reasons for acting. Williams (1980), for instance, has argued effectively that once we understand what is involved in the existence of a justifying reason - what Williams calls "an external reason" in contrast to an "internal reason" - we will have to conclude that there is none. The fact that we continue to speak as if there were, is to be explained by the fact that we often hold false beliefs or mistake what our desires ultimately are. When I say that Jeppe has a reason to stop abusing alcohol, I assume that he wants a healthy long life but is mistaken about the

47. Already in *LM*, it becomes evident that even if a person were to justify his principle by showing that it was founded "upon a consideration of everything upon which it could be possibly founded" (*LM*, p. 62) it would not be immune for further requests for justification. For a brief but recent discussion of Hare's view in *LM*, see Wallace 1988, pp. 124–128.

48. Monro (1967) traces the distinction back to Francis Hutcheson. That moral reasons should be justifying reasons has been suggested by Frank Snare (1972) who at the same time makes it clear that he considers this to be a normative claim.

damaging effects of heavy drinking. Suppose, for instance, that Jeppe convinced me that his desire for drinking was not based on any misinformation about the effects of alcohol abuse. Given my sincere belief that he does not desire a long and healthy life, I cannot escape the conclusion that Jeppe does not have any (or at least no overall) reason for becoming a sober citizen.

Let me now consider the second objection that I mentioned earlier. It concerns an issue closely related to the one discussed in the present section, namely, whether Hare is correct in maintaining that we are logically committed (cf. *ET*, p. 198, 1978b) to take not only actual but also hypothetical cases into consideration.[49] Notably the works of the late J. L. Mackie (1977, 1985) have had a great influence on this matter.

3.7 The Scope of Universal Principles

Mackie questions the claim that we are making a logical mistake when we refuse to apply the universalizability thesis to hypothetical cases. He introduces (Mackie 1985) a distinction between a practical universalizability thesis "which would instruct one to act only in ways which one can prescribe to oneself in universalizable judgements" from the logical universalizability thesis which states that "moral terms, concepts, and so on are such that every genuinely moral judgement is universalizable" (Mackie 1985, p. 170). Mackie then sets out to show that by universalizing our singular moral judgement we may actually proceed in three different ways, which correspond, he thinks, to three different stages of universalizability. He maintains that the person who stops at stage 1 or stage 2 does not commit any logical mistake. Moreover, he claims that it is doubtful if any of these stages, let alone all three, can be said to be logically required. The distinction between the second and third level raises questions about Hare's application of the universalizability thesis, which we are not yet in a position to examine, and I will therefore confine my attention to the difference between stage 1 and stage 2.

If I have understood Mackie correctly, the difference between stage 1 and stage 2 universalizability is entirely a question of whether one is ready to

---

49. In the last chapter I will comment on a different objection that has been raised regarding the applicability of universal principles to hypothetical cases. This criticism centres on the idea that it somehow would be logically impossible to imagine oneself having a different set of properties than the properties one *de facto* has (see e.g. Taylor 1965; Locke 1968).

apply one's universal judgement not only to the actual world but to various possible worlds. Thus, it should be noticed that the difference does not amount to different ways of formulating different universal judgements. The difference between stage 1 and stage 2 "turns upon the contrast between the actual world and various merely possible worlds" (Mackie 1985, p. 177). Mackie considers two different kinds of universalizability on stage 2. To universalize on what he calls stage 2a is to be prepared to commit oneself to "possible worlds . . .which are generically identical with 'the actual world' (*loc. cit.*), whereas on 2b, we will be prepared to universalize to all possible worlds whatever. Since neither Hare nor, to my knowledge, anyone else has argued that we are *logically* committed to universalize our judgements in the fashion of stage 2b, we may here disregard it.

On both stages 1 and 2

> a judgement is universal if it contains only general terms, but no proper names or indexical expressions. To universalize a judgement would be to replace any such singular terms that it contains with terms containing only predicates – including relational or many-place predicates – logical constants, and variables, and to bind the appropriate variables with universal quantifiers (Mackie 1985, p. 171–172).[50]

With one exception, the two requirements (**U1**) and (**U2**) appear to express much the same as Mackie's account. The main difference derives from the fact that (**U1**) and (**U2**) are formulated in terms of exact similarity. The examples which Mackie uses of universal judgements are problematic in two ways. To begin with, he is not speaking of 'ought' judgements or other judgements containing moral terms. He prefers for some reason to discuss straightforward prescriptions such as "Let Smith pay Jones £5". It is controversial whether such expressions of preferences and desires are universalizable. Hare maintains that it is not a logical requirement that they should be universalized (see *FR*, pp. 36-37). Secondly, when Mackie gives examples of universal prescriptions, these are not formulated in terms of exact similarity (with regard to universal properties). Thus, Mackie says:

> For example, the judgement 'Let Smith pay Jones £5' would be universalized as follows. We would find some set of general features, which

50. Mackie speaks here about the first stage. However, what he expresses here is the logical universalizability thesis which states what conditions a judgement has to meet to be universal. It becomes clear later on that the requirement expressed in the passage holds for all three stages.

> we can sum up as '*S*', possessed by Smith, and some set of general features, which we can sum up as '*J*', possessed by Jones, and also some set of relations which hold between Smith and Jones, which we can sum up as '*R*', and then say 'For all x and for all y, if x has *S* and y has *J* and x has *R* to y, then let x pay £5 to y' (*op. cit.*, p. 172).

However, as I see it, Mackie's point does not hinge on whether we speak of 'ought' judgements or straightforward prescriptions. Moreover, Mackie's point may well be expressed even if, as before, we formulate the universalizability requirement in terms of exact similarity.[51]

The point which Mackie makes is the following:

> The logical description of stage 1 tells us only what form a universalized prescription must have; it says nothing about the range of actual or possible situations to which it is to be applied. One could perfectly well adopt a prescription which contained no singular terms, no essential reference to individuals, and yet apply it only to actual states of affairs. What one is prepared to commit oneself to with regard to merely possible states of affairs is another matter. So stage 1 does not *logically* carry stage 2a with it (Mackie, 1985 p. 177).

It is not altogether clear to me just what Mackie means by "the logical description of stage **1**". However, the main idea, that the logical universalizability-thesis does not contain any requirement to the effect that one must apply the universal judgement to hypothetical situations, accords well with my own interpretation of Hare; neither **(U1)** or **(U2)** mentions anything about the scope of the universal judgement.

A puzzling feature of Mackie's treatment of this matter is brought out in the following passage:

> But it might be replied that even if stage 2a is thus logically distinguishable, one would be violating the spirit of stage 1 universalization if one refused to go on to 2a. One's possible reason for such a refusal would be a special attachment to some individuals – notably oneself – in contrast with others. That is true, and if the spirit of universalization is identified as complete non-regard for individuals, then it will indeed commit us at least to 2a. But the fact remains that stage 1 universalization, which fails to embody such complete non-regard is coherently adoptable (Mackie 1985, p. 177).

Here Mackie claims that having a special attachment to some individual would be a reason for refusing to go from stage 1 to stage 2 universalization. However, given that one is attached to some individual,

---

51. In one place Mackie says "Hare has, indeed, suggested an ingenious argument to show that stage 2 universalization follows automatically from stage 1" (Mackie 1985, p. 176). However, no reference to Hare's work is mentioned, and it is therefore not clear exactly which argument he is referring to.

how is it possible to sincerely hold a universal prescription even on stage 1? Being attached to an individual would enable us to refuse to go to stage 2a, not because we can be content with staying on stage 1, but because we would express a non-universalizable judgement.

Is Mackie's contention then valid? Are we not logically committed to apply our universal prescriptions to hypothetical situations? This question is important. In light of the claim that it is likely that there will not be two exactly similar (with regard to universal properties) situations in one world, the idea that we must take hypothetical cases into consideration is of crucial importance to Hare. Whether the requriement to universalize our moral judgements will have the effect which Hare thinks it will, depends to a large part on the idea that we must pay attention to hypothetical situations which differ from the actual one only in that the roles of the involved parties are reversed.

Hare has addressed the issue of hypothetical situations in various places.[52] In *MT*, for instance, he maintains:

> There is no way of framing a properly universal principle which prescribes for actual cases but does not similarly prescribe for non-actual cases which resemble the actual cases in all their universal properties and differ from them only in the roles played in them by particular individuals (*MT*, p. 114; cf. Hare 1978b).

The reason which emerges here seems to be that a refusal to consider hypothetical situations would be tantamount to endorsing a principle that contains a reference to an ineliminable individual, viz., the actual world. However, this reply does not really go to the crux of the matter.

Suppose I formulate a universal prescription, say, 'For any person x, let x paint all swans black'. I could without any logical offence prescribe such a thing for only this actual world without committting myself with regard to what should happen in any logical possible world. This would be the case when I issued a universal type A prescription (see section 3.6). So far Mackie has a point.

However, once we take into consideration that moral terms are supervenient, in at least the minimal sense that moral judgements have to have reasons or grounds, light is shed upon Hare's adherence to the claim

52. However, in his reply to Mackie, (Hare 1984a), there is no explict attempt to meet Mackie's point, other than that he maintains that his insistence that universalizability, as he understands it, is founded on linguistic intuitions.

that we are logically committed to apply our universal principle to hypothetical situations. The supervenient nature of 'ought' brings with it that it will always be logically legitimate to ask something to the effect of 'Why is it that so-and-so ought to be done?'. In other words, 'ought'-judgements must logically be supportable by reasons. Now, suppose my reason for saying that I ought to do act x does not contain any reference to an ineliminable individual. In such a case, I fail to see how it could be denied that we are logically committed to say of someone else who in a hypothetical situation is exactly similar to me (with regard to universal properties) that he or she ought to do act x. The only reason I could have for refusing to apply the reason to the hypothetical situation would be by invoking a reference to an individual.

My conclusion is therefore that Mackie, in maintaining that we are not logically committed to universalize for hypothecal situations, overlooked the reason aspect of moral judgements.

3.8 Supervenience

Alongside the above explanation in terms of descriptive meaning-rules, Hare endorses an account of universalizability which centres on the concept of supervenience. Lately the supervenience phenomenon has been discussed to a great extent not only in moral philosophy but in other philosophical areas such as philosophy of mind (and in fields concerning the relations between primary and secondary qualities, in aesthetics with regard to relations between secondary and tertiary qualities, and statements about 'natural kinds' and statements about underlying physical properties). This increasing interest in supervenience is partly due, it seems, to the availability of new distinctions between different kinds of necessity, partly because the notion of possible worlds has gained footing (see e.g. Blackburn 1985, Hare 1984b).[53]

However, I will not here make any attempt to contribute to this issue. My

53. Hare's account of supervenience has the advantage over, for instance, Blackburn's explanation, in that it is simpler. Blackburn's account rests in large measure on a proper understanding of 'underlying'. Thus, consider what Blackburn calls his (S) thesis: "as a matter of necessity, if something x is F and G* underlies this, then anything else in the physical or natural or whatever state G* is F as well" (Blackburn 1985). If 'underlying' is interpreted as referring to a causal relation, (S) runs the risk of being analytic – given that causal relations have a nomological character expressable as 'For all x, if Gx then Fx'. That (S) runs this risk of being trivial was pointed out to me by Ingmar Persson at a seminar in Lund 1990.

aim in the remaining sections of this chapter is partly to present Hare's view on the supervenience relation, partly to examine whether Hare's supervenience claim adds (or withdraws) anything to (or from) the picture presented in *FR*.

Now, the notion of supervenience plays a part in *LM*, whereas there is no explicit trace of it in either *FR*[54] or *MT*. However, he has recently, as mentioned, returned to this topic in Hare 1984b, and I will therefore mainly focus attention on this recent paper rather than on *LM* (for an examination of the account in *LM*, see Hudson 1978).[55]

In moral philosophy the issue of supervenience concerns the relationship which is said to hold between moral properties and natural or non-moral properties (e.g. David O. Brink 1989). Alternatively, it is put forward as a claim about a certain feature of moral terms or moral predicates (e.g. Narveson 1985). When it is said of some *a* that it is, say, good, *a* is good *because* or *in virtue* of some subjacent or underlying property of *a*. Generally it is held that these subjacent properties are natural properties of *a*. Another way of putting it would be to say that when it is predicated of some *a* that it has some moral property, then it is necessarily the case that *a* exemplifies some other (descriptive) predicate. This seems to imply that there is a law expressing necessity of some kind, to the effect that for any other x that is similar to *a* with regard to natural properties, x is good too. Much of the recent controversy about supervenience turns around the status of this law or generalization. Hare's latest account of supervenience, as we shall see, is no exception here.

3.9 Hare's Latest Account of Supervenience

In order to elucidate what is involved when '*Fa*' ascribes a supervenient property to *a*, Hare examines the following syllogism, where '*Ga*' denotes a subjacent property (*ET*, p. 69, 1984b):

---

54. Hare mentions the supervenient character in *FR* but it plays no explicit part in his attempt to establish the universalizability of moral judgements (*FR*, p. 19)

55. Hudson gives an accurate account of Hare's view on supervenience. However, on one point it needs a revision. According to Hudson (1978, p. 183), Hare maintains that "value-judgements alone are supervenient". Whether or not this claim at that time was justified, Hare makes it clear in 1984b that he now regards supervenience as a feature "not just of evaluative words, properties or judgements, but of the wider class of judgements which have to have, at least in some minimal sense, reasons or grounds or explanations" (*ET*, p. 66, 1984b).

p: For all x, if *G*x then *F*x

q: *G*a

So **r**: *F*a

According to Hare there are two kinds of supervenience relations that must be distinguished from each other. However, although different they have this in common that when *Fa* ascribes a supervenient property, the following holds true:

> Necessarily, if **r**, then there is a valid inference of the '**p**, **q**, so **r**' form, the two premisses of which hold (*op. cit.*, p. 70).

The necessity involved here is, according to Hare, "conceptual or logical necessity". Consider the following passage:

> There is, all the same, a necessity *that there be* a universal premiss (itself perhaps optional) from which, in conjunction with the subsumptive premiss, the conclusion necessarily follows. To use an example I have used before (H 1963b s.f.): the food was poisonous because it contained cyanide; and to say this is to say *that there is* some universal premiss which holds (which might be, in this case, 'Everything containing cyanide is poisonous') and some subsumptive premiss which also holds (e.g. 'This food contains cyanide') and that the two premisses together logically entail the conclusion 'This food is poisonous'. It is not to claim that the universal premiss holds necessarily (*ET*, p. 75, 1984b).

Given that 'poisonous' is a descriptive term, the above passage suggests that what Hare means by 'supervenience' could be stated in the following way: What is involved in ascribing a supervenient property is that there is by logical necessity a universal premiss from which, in conjunction with a subsumptive premiss, the conclusion necessarily follows. According to Hare, moral judgements and descriptive judgements have in common this feature, that there is of necessity a universal premiss involved. This claim is compatible with the account in terms of meaning-rules given in, for instance, *FR*.

Now, the two kinds of supervenience that Hare speaks of are what he calls *substantial* and *trivial* supervenience. The distinction can be brought out by considering the nature of the universal premiss in **p** above. Hare considers three possibilities. The universal premiss can be (i) true in virtue of the words, (ii) necessarily true in some other sense than logical, and

finally it can be (iii) neither (i) or (ii) but held in the sense of being subscribed to. These are the options that Hare discusses.

The substantial cases of supervenience are those in which the universal premiss is not true in virtue of the words, as is the case with trivially supervenient terms.[56]

According to Hare, substantial supervenience is what is involved in moral contexts. Moreover, in moral cases, he thinks the universal premiss must satisfy two further requirements. First, it must not contain, as has already been mentioned, any ineliminable references to individuals.[57] Thus, he says about the universal premiss which a singular moral judgement is dependent on, that it is not:

> allowed to contain references to individuals. The doctrine of the universalizability of moral judgements standardly imposes this requirement, which is the second main element, besides supervenience, in the doctrine. Supervenience by itself does not impose the requirement; it only says there has to be some universal premiss (which might be universal in some weaker sense not requiring the elimination of individual references, but only universal quantification). However, it is generally held that moral principles have to be universal in a strong sense which excludes such individual references (*ET*, p. 74, 1984b).

Secondly, the universal premiss in at least moral and causal contexts must be nomological in character (*ET*, p. 75, 1984b).

At this point, let me connect with the discussion earlier in section 3.4. The fact that Hare puts forward (notably in *FR*) on the one hand a parallel between moral judgements and descriptive judgements, and on the other a claim to the effect that moral terms are supervenient terms, raises the question about the relationship between these two explanations as to why we must universalize our judgements. As may be recalled from the earlier discussion, several writers maintain that it is misleading to say that descriptive judgements and moral judgements are universalizable in the same sense, as Hare did in *FR*.

---

56. See here Schrag (1975, p. 80), who seems aware of this possibility. On this issue, contrast also Singer (1985, p. 56)

57. Narveson says that "Hare uses the word 'universalizable' as a synonym for 'supervenient', e.g. [*FR*], p. 10, and his recent [*MT*], p. 7, among many other places" (1985, p. 42). Apparently Narveson has not read "Supervenience" in which Hare makes it clear that the universalizability-thesis consists of two theses, of which the supervenience thesis is one (*ET*, p. 74). Besides, it must be noticed that page 10 of *FR*, to which Narveson refers, does not contain the term "supervenience", nor does it otherwise support his contention. It is therefore unclear what Narveson's reasons are for his claim.

At first sight a conclusion similar to the one reached by the above-mentioned writers could be drawn from Hare's recent statement of the supervenience relation. That is to say, from the premiss

(**M1**) Moral judgements are universalizable in the same sense as descriptive judgements are

and what we now can ascribe to Hare, viz.,

(**M3**) Moral judgements are universalizable because moral terms are supervenient terms

it seems to follow that descriptive terms must also be universalizable because they are supervenient terms. But, it has been argued, since at least certain descriptive terms, do not ascribe supervenient properties, Hare must give up (**M1**).

Colour terms are often given as examples of descriptive terms that are not considered to be supervenient terms.[58] The reason for this, I suspect, is the following: Physics teaches us that it is in virtue of some underlying molecular structure of the coloured object that make us see it as, say, red. However, this is an empirical claim, and there appears to be no logical blunder involved in imagining a logically possible world where such a claim would be false. In other words, if it is necessarily true that if an object has a certain molecular structure, it is red, the necessity here is not logical necessity. However empirically improbable it is, colours could logically exist without there being any further underlying fact. This door, however, does not seem to be open with regard to moral properties. If some act or person is good, it is always because of some good-making properties of the act or person, and not merely because the act or person possesses the property of goodness. The necessity of there being a (or a certain) law or generalization, with regard to moral properties, is not merely empirical improbability, but something much stronger.

Now, by distinguishing between the two kinds of supervenience above, Hare appears to concede this criticism. Descriptive terms are, on Hare's

58. Not all writers agree that there is a difference, with regard to supervenience, between moral properties and secondary qualities. Notably McDowell's (1985) argument for "moral realism" has been met with interest. The argument rests to a great extent on an analogy between moral properties and secondary qualities. See here also Peter Sandøe (1988, chapter 4). A recent critical examination of McDowell's views is found in Alan H. Goldman 1987. See also Williams 1985, pp. 149–150.

account at least, not substantially supervenient terms, but ascribe only trivially supervenient properties, whereas moral terms are supervenient in a substantial way. 'Red', for instance, would not be a substantial supervenient term, but merely trivially supervenient. On the other hand, 'red' would have that in common with a substantial supervenient term such as 'good' that they both would by logical necessity involve a universal premiss of the above kind (cf. *FR*, p. 11).

This idea - that substantial supervenience must be separated from what he calls trivial supervenience - is brought out in the following passage:

> Note also that in order to separate off cases of 'substantial supervenience' from cases which are supervenient in only a trivial sense we have to insist that the universal premiss should not be analytically true, or true in virtue of the meanings of words. In this trivial way a great many descriptive properties are supervenient (*indeed all, if we admit even unformulated meaning-rules as universal premisses*) for the same reasons as all descriptive statements are trivially universalizable (*ET*, p. 70, 1984b; my italics).

Thus, Hare apparently maintains that descriptive properties are supervenient in a trivial way, whereas moral properties are supervenient in a substantial way. A further passage which illuminates his position is the following one:

> For supervenience is a feature, not just of evaluative words, properties, or judgements, but of the wider class of judgements which have to have, at least in some minimal sense, reasons or grounds or explanations. And, as we shall see later, it is possible, at a certain cost in queerness, to remain a descriptivist or a realist, even of a non-naturalist stamp, and still believe in supervenience (*ET*, p. 66, 1984b).

Words such as 'red' and 'bachelor' would on Hare's account be trivially supervenient terms, since there is of necessity an analytically true universal rule (x) (*Gx* --> *Hx*) involved. That *Gx*, in the case of 'red'[59], for instance, is the *sui generis* property of being red would thus not count as an obstacle to saying it is a trivially supervenient term.

However, this sense of "trivial supervenience" is puzzling. Between

59. It is a peculiarity of colour terms that we say of objects that they are, for instance, red, not in virtue of some underlying natural property but only because they have the property of redness. Of course, if I predicate of *a* that it is 'red' I must predicate of x that it is red, if x is similar to *a*. However, what logically prohibits me from making divergent judgements here, is not that the two objects have identical sets of underlying natural properties, but that both objects have the property of being red. Thus in Hare's sense of 'supervenience' the fact that '*Gx*' in the universal premiss at issue would refer to the property of being red would not disqualify this from being supervenient in a trivial way.

exactly what is this attenuated supervenience relation said to hold? On what does the property of being 'rotten', for instance, supervene, if I consider 'this stick is rotten' to follow from a meaning-rule (e.g. 'Anything exactly similar to this stick is rotten'), which is true in virtue of the words involved (and a certain subsuming premiss)? Surely it cannot be that it is the analytic *meaning-rule* which supervenes on some subjacent *properties*? It would be equally strange to say that the *term* 'red', for instance, would supervene on the *property* of redness. To speak of terms supervening on properties in a discussion of whether some properties are consequential on others is, I believe, confusing. The supervenience relation is bewildering enough as it is without making it hold across what seem to be different categories, viz., properties and words/meaning-rules.

What Hare's account of supervenience shows is that he regards universalizability and supervenience to be similar in that both involve a universal premiss.[60] However, universalizability is, as mentioned, a more comprehensive concept in that it contains the further requirement that the universal must not contain any ineliminable references to individuals.[61]

Let me now round off this chapter by considering an objection to Hare's claim of what is involved in substantial supervenience.[62]

3.9.1 Substantial Supervenience

Now a problem arises that concerns whether Hare considers his explanation of the supervenience relation to state a necessary and/or sufficient condition of what is involved in ascribing a supervenient property. A. J.

---

60. In *FR* Hare says that a person who considers goodness to be a *sui generis* non-natural property could admit that moral judgements are universalizable. Hare then claims that such a person "would be maintaining a thesis which is obviously false (for the reasons given in *LM* 5.1 ff)" (*FR*, pp. 19). The argument to which he refers in *LM*, is, I believe, that moral terms are supervenient terms. Given what he now says in "Supervenience", his argument must be that 'good' is a substantial supervenient term.

61. Hare's explicit claim that in moral contexts the universal rule must not contain ineliminable reference to individuals, would appear to suggest that in descriptive context the universal rule could. This would, in its turn suggest a further difference between descriptive and moral judgements. However, as far as I can see, a term containing a meaning-rule involving an ineliminable reference to an individual would not be a descriptive term, but rather a name.

62. In light of the fact that Hare stresses in *FR* that evaluative terms do not have their meaning fixed, the question of whether the arguments at issue are also extensionally equivalent depends on how "nomologicality" and "subscribed to" should be understood. However, I do not see any reason for interpreting Hare as requiring more of the speaker in "Supervenience" than he does in *FR*.

Dale (1985), for instance, assumes that Hare's suggestion of what is involved in ascribing a substantial supervenient property states a necessary and sufficient condition. Dale then brings to attention that on such a reading it can be shown that if there is a supervenient property, then all contingent properties are supervenient. The argument runs as follows: Assume that there is some supervenient property *F*. Given what Hare says there then exists a universal of the form (i) "(x) *G*x --> *F*x". Now, Dale points out that for any arbitrary property *H*, (i) entails

(ii) (x) {[(*H*x v *G*x).(*H*x v –*F*x)] --> *H*x}.

Thus, if there necessarily exists a universal of the form (i) then there necessarily exists a universal of the (ii) form, which itself has the form (iii) (x) (*G'*x --> *H*x). In other words,

> if there is a supervenient property then all contingent properties are supervenient (Dale 1985, p. 600).

Since Hare is here discussing substantial supervenience, the consequence to which Dale draws attention is clearly not desirable. In order to avoid it Hare must place some restrictions on the antecedent of the universal "(x) *G*x --> *F*x". It is not all together clear to me just how this could be done. Perhaps if we took "*G*x" to refer to a list of all of x's properties, and required that, to avoid circularity, no property were mentioned more than once, we would, it seems, solve the problem. Although Dale appears to agree that such a move clears away the problem, he believes it gives rise to the further complication of how to separate properties from each other (Dale 1985, p. 600 s.f.). Should Hare then be interpreted as suggesting a necessary and sufficient condition? Recently, Horacio Spector (1987) has objected to Dale's criticism that it takes for granted that Hare would view his suggestion as stating a sufficient condition. Spector notes that Hare speaks of what is "involved in" ascribing a supervenient property, which he thinks lends some support to an interpretion that differs from Dale's, namely that Hare is only speaking of a necessary condition.[63] Moreover, Spector draws attention to what he calls an "element of supervenience"

63. Recall that the universal premiss in causal and moral cases should be nomological, which also suggests that Hare does not have a sufficient condition in mind.

that once realized will make Dale's objection vanish.[64] He expresses his point in the following way:

> It is sometimes said that if *Fa* ascribes a supervenient property to *a*, then the question 'Why *a* is *F*' has necessarily an answer, and this answer must state something that is not stated by *Fa*. In other words, if *Fa* ascribes a supervenient property to *a*, then there must be a *sufficient* reason why *a* is *F*, and the giving of a sufficient reason R of a fact E conceptually excludes the repeating, wholly or partly, that E takes place (Spector 1987, p. 94; cf. Caton 1963, p. 51).

The supervenient expression *F*x, Spector argues, must be viewed as being logically independent of the subjacent expression *G*x. But the complex property which Dale mentions, viz., the one of 'being *H* or *G*, and *H* or not *F*' is logically dependent upon the propery of being *H*.

Spector's article captures an important feature of the supervenience relation, which appears much in line with what Hare says about this issue. Whether Hare would agree entirely with Spector's account is a question we need not go into more deeply. That Hare never aimed at stating a necessary *and* sufficient condition of what is involved in ascribing a *substantial* supervenient property is a plausible interpretation. Just how we should avoid the consequences to which Dale draws our attention is a matter on which we may hope that Hare will soon shed some light.

There is a further, and more general issue that needs to be elucidated. However, since Hare does not touch upon this subject I will here confine myself to making a brief comment. It concerns to what extent a prescriptivist, such as Hare, is served by explaining the logical behaviour of moral terms by invoking the phenomenon of supervenience. Hare uses, for instance, the expression "ascribe supervenient properties" (e.g. *ET*, p. 70) about moral terms or moral judgements. However, supposing that moral terms and moral judgements are primarily prescriptive, as Hare argues elsewhere, it is unclear exactly how we should understand the expression above. As we have seen, what a person primarily does in saying that some x is good, is not telling that something is the case. The speaker is not primarily ascribing a property but rather telling the listener that something should be made the case.[65] Presumably the phrase above could be made

64. Spector's argument will also affect Dale's second objection, viz., that given Hare's characterization, logically impossible properties will pass as supervenient properties.

65. Cf. what is expressed, for instance, in Hare 1985, p. 46: "when we say something moral, we do not have to be ascribing any kind of property, subjective or objective".

consistent with a prescriptive analysis of moral terms. Hare owes it to the reader to show this.

IV. Relevant Facts and Rationality

4.0 Introduction

As may be recalled from the introduction, Hare distinguishes in *MT* between two levels of moral thinking, namely critical and intuitive thinking. What critical thinking consists in - which is what interests me in this work - is expressed in the following passage that I will call *T1*.

> Critical thinking consists in making a choice under the constraints imposed by the logical properties of the moral concepts and by the non-moral facts, and by nothing else (*MT*, p. 40).

Moreover, as was also shown, Hare makes the further contentious claim *T2*:

> if we assumed a perfect command of logic and of the facts, they would constrain so severely the moral evaluations that we can make, that in practice we would be bound all to agree to the same ones (*MT*, p. 6).

Hitherto I have been concerned with the logical properties which Hare thinks that moral terms have, and which are relevant for *T1* and *T2*. These were, as we saw, universalizability and prescriptivity. Before examining the third key thesis, which concerns Hare's views on what I will call conditional representation, it will be convenient to consider in more detail just how we should understand *T1*. In this chapter I will, in other words, prepare the ground for an assessment of *T2* by outlining how Hare thinks that taking account of the facts and the meaning of 'ought' will constrain our moral reasoning. Doing so enables me to focus in the next chapter on his view from *MT* on conditional representation, which together with his distinction between critical and intuitive thinking, constitutes the major difference from his earlier works.

In section 4.1 I set out to outline more in detail just how, according to Hare, the logical properties of prescriptivity and universalizability constrain our moral reasoning. I will do so by considering an example from *FR* - sometimes referred to as an example of Harean Golden Rule argument - that has been much discussed, and which is instructive for my purposes. In

4.2–4.3 I will be concerned with Hare's claim that we are rationally required to take facts into consideration before deciding what we are ready to prescribe universally. That we should take facts into consideration before passing moral judgements is hardly a controversial claim. What is controversial is whether Hare can do without any substantial criterion of what are the relevant facts. There is, for instance, ample evidence that Hare thinks that to omit to take facts about other people's preferences or, in general, motivational state into consideration, when we are prescribing universally, is to prescribe irrationally (e.g. *MT* p. 221). This raises the further question of how Hare argues for the relevance of these particular facts. Although a complete answer to this question will have to await an examination of his views on conditional representation – for reasons that will hopefully be clear then – sections 4.2–4.3 contain in addition a presentation of the central part of his view on relevance. Section 4.4, in its turn, will be devoted to a passage from *MT* in which Hare outlines how we should go about determining the relevant facts in a certain kind of case, namely, when we are not guided by *prima facie* moral principles from the intuitive level. Examining this passage gives me an opportunity to consider some objections that have been put forward against Hare's views – notably a fairly frequent one that concerns his version of the so-called Golden Rule argument.[1] I round off this chapter by offering a first, preliminary interpretation of *T1*.

In the next chapter I will follow up the discussion by a critical examination of some important problems that I connect with critical thinking and his *T2* claim.

4.1. The Creditor and the Debtor

The person who is reasoning on the critical level – I will henceforth refer to him as the "critical thinker" – is a person who will try to consider only facts and logic when answering his moral 'ought'-questions. Whatever moral principles or intuitions he has endorsed in the past, he must try to set these aside when reasoning on the critical level. This will clearly often be difficult, as Hare himself recognizes. Some of us will be able to do this better than others. People with weak moral intuitions, for instance, will have less difficulties in putting these aside than persons who have firm intuitions

1. The process of putting onself into the shoes of someone else has in recent literature, following MacKay (1986), been referred to as the mental shoehorn manoeuvre.

about moral issues. Let us, however, concentrate on the critical thinker who feels confident in his abilities with regard to this matter. How will he proceed?

Hare sometimes compares the method of moral reasoning which he advocates with Popper's method for testing scientifical hypotheses:

> Just as science, seriously pursued, is the search for hypotheses and the testing of them by the attempt to falsify their particular consequences, so morals, as a serious endeavour, consists in the search for principles and the testing of them against particular cases (*FR*, p. 92).

Moral reasoning, as Hare conceives it, is not a matter of linear reasoning from facts to substantial moral principles. Rather, by paying attention to logic and facts, we will be able to reject and single out those 'ought'-judgements that we cannot hold in an evaluative way from those that we are ready to endorse. An important difference, though, between scientific and moral reasoning is that, whereas suggested universal hypotheses in science are rejected by observational statements, moral judgements are rejected, according to Hare, not by facts but by the desires, preferences, of the critical thinker.[2]

In order to get a clearer and more detailed picture of *T1*, it will be wise to consider one of Hare's own examples. The following one from *FR* is a fairly simple one that Hare thinks displays the "bare bones" (*FR* p. 90) of his method:[3]

> *A* owes money to *B*, and *B* owes money to *C*, and it is the law that creditors may exact their debts by putting their debtors into prison. *B* asks himself, 'Can I say that I ought to take this measure against *A* in order to make him pay?' He is no doubt *inclined* to do this, or *wants* to do it. Therefore, if there were no question of universalizing his prescriptions, he would assent readily to the *singular* prescription 'Let me put *A* into prison'. But when he seeks to

2. Hare's view on moral reasoning has sometimes rendered him the epitet "quasi existentialist". See e.g. Guisán (1986, p. 246) and Norbert Bilbeny (1990 ff. 102). Some point of contact does exist, but any deeper connection is difficult to spot. Notwithstanding, Hare does sometimes express himself in an "existentialist" sounding way. In an interview with Magee (1978, p. 163), for instance, Hare commented that "If you want facts, you must look; and if you want values, you must choose". See also Griffiths 1983, p. 502.

3. In *MT* Hare says the following about the section from which the example comes: "It is now time to set out in more detail how the logical properties of the moral concepts help us to construct moral arguments. I have already done this in a preliminary way in *FR* 6.3.ff. The method I shall be outlining here will be the same in essentials; but much has been written on the the subject since then, and I have myself seen some of the moves more clearly (Hare 1976a, 1978b)" (*MT*, p. 87).

> turn his prescription into a moral judgement, and say 'I *ought* to put *A* into prison because he will not pay me what he owes', he reflects that this would involve accepting the principle 'Anyone who is in my position ought to put his debtor into prison if he does not pay'. But then he reflects that *C* is in the same position of unpaid creditor with regard to himself (*B*), and that the cases are otherwise identical; and that if anyone in this position ought to put his debtors into prison, then so ought *C* to put him (*B*) into prison. And to accept the moral prescription '*C* ought to put me into prison' would commit him (since, as we have seen, he must be using the word 'ought' prescriptively) to accepting the singular prescription 'Let *C* put me into prison'; and this he is not ready to accept. But if he is not, then neither can he accept the original judgement that he (*B*) ought to put *A* into prison for debt (*FR*, pp. 90–91; cf. *MT*, pp. 108–109).

This much discussed example exhibits the essential features of a moral reasoning conducted in accordance with logic and facts. A caveat must here be inserted: as the example is outlined, person *c* is an actual person who stands in an exactly similar relation to *b* as *b* stands to *a*. However, as shown in chapter III, and as Hare later on in *FR* stresses,

> it is not necessary for the force of the argument that B should *in fact* stand in this relation to anyone; it is sufficient that he should consider hypothetically such a case (*FR*, p. 93).

For this reason, I will leave out any reference to a third person. Nothing substantial is lost by letting the creditor consider a hypothetical situation rather than an actual, analogical situation.[4]

With this simplification in mind, let us consider the example. To begin with, the creditor faces a situation that makes him wonder what he ought to do. He asks himself whether he is ready to assent to 'I ought to put the debtor into prison in order to make him pay'. We saw also that the creditor actually wants to put the debtor into prison in order to make the debtor pay his debt. What prevents him from saying 'I ought to put the debtor into prison' is that he is logically committed to universalize his judgement. He realizes, in other words, that it follows from universalizability that if he now says that he ought to put the debtor into prison, he is logically committed to the view that the very same thing ought to be done to him, were he in the

4. Contrast Pollock (1972) who makes it an issue that Hare's Golden Rule argument in contrast to other examples centres on analogical situations rather than hypothetical ones. Pollock seems to have overlooked Hare's claim that even with regard to the creditor's case there does not have to be an actual similar case. Perry (1976, p. 182) raises the objection that "even while restricting oneself to 'universal' characteristics, one can easily make one's own case practically unique, if not logically unique, by listing all kinds of things that others would doubtless think morally irrelvant". However, this objection misses its target if Hare is right in that only hypothetical situations will be invoked.

debtor's situation, including having the same personal characteristics and in particular the same motivational state as the debtor. In Hare's example it is safe to assume that the debtor does not want to be put into prison.

However, invoking universalizability alone will not explain why the creditor in the end decides not to follow his original inclinations. To explain this we need to bring in the further claim that the creditor must in addition prescribe that he too ought to be imprisoned, in the hypothetical situation in which he is in the position of the debtor, with the debtor's universal properties. To issue such an 'ought'-judgement would mean, in virtue of what Hare claims is involved in prescribing, to accept the singular prescription 'Let x put me into prison'. Moreover, to accept this sincerely, as may be recalled from chapters I and II, it is for logical reasons necessarily the case that the debtor has a desire to be put into prison. But since the creditor does not want, *ex hypothesi*, to be put into prison, he cannot sincerely accept the prescription 'Let x put me into prison'. The creditor therefore faces the following dilemma: Either he rejects the judgement 'I ought to put the debtor into prison' or he makes himself guilty of making two logically inconsistent (cf. *FR* p. 33) judgements about two situations that are exactly or relevantly similar with regard to universal non-moral properties. To avoid the latter alternative - to avoid this, as Hare puts it in *MT*, "Contradiction in the will" - the creditor therefore concludes that he cannot prescribe that the debtor ought to be put into prison.

Hare's creditor/debtor example is aimed at showing how the logical properties of universalizability and prescriptivity, together with the facts of the situation, constrain what moral judgements a person can prescribe. But if this is the case, should not Hare's method then rather be characterized as a method for reaching knowledge about answers to 'What ought I to do?' that the individual *cannot* embrace (cf. *FR* p. 93, *MT* p. 112)? It has sometimes been held against Hare that his method will not help us find out what we ought to do.[5] It will at most make clear what ought-judgements we cannot endorse. This is a matter of some substance that ultimately derives from Hare's claim that there are, as may be recalled, three positions to take when considering what ought to be done. The creditor, for instance, may assent to 'I ought to put the debtor into prison', or 'I ought not to put the debtor into prison' or he may take an entirely different stand to the

5. For a discussion on whether Hare's work explains satisfactorily the step from having rejected a moral judgement to actually coming to embrace one, see Hoche (1983) and Kese (1990).

problem and not make any judgement at all, or make a judgement of indifference such as 'It is neither the case that I ought to put the debtor into prison, nor the case that I ought not to put the debtor into prison' (cf. *FR*, pp. 100-101; *MT*, p. 185). Rejecting, as in the creditor's case, the first one leaves the creditor with an option. He can maintain that he ought not to put the debtor into prison or he can abstain entirely from making any moral judgement: He can take the position of what Hare calls an amoralist and refrain from making moral judgements or make only ones of indifference of the form: 'It is neither the case that I ought, nor the case that I ought not . . .'. According to Hare this latter possibility is logically open to him. Moreover, it is consistent with assenting to the singular prescription 'Put the debtor into prison'.

As mentioned earlier, Hare thinks there are prudential reasons for not becoming an amoralist. This claim, as was shown, rests in part on some empirical assumptions (cf. Singer 1988, p. 153) which I find reasonable. Let me therefore here put this issue aside, and concentrate on cases where the desire to make a moral judgement outweighs the desires of the amoralist (for Hare's view see *MT* 10.7).[6]

Let me now turn to the first matter I wish to look into, namely Hare's claim that we are rationally required to take the facts of the case under consideration.[7]

4.2. Facts and Moral Reasoning

To begin with let us examine in more detail the arguments that Hare, in *MT*, endorses for the requirement that the facts of the case must be taken into consideration:

---

6. Hare's view on amoralism raises questions that I will not deal with in this work. For a recent interesting discussion see Sandøe (1989a) and Hare (1989e).

7. It has been suggested by Locke (1981) that Hare introduces a new argument in Hare (1976a) that was not found explicitly in *FR*. In contrast to the latter work, the requirement that we should consider the interests of people impartially is in Hare (1976) not only "an argument . . .from universalizability, prescriptivity, and role reversal alone. It is based rather, on an appeal to prudence, or rational self-interest, on the part of the person making the moral judgement"(*op. cit.* p. 545). Keekok Lee (1985, p. 64: see also p. 139), in comparing *FR* and *MT*, makes a similar observation. I agree with these writers, if they are suggesting that Hare makes it clearer in works after *FR* what follows from being a rational, prudent person who considers what to universally prescribe. However, I would like to make one proviso: Hare should be understood as speaking of rational hypothetical self-interests rather than self-interest (see e.g. *HC* p. 218 *s.f.*).

> It is irrational to prescribe anything without regard to what, concretely, we are prescribing; and this involves cognizance of what our prescription means, and of what its execution in this concrete situation would entail. Obviously we cannot be cognizant of this without attention to the facts of the situation (12.4). If this is right, then it can be shown that rationality in prescriptions depends on cognizance of facts, without bringing universalizability into the argument. When the prescription is universal or universalizable, however, the requirement in consequence becomes stronger; we are required to satisfy ourselves that we can accept the *universal* application of the prescription; and this includes its application were we in the other's position (*MT*, p. 89).

From this passage, which contains not only a claim about the rationality of prescriptions in general, but also one about the rationality of universal prescriptions, we may isolate the following reason:

> *R1*: Rationality in prescription depends on cognizance of what our prescriptions means, and of what its execution in the situation under consideration would entail (*MT*, p. 89).

Now, by turning to the paper "What Makes Choices Rational?" (Hare 1979c), it will become clear what Hare has in mind when he says above "this involves *cognizance* of *what our prescription means*, and of what its execution in this concrete situation *would entail*" (*loc. cit*; my italics).

Hare maintains that there are two important elements in the rationality of prescriptions:

> The first is to understand what the prescription which expresses the choice *means* in the strict sense of understanding what the *words* mean, or what the speaker means by his words. The second (which is also sometimes confusingly expressed by asking what the choice *means*) is to find out what, concretely, the different answers to the question, the different choices, *involve* by way of actual differences in the future history of the world. To find this out, we have to make certain *factual* predictions about the consequences of carrying out the prescription 'Yes, please do', and of carrying out the negation of this prescription 'No, thank you' (*ET*, p. 37, 1979c).

Hare approaches the first point by considering the *prescription* "Yes, please do" issued as an answer to the question "Shall I pour you a cup of coffee?" To rationally issue the prescription "Yes, please do" one needs to understand it. But this is, according to Hare, not sufficient.

> Secondly, it might be that, although I knew the language and therefore knew I would get coffee, I did not know the properties of the coffee I was being offered. I might not know what *sort* of coffee it would be; for example, perhaps we are in Egypt and the coffee is Turkish, which I have never had and shall much dislike when I taste it (*ET*, p. 36, 1979c).

There is not much to say about the first point. I suspect that in most situations it is as irrational to make a request for a cup of coffee if one does not know what 'coffee' means, as it is irrational in most situations to accept any offer that one does not understand. Still, although this may generally be the case, it is questionable whether it is always the case. Hare says in one place that he does not want to deny "that it is sometimes rational to deprive oneself of information" (*ET*, p. 38, 1979c). Moreover, he speaks of both these criteria of rationality as being the "main criteria" of rationality of prescriptions and choices (*ET*, p. 37, 1979c). Sometimes an inquiry into the nature of what you are offered could be understood as an insult or as a sign of mistrust. In a situation where, say, your life is in the hands of some maniac who threatens to kill you, you may have good reason to believe that an acceptance to drink, say, a *carajillo*, a liquid which is unknown to you, will give you a better chance to save your life. Turning down the offer could make the maniac furious. In such a situation the rational thing to do, all things considered, seems to be to accept the *carajillo* (coffee and brandy) from the maniac.[8]

What should we then say about the second point, that we need to know the properties of the particular coffee? The only way to reach knowledge of some of the particular properties, is, it seems, to actually try the coffee. To know whether I want this coffee, I must on most occasions know, among other things, whether I like it; and to know whether I like it, I must try it. However, it seems odd that rationality of prescriptions should depend on the knowledge obtained by actually prescribing ('Please give me a cup'). Hare maintains in the former passage above that we must make "*factual* predications about the consequences of carrying out the prescription". This suggests indirectly that he thinks that the knowledge gained by actually trying the coffee is unavailable. This knowledge is unavailable since knowledge of the properties involved presupposes that I do precisely that which I wonder whether I rationally should do. Nevertheless, they are relevant since having a knowledge of these facts could make me regret having prescribed as I did.[9]

---

8. Notice here that the person who accepts the offer does so by issuing a prescription, viz., 'please, give me a cup of coffee'. I have, in other words, not shifted attention here to what might be considered as another problem, viz., that of what is rational for the person to whom a prescription is addressed.

9. Perhaps we should interpret Hare as maintaining that, to draw on Russell's distinction between "knowledge by description" and "knowledge by acquaintance", only the former kind is available (Russell 1905)? For a recent discussion of Russell's idea, see

Some pages further on, Hare makes the following claim that suggests a way of reformulating *R1* in terms of relevant available facts:

> What is irrational is to ignore relevant facts which are available, or to fail to seek information which would be relevant and could be readily obtained (*ET*, p. 38).

Now, this suggests that we may reformulate *R1* as stating a sufficient condition:

> *R2*: A prescription is rational if formed in the light of relevant non-moral facts that are available.[10]

R2 corresponds well with how he uses 'rational' in the following passage, for instance:

> It is not necessary, in order to show the rationality of moral thinking, to bring into existence any specifically moral facts additional to these ordinary facts ([Hare 1979c]). If somebody in his moral thinking has taken cognizance of such ordinary facts as are available and relevant (3.9), he has done all the fact-finding that rationality requires (*MT*, p. 217).

R2 seems a reasonable interpretation. However, I must confess that I find Hare's notion somewhat difficult to grasp.[11] The want of an account in terms of necessary and sufficient conditions for rational prescription, makes it difficult to evaluate Hare's notion properly.[12] Moreover, a warning must be voiced here. Hare sometimes speaks of rationality of prescriptions (e.g. *MT*, p. 21, 22), sometimes of rationality of preferences (*MT*, II.8.7). However, as long as it is recalled that, according to Hare, "all prescriptions, including moral ones, are expressions of preferences or of desires in a wide sense" (*MT*, p. 185), this should not give rise to any misunderstandings.[13]

McDowell (1986) in McDowell and Pettit (1986).

10. Since we are here concerned with critical thinking, it is reasonable to add the expression 'non-moral'. See also the passage that follows immediately afterwards.

11. Göran Hermerén drew my attention to the fact that *R2* is unclear with regard to whether it is sufficient to take *some* relevant fact into account, or whether we must take *all* the relevant, available facts into account. Although Hare has not made this clear, the latter option seems the more reasonable.

12. I will return to Hare's view on rationality in section 5.5, where I will be concerned with aspects of his view which there is no need to bring up here.

13. In *MT* Hare says "It will be convenient if we continue sometimes to speak, not in terms of motivational states, nor even of preferences, but in terms of the prescriptions

Thus, *R2* is formulated in terms of 'prescription', but could in Hare's account have been formulated as a requirement on desires and preferences.

Is *R2* acceptable? To answer this, we need to be clear about Hare's view on what constitutes a relevant, available property. At first sight, the underlying idea of *R2* seems reasonable: if I consider some facts to be relevant for whether or not I should accept an offer (and by extension, issue the prescription 'Please, give me some coffee), I would be irrational, it seems, if I did not inform myself whether these relevant properties were present or not. That some property is relevant entails at least that it is worthy of consideration, and to disregard what we believe is worthy of consideration must, in any sense of 'rational', be judged irrational.

4.3 Relevant Available Facts

R2 gives rise to two issues. To begin with there is the problem of what constitutes an *available* fact. Hare does not further elucidate his notion of an available or readily obtained fact. This is unfortunate, since it is far from clear where we are to set the borders for which facts are to be considered as available or not. In the coffee-example above, it seemed reasonable to exclude the information that we would get from knowing what would happen if we actually did the contemplated act as being unavailable. But there are other cases when it is less clear when we are to say that a fact is available or not. Suppose I was offered to buy a journal of philosophy, the language of which I was unfamiliar with. Suppose further that I was told or suspected that it contained a very good argument for why it is rational to pay attention to people's preferences. Since the question interests me I would certainly like to examine the argument thoroughly. Of course, if a friend of mine knew the language and the article, then it would be irrational, it seems, not to pick up the telephone and make him restate the argument, before making a purchase of the journal (i.e. before issuing the prescription, say, 'Please, let me buy the article'). But what if I had no such friend? Should I begin to investigate who in my neighbourhood, town or country, spoke the language? Should I ultimately learn the language myself? None of these alternatives appear to involve any impossible task to undertake, and it is therefore not easy to say just where to draw the line. It seems safe to say that the extremes to which I actually go will depend on the strength of my desire. If I want very much to know the content of the

which are their expressions in language" (*MT*, p. 107).

article, I will go to great trouble to find the means to reach this knowledge. If my desire here is weak, I may not even try to put myself in contact with my learned friend who spoke the language, but merely make a note that if I see him, I will ask him about the article.

To avoid making 'available' too demanding, we might wish to limit its scope to the moment of decision. However, it is questionable whether this will do. Sometimes we are aware that we lack information that we predict will be relevant for our decision. For instance, I want very much to go to *a*'s party, only because I suspect *b* will be there. I know also that if I wait until tomorrow with my decision, I will know for sure whether *b* will show up or not. Moreover, I may be affected by a situation to such an extent that postponing my decision to a calmer moment would probable lead me to make another decision. In such circumstance it seems irrational to disregard the future information just because it is not available at the moment of decision (cf. Egonsson 1990, p. 97).

Thus, on this horn of the dilemma *R2* is not very enlightening. However, even if the problems associated with the term 'available' are overcome, other ones appear to tower aloft when we turn our attention to the term 'relevance'. Judging from the great number of writers that have argued that the kind of "Golden-Rule" or role-reversal argument which Hare endorses, by not providing us with any criterion of relevancy, is trivial or at least useless except that it logically prohibits us from invoking reasons that contain ineliminable references to individuals.[14]

Hare does deny that the universalizability-thesis by itself provides us with any criterion of relevancy. However, he argues, as may be recalled, that the method of critical thinking by

> itself will provide criteria of relevance *pari passu* with the selection of principles, both critical and prima facie (*MT*, p. 63).

Moreover, some lines further on he maintains that it is only when we are at the intuitive level that it will be a problem to select which properties are relevant or not:

---

14. Various writers have expressed the idea that the universalizability-thesis is something trivial and of little or no use in moral reasoning. Silverstein (1974) is one of to those who have most forcefully endorsed such a view. Other writers with similar views include Gilbert (1972), Pollock (1972), Monro (1967), and Emmet (1963). In the next chapter I will have occasion to consider a more recent criticism by Pettit (1987).

> It is only when we come to intuitive thinking, guided by relatively general prima facie principles, that we need to be able to pick out the morally relevant features of situations, so as to leave out of consideration all the other features (*MT*, p. 63).

Notwithstanding, several pages later on he qualifies this statement considerably, maintaining now that even on the critical level we need to make judgements of relevance:

> We cannot, when making a moral judgement, be expected to ascertain all the facts that there are (though an archangel would know them, and therefore what I am going to say applies only to human critical thinking, not to critical thinking as such). We need some way of selecting those facts to which we are going to attend. We need, that is to say, to make judgements of relevance (*MT*, p. 89).

In chapter V I shall have occasion to comment on the role played by the archangel in Hare's establishment of *T2*. Meanwhile, it is important that we get a clearer picture of what Hare means by saying that we humans need to make judgements of relevance.

Hare has addressed the topic of relevance in several of his works. In his paper "Relevance" (Hare 1978b), for instance, he sets out to explain how we decide what features of a situation or action are relevant to its moral appraisal. From the preceding chapters we already know that, according to Hare, two features are irrelevant on what he calls purely formal grounds. Firstly, the universalizability-thesis requires that our reasons must not contain any ineliminable reference to individuals. Secondly, other features are irrelevant because a mention of these would constitute a breach against the idea, often expressed intuitively as "'Ought' implies 'can'".[15] In "Relevance", for instance, Hare expresses this in the following way:

> if I am being blamed for missing an appointment, the fact that there was a flight at such-and-such a time which would have enabled me to keep the appointment if I had caught it is irrelevant, if I could not have caught it (*ET*, p. 191).

With these two restrictions in mind, let us see what Hare has to say about relevance:

> The first thing that must be said is that the decision to treat certain features of a situation as morally relevant is not independent of the decision to apply

15. For a critical appraisal of this idea, see e.g. Margolis (1971, pp. 60–87) and Tännsjö (1974).

> certain moral principles to it, i.e. to make certain moral judgements about it. It is a great mistake to think that there can be a morally or evaluatively neutral process of picking out the relevant features of a situation, which can then be followed by the job of appraising or evaluating the situation morally. We can indeed *describe* a situation without committing ourselves to any moral judgements about it (that this is not so has become a dogma in some circles, but I know of no good reasons for accepting this dogma) (*ET*, p. 193; cf. *MT*, pp. 62–63, 90–91).

Thus, since moral principles, according to Hare, are universal, prescriptive overriding judgements, what in effect he is saying here is that what we consider as a relevant fact for making our moral judgement depends on what universal overriding prescriptions we are ready to issue sincerely. And this, in its turn, is determined by what our universal overriding preferences are.

Hare's notion of relevance is apparently a causal one (in this it is similar, for instance, to the one proposed by Brandt 1979, p. 12). A certain property will be relevant for person *a*, if a knowledge of this fact will actually affect *a*'s preferences (cf. Hermerén 1972; Millican 1983; Dahl 1987).[16]

Let us for a moment turn back to the creditor/debtor example. In order to make a rational decision concerning what he morally ought to do to the debtor, the creditor should try to ascertain the relevant available facts. The creditor must accept - in virtue of the universalizability requirement - that whatever he prescribes should happen to the debtor, also should happen to him, were he in an exactly similar situation to the debtor (with regard to universal properties). Those properties of the debtor will therefore be relevant that were they, *ex hypothesi*, properties of the creditor, he could not accept the prescription that he should be put into prison.

Often we will be guided by our prima facie principles when looking for relevant facts. However, there are situations, Hare thinks, when this is not the case. To get a clearer picture of what Hare means by saying that even

16. Both Millican and Dahl suggest interpretations of Hare's notion of relevance, with which on the whole I am in agreement. Millican (1983 p. 208) says, for instance "A fact is relevant if its non-existence would make a difference to the decision of our omniscient archangel". Cf. Dahl (1987, p. 400): "A morally relevant feature will be one whose consideration under conditions of rational thought will give rise to desires that determine which imperatives, and hence which universal principles, we can accept" (Contrast Roxbee Cox 1986). Dahl raises several interesting but, as I see it, external objections. For instance, he claims that Hare has not shown the moral irrelevance of "any number of features taken to be relevant by non-utilitarian theories" (*op. cit.*, p. 409), and he mentions that there are other features of actions other than those which have an affect on the desires on persons that are relevant. A relevant feature, he proposes, is whether the person's autonomy is respected or not (see also *op. cit.*, p. 410). Hermerén (1972, ff. 168) differentiates between two notions of relevance, viz., causal and logical relevance ("kausal relevans" and "logisk relevans").

on the critical level we must make judgements of relevance, let us consider a passage from *MT*, in which Hare sets out to explain how a rational agent should go about when he begins his critical thinking in the absence of *prima facie* moral principles.

4.4 The Moral Tiro

The case which Hare considers is the following one:

> If we are beginning our moral thinking, however, we are clearly not in a position to assess moral relevance in this way, because we have not yet decided what moral principles are to be applied. We have therefore to proceed initially by guesswork. If we think that a feature of a situation *might* be relevant, we experiment with principles mentioning the feature; to accept the principles will be to accept the relevance of the feature, and to reject them will be to reject *those* reasons why it might be relevant, though there may yet be *other* principles which make it relevant (*MT*, pp. 89, 90).

The idea expressed here merits closer scrutiny. Let us call the person who is experimenting in the way outlined above, the *moral tiro*. Thus, what is characteristic of the moral tiro is, it seems, that none of his moral principles (if he has any) applies to the situation. Moreover, what Hare seems to be suggesting is that in order to find out what moral judgement to hold, the moral tiro should proceed in the following way:

(*) The moral tiro should by guesswork select some feature of the situation, and then ask whether he is ready to accept a moral principle that mentions this feature.

But this advice is not quite clear. What does it mean to "select" some feature of the situation?

Assume that what the moral tiro believes is that one state of affairs *s1* is about to turn into another state of affairs *s2*, and that he wonders whether this ought to happen. I suggest that when Hare says that the moral tiro should proceed by guessing some feature, what he reasonably must mean is that the moral tiro ought to assume that what he thinks are going to be the facts of *s2* are the relevant facts. Given that the moral tiro has not full information about what will be the facts of *s2*, what he believes will take place in *s2* could then be considered as "some feature" of the situation.

The possible interpretation that I wish to guard against is brought out by the following: Suppose that the moral tiro believes that a tall person is about to aim a hard blow against a red haired child, holding an ice-cream in

his hand. Assume that the moral tiro has no moral principle that to his mind applies to the situation. Surely Hare is not giving the moral tiro the advice that in such a case he should pick out one feature of the situation, say, the fact that the ice-cream will with great certainty fall to the ground or that the child will get his hair messed up, and then by *disregarding* what he otherwise thinks will happen ask whether he is ready to accept a principle that mentions this feature. The rational thing to do for the moral tiro is to base his decision on all of his beliefs about the situation.

There are all sorts of problems involved in forming correct beliefs about a situation and about what will happen if we try to change the act.[17] Nevertheless, what seems reasonable to expect from the moral tiro is that he bases his moral prescription not on some of his beliefs about the situation, but on all of his beliefs that concern the situation. Having beliefs about what is going to happen, about who is who, who is doing what to whom, etc., the moral tiro should be able to form new beliefs and make predictions about what will be the case if some act is carried out or not. Clearly such beliefs and predictions can be formed in the absence of any universal overriding preferences about the outcome. Given this, what Hare means by 'guesswork' presumably must be that it is a guess from the moral tiro that what he believes are the facts also are the relevant facts of the case. Learning even more about the situation would perhaps show that he has missed some feature to the effect that he now no longer would be able to prescribe the same act.

The reason why we should interpret the passage in (*) above with caution is the following: By maintaining that the moral tiro should try his way with principles mentioning different features of the situation under consideration, grist is seemingly brought to the mill of those writers who object to the use of universalizability-tests on the ground that an act or situation can be described in several true ways. The universalizability-thesis cannot alone, it is argued, serve as a workable criterion for knowing what to do morally. This line of reasoning is most notably endorsed by David L. Norton (1980) and Kenneth A. Milkman (1982):[18]

17. Hardin (1988) contains a recent discussion of this topic. There are, he thinks, at least three obstacles involved in calculating the consequences of our actions: "(1) we lack the information required to carry out such calculations . . .(2) we lack relevant causal theories of the implications of our actions, and (3) we could not do the necessary calculations in any case because our minds have limited capacity" (*op. cit.*, p. 9).

18. Relevant here is also Rees, who deals with the problem by excluding those descriptions of an act or situation "which make no reference to any one's interests" as being morally irrelevant. Rees (1970–71, p. 250). Cf. Pollock (1972, pp. 31, 32), which

> in fact any act can be validly described in a great many ways, and its universalizability necessarily depends, in part, upon how it is described. If this is so, then universalizability as a valid criterion requires standardized act-description (Norton 1980, p. 522).

> Whenever I find myself in a situation marked by general descriptive characterisitics X, Y, and Z, I will perform in a given way. But the number of ways that an act or situation can be truly described is indefinitely large, so that the number of maxims which may be generated on the basis of these true descriptions is also indefinitely large (Milkman 1982, p. 20).

Another writer that pays attention to the above issue is Onora O'Neill (she is not, though, focusing on Hare's universalizability-thesis). She says, for instance, the following:

> The possibility of applying a theory or principle, at least in some cases, requires only that we have at least some strategies for selecting among the many true descriptions of a situation ones that are significant for moral decision (O'Neill 1989, p. 180)

> The problem of reflective judging is that any actual example may fall under many descriptions and so exemplify numerous principles or practices, many of them prima facie of moral significance (*op. cit.* p. 181; cf. O'Neill 1985, see also Emmet 1963, p. 219).

Thus, what the above writers all seem to agree on - at least the first two - is that the fact that an act or situation can be described "in great many" or "indefinitely" many ways, poses a problem for anyone who thinks that what moral judgement we decide on logically is determined by what moral judgement we are ready to universalize.

However, it is not clear just what their objection is. Admittedly, it will sometimes for psychological reasons be difficult to form a purely descriptive picture of a situation. Sometimes we may even be mistaken in thinking that we have described a situation in a way that is free from evaluations. However, as I understand it this is not the problem addressed by the above writers.[19] The fact that a situation can be described in several true ways does not pose, I believe, any serious problem for Hare's view. What the moral tiro has to do is to decide what to universally prescribe given what he

---

contains a discussion of the problem of different act-descriptions.

19. Nor do I take these writers to question what I take to be an uncontroversial view, viz., that to give a new, true description is tantamount to holding a new belief or to become clear about one which was held only tacitly. See the distinction between beliefs and representations in the next chapter.

believes is a true and (as) complete (as possible) description of the case. If he realizes that the act or situation can be described in another way, then he must add this new description to his original one, and then base his decision on this new set of descriptions. This will be the rational thing to do, since if there is another description which is true, this implies that the original true description was incomplete.[20] The moral tiro could hardly be said to be rational if his reason for prescribing some act was based on a description, if he knew all the time that on another description which he considered to be true and complete he would not be ready to prescribe the same act.

Consider, for instance, the example which Milkman thinks demonstrates "Hare's inability to produce an action-guiding system" (Milkman 1982, p. 23): A student is offered by a friend a copy of the exam which he is about to take. The friend says he would be terribly hurt if the student did not make use of the exam:

> Our student may consider two construals of the situation which lead him to consider two particular moral judgements: (a) – I, in this situation, being the kind of person I am, ought not insult my friend; and (b) – I, in this situation, being the kind of person I am, ought not cheat on examinations. These judgements yield minimal universal forms: (a) – Anyone relevantly similar to me in relevantly similar situations ought not insult their friends; and (b) – Anyone relevantly similar to me in relevantly similar situations ought not cheat on examinations. . . .We have here a case, however, in which two true descriptions yield morally incompatible prescriptions, according to the dictates of universal prescriptivism). If our student cheats on the examination, what stops him from justifying his action as a case of acting well toward friends, which is a true (incomplete) description of the situation (if he cheats)? (Milkman 1982, pp. 23–24).[21]

Let us call the two different construal, which Milkman speaks of description *D(a)* and *D(b)* respectively. Assume further that both *D(a)* and *D(b)* are purely descriptive, since otherwise we would no longer, according to Hare, be on the critical level of moral reasoning. Let us therefore grant Milkman that words such as 'insult' and 'cheat', for instance, are used here in a purely descriptive fashion.

According to Milkman, if the student considers *D(a)*, he thinks that in

20. Perhaps it is this assumption which is denied by these writers. However, if this is the case it is at least not clearly stated.

21. Notice that Milkman at the end includes in the description the expression "acting *well* towards friends". On Hare's account "well" would make the description carry prescriptive meaning in addition to whatever descriptive meaning it had.

situation *S* he ought not to insult his friend; and if he considers *D(b)* he thinks that he ought not in *S* cheat on examinations. The alleged dilemma of choosing which moral judgement to endorse arises because the student believes that *D(a)* and *D(b)* are both true descriptions. Unfortunately, Milkman does not tell us precisely what is the content of *D(a)* or *D(b)*. However, presumably Milkman thinks that *D(a)* mentions at least one property of the situation, say, that the friend will be hurt, which is not mentioned by *D(b)*, which in its turn mentions one property that is not included in *D(a)*, say, that he can cheat. But once this is clear, then it will also be clear that *given what the student believes* are the facts of the situation (or what is a true description of the situation), *D(a)* and *D(b)* are both *incomplete* descriptions of the case. Now, it is reasonable to say that whatever the student decides, it will be on the basis of what he believes are the facts of the case. Unless he is insincere or irrational, he cannot just put some beliefs about what are the facts of the case aside.

The situation which Milkman describes is that of a student who is offered an exam by a friend, who tells him he will be hurt unless the student uses the exam. Given this and the further fact that the student believes that by using the exam he will be cheating, what the student must ask himself is not whether he ought to cheat on the exam or whether he ought to insult his friend but whether he ought to cheat in order not to insult a friend. This will not just be another, third description of the case but one which is more complete than *D(a)* or *D(b)*.

Ultimately what eventually prevents the student from justifying his act of cheating in a certain way, according to Hare, is the combination of universalizability, prescriptivity, and the facts of the case. Assuming that the student is prescribing sincerely, he must logically meet the requirement that he too should want in an exactly similar situation to be treated as he is ready to treat the friend.[22] To decide this rationally involves knowing, according to Hare, what it is like both to be the friend and to be the person(s) who are cheated. My conclusion is that Milkman has not showed that the above case constitutes any problem for universal prescriptivism.

Hare's characterization of what I called the moral tiro has been

22. Actually the example is more complicated than I present it here. Being not a bilateral but a multilateral situation, what the student must want is what an act-utilitarian would prefer in the situation. However, since my point here is to stress the feature of prescriptivity, there is at this point no need to complicate the matter. In chapter V, I will return to multilateral cases that I do think pose a serious problem for Hare.

questioned from a different angle. There is, Hare thinks, an analogy involved between the moral tiro's procedure of testing whether or not he will be ready to accept some universal judgement, and how a chessplayer proceeds before deciding what to do:

> This procedure is no more circular than that of a chessplayer who says 'Any feature of the present position *might* be relevant to my decision on the best move to make next; so I'll look first at such and such a feature: ah, yes, it means that I shall lose my queen if I take his bishop; so, since it is a good principle (in general) not to sacrifice one's queen for a bishop, that feature is relevant'. He selects the relevant features by looking for features which will bring the present position, and the moves possible in it, under principles which he can accept; and the moralist has to do the same (*MT*, p. 90).

Roger Paden (1988) has challenged to what extent there is an analogy involved here. He claims:

> There are three considerations which together determine for the chessplayer the relevance of a principle. These are the player's desire to win, 'the rules of the game' and 'winning' itself. When a player evaluates a proposed principle, it is always (given the desires of the average chessplayer) with respect to the rules of the game and ultimately to the chances that the principle will aid him in winning the game. These considerations, which are all conceptually distinct from the proposed principles, allow the chessplayer to judge proposed principles. Hare pays insufficient attention to these considerations. While the chessplayer can take all three of these considerations into account in determining the relevance of a principle, it is not clear if the moral thinker can appeal to anything similar. For example, chess has a point, a goal, which makes such pragmatic reasoning possible. The same cannot be said for moral thinking, at least at this critical stage. Presumably, the goal of moral thinking is to discover those principles that would lead us, as Hare puts it, to 'act rightly.' ([*MT*] p. 37) However, to make use of this phrase, it must be given content (Paden 1988, pp. 229–230).

To give content to such a principle would involve a violation of what Paden calls Hare's principle of Rationality, which he understands as the following restriction expressed by Hare:

> no appeals to received intuitions are allowed, because the function of critical thinking is to judge the acceptability of intuitions, and therefore it cannot without circularity invoke [moral] intuitions as premisses (*MT*, p. 131).

The passage from Paden raises mainly two questions. First, there is the matter concerning in what sense the critical thinker can be said to have a goal in the way chessplaying has a goal or a point. Secondly, there is Paden's contention that invoking such a goal would be equivalent to leaving

the critical level for the intuitive one. This would be the case, since invoking a goal would be equivalent to invoking an intuition.

I am not entirely sure what Paden means by the "goal" or "point" of chess. Most likely, what he implies is that when playing chess we have as a goal to put our opponent's king checkmate. A player who did not intend this would have missed the point of the game. But if this is what Paden claims, it is in no obvious way true. Admittedly, on most occasions when two persons decide to play chess each will have a desire to checkmate the opponent (Paden, for instance, speaks of the desires of the average chessplayer). Notwithstanding, I can think of situations when, say, a chessplayer in a tournament strives to reach a draw; or two players who intend to reach a stalemate. These players have not missed the point of the game. They are surely playing chess. They know the constitutive rules (cf. Hargrove, 1985) for how to move the pawns, bishops, etc. Moreover, they know when the game comes to an end, viz., in the event of checkmate or stalemate, or because the players decide that they have reached a draw. When I play chess with someone who never has played before, I have taken it as a habit to always try make my opponent win. By so doing, I hope, maybe naively and somewhat presumptuously, to make the beginner favourable to chess. Have I on such occasions missed the point of chess? I do not see why. That as a matter of fact nearly all chessplayers try to put their opponent's king into trouble does not necessarily say anything about the point of chess. It tells rather something about the practicians of the chess game.

Hence, Paden's claim that Hare does not pay sufficient attention to what, by analogy with the point of chessplaying, is the point is of critical thinking, rests on the unsubstantiated assumption that there actually is *a* point of chess. As far as I can see, there is no such fixed point or goal in playing chess. Therefore, Paden has not given convincing reasons for why, on this horn of the dilemma, the analogy between chessplaying and doing critical thinking does not hold.[23]

However, Paden's main objection is not that there is no actual analogy involved between chessplaying and critical thinking. Rather he seems to be claiming something to the following effect: (i) No appeal to intuitions is

23. Paden also speaks of the "rules of the game". Hare would presumably say that by analogy with the "rules" of chess, moral reasoning has to follow the logic of the terms employed (see e.g. *FR*, p. 89). Moreover, the "desire to win" could, in its turn, correctly be paralleled with the desire to make a universal prescription.

allowed on the critical level of moral thinking; (ii) Doing moral thinking requires a goal or a point (presumably the goal of discovering principles that will lead us to act rightly; (iii) An appeal to the goal for critical moral thinking will be an appeal to an intuition. But, Paden appears to conclude, since such an appeal to intuitions would violate (i), critical thinking will not rank as equal to moral thinking. Moral thinking will always (in part) be on the intuitive level.

Paden's objection, I suspect, ultimately focuses on what in the first place motivates the agent to start doing critical thinking. Hare, as we have seen, will eventually fall back on to his argument from prudence, in order to explain why we ought to use a moral language. But although I found this a reasonable reason, it will not in any obvious way give us an explanation why the person in the first place starts doing critical thinking, with regard to a concrete case.[24]

Suppose it is a reasonable analysis of the moral tiro's situation that he will actually be moved to do critical thinking because he holds some *prima facie* principle, or has some moral intuition, or at least some universal desire that would take the form of a moral prescription, were it expressed in words. As far as I can see, the existence of this *prima facie* principle will not constitute any drawback for Hare's method. Paden is right in that no appeal to received intuitions is allowed, according to Hare, on the critical level. But the intuition here would not have to be appealed to in critical thinking. It would be invoked solely in order to explain why the moral tiro wanted, say, to express a moral judgement, or to find "the right act" to do in a certain situation. But what moral judgement we should endorse or what act we should do in this specific situation, could still be answered by taking logic and facts into consideration. That the moral tiro/critical thinker eventually needs to have recourse to an intuition in order to explain why he in a certain situation does critical thinking, does not prevent him from being able to reason in accordance with facts and logic. Such reasoning can be

24. Relevant here is also Hargrove (1985), who discusses at length similarities and dissimilarities between chess and ethical decision making. However, there is no need here to go further into details. Hargrove also comments on Hare's views. His main objection focuses on what he considers as Hare's recommendation that ethical decision making should involve "conscious rule following". By this Hargrove seems to mean the decision-making process in which the agent recites some rule while trying to decide an ethical issue. In chess such reciting would only delay the decision process, and in moral thinking, it is not obvious in what way such conscious rule following improves the decision process (*op.cit.*, p. 26). Presumably Hare would say, as seems reasonable, that it helps us avoid committing logical mistakes such as prescribing different actions for situations that we consider are exactly similar in universal respects.

performed without violating what Paden called Hare's Principle of Rationality.

Now, we have seen that, according to Hare, we are rationally required to take facts[25] into consideration before issuing our prescriptions. Moreover, it was maintained by Hare that this requirement becomes stronger when we are prescribing universally. In the case of the latter kind of prescriptions, we need to know the properties of the persons who will be affected by our prescriptions. The creditor, for instance, in attempting to rationally answer 'What ought I to do to the debtor?' must rationally prefer that the answer should be applied to him, were he in the role - and with whatever universal properties - of the debtor. The creditor will therefore be rationally required, before deciding what he ought to do to the debtor, to know as much as possible about the debtor.

In *FR* as well as in *MT* Hare actually tries to show that facts about other people's preferences and their satisfaction will be relevant facts. An important difference between these two works is that in contrast to *FR*, *MT* actually contains an argument to the effect that other people's preferences are relevant facts. However, this argument will, as mentioned earlier, not occupy us until chapter V. Here it suffices to note that facts about other peoples preferences will be relevant for the prescriber in so far this knowledge will causally affect what prescriptions the prescriber is ready to accept. *FR* rested on the empirical assumption that apart from some persons - fanatics (and amoralists) that were said to be few in number - having knowledge of other people's preferences would affect what we were ready to prescribe universally.

It will be instructive for the discussion in next chapter to present a summary of how *T1* should be understood. Let me do so by outlining one way of decomposing the line of reasoning of a critical thinker, which exposes the essential steps.

4.5. Critical Thinking Decomposed

Call the critical thinker *a*, the act which he wonders whether to do *B*, and the person who he is thinking of doing *B* to *c*:

25. A word must be said here about the use of the term 'facts'. In *MT* (p. 212) Hare says that 'factual' is an "elusive notion", and that he prefers the expression "descriptive". Nothing of substance would be lost if instead of speaking of statements of facts, we spoke of descriptive statements, the truth-conditions of which are determined entirely by the meaning of the statements.

(i) *a* asks himself whether to issue the moral judgement (s) 'I, *a*, ought to do *B* to *c*.

(ii) *a* has a perfect command of the logical properties of moral terms (universalizability, prescriptivity, overridingness).

(iii) Since *a* logically must universalize his judgement (s), rationality requires that *a* asks himself whether he is ready to sincerely prescribe that *B* ought to be done to him, in the hypothetical situation in which he has *c*'s universal properties (including in particular *c*'s preferences).[26]

(iv) To know whether he is able to rationally prescribe that *B* ought to be done to himself in the hypothetical situation that he were in *c*'s place, and with whatever universal properties of *c*, *a* is rationally required to know as much as possible about *c*'s universal properties that are available and relevant.

(v) To acquire the knowledge about *c*, *a* must represent to himself what it would be like to have *B* done to him in the hypothetical situation that he had *c*'s universal properties.

According to the theory outlined in *FR* as well as in *MT*, *a*'s moral reasoning will be constrained in the event that the motivational state of *c* actually runs counter to *a*'s present one. Let me therefore confine my attention to this possibility.

(vi) Since *c*, *ex hypothesi*, does not want to have *B* done to him, representing what it would be like to be *c* will have the result that *a* cannot accept the singular prescription 'Let *B*

26. Various writers who have criticized Hare's so-called Golden rule arguments have focused nearly all their attention on the requirement of universalizability. Gilbert (1972, p. 438), for instance, is a clear example. He argues that the universalizability-requirement is trivial: "suppose one is wondering whether or not Gauguin did well or ill to leave his wife and children; in what sense is he helped by being informed that if he decides that he did well he must also decide that anyone like him who acted as he did in a sufficiently similar situation did well too?". Gilbert's conclusion is that he is not helped, since "the answer to his question is compatible with any coherent position he might take". However, Gilbert in concluding this does not pay any attention to the fact that the decision reached must be one which we are ready to *prescribe* universally.

be done to me in the hypothetical situation that I had all the universal properties of *c*.

(vii) But if he cannot accept this prescription, neither can he prescribe sincerely that *B* ought to be done to him in the hypothetical situation that he had all the universal properties of *c*.

The reason why he cannot prescribe this is that it is, as may be recalled, a sincerity condition on issuing moral prescriptive judgements that we accept, in this case, the singular prescription 'Let *B* be done to me. . .'.

(viii) And if *a* cannot sincerely prescribe that *B* ought to be done to him in the situation that he was in *c*'s role, neither can he, on pain of breaching the universalizability-requirement, sincerely prescribe his original judgement that *B* ought to be done to *c*.

Steps (i)-(viii) summarize how, according to Hare, a person's moral reasoning will be constrained if he pays attention to logic and facts. Step (vi) is crucial to this line of reasoning. It will also be to this claim that I will turn my attention in the next chapter.

As mentioned already in the introduction, there is a difference between *FR* and *MT* as regards what Hare thought could be accomplished by reasoning in accordance with *T1*. Hare drew the conclusion in *FR* that nearly all moral conflicts or disagreements could be solved by pointing out that some party had committed a logical and/or factual mistake.

There was, however, the possibility that one of the opponents was what Hare calls a fanatic.[27] Such a fanatic would be a person who was ready to universally prescribe some course of action, although it would have disastrous consequences for himself were he to be in the place of the victims of his acts. Without committing any logical error, or being mistaken about facts, such a person could still be able to disagree with non-fanatics about some moral issue. This possibility gave rise to much criticism that to a

27. A problem with the account of fanaticism in *FR*, which, moreover, has been noticed by various writers (e.g. Donagan 1965, p. 9; cf 1977, pp. 217ff.) is that Hare would in effect have to call people fanatics that we would not, I venture, normally call fanatics. To discuss this and other related questions would unfortunately soon lead to a digression. I have therefore chosen to leave it out.

great part centred on the idea that Hare's theory seemingly paves the way for a legitimate entry into moral thinking of malevolent preferences.[28] By stressing that moral terms do not carry any descriptive meaning that might rule out on logical grounds such malevolent preferences as we associate with a fanatic, Hare must include among moral judgements those of a fanatic who is ready to prescribe universally some action such as, for instance, racial discrimination.

Hare's way of dealing with the problem of fanaticism in *FR* consists in accepting as a fact that supposing that the desire[29] of our opponent, a Nazi for instance, "were sufficiently eccentric they might lead him to hold eccentric moral opinions against which argument would be impossible" (*FR*, p. 173).

In *MT* this possibility is no longer present. The only kind of moral disagreement to be found between persons who are sincerely trying to do critical thinking, will derive from logical mistakes or because there is factual disagreement. The force of the argument in *MT* is reached by a fresh view on what follows from imagining or representing oneself in the position of someone else. By an ingenious combination of this view with the universalizability- and prescriptivity-theses, Hare hopes to show how paying attention to logic and facts will help us solve our moral problems and disagreements.

---

28. The list of writers who find reason for criticizing Hare in relation to what he says about the fanatic can be made very long. See for instance, Donagan (1965), Kerner (1965), Reed (1969, 1979), Curtler (1971), Trigg (1973), Fullinwider (1977), Gettner (1977), Norton (1977), Fotion (1983), Gorr (1985), Dahl (1987), Wetterström (1989), and Sandøe (1989a). Hare replies to the two latter writers in Hare 1989f, 1989e. See also Narveson (1978) and Shadish (1982), who each attempt to find a way to exclude the fanatic from the moral realm without damaging Hare's theory.

29. In *FR* Hare discusses the case of the fanatic to a great extent in terms of "ideals". However, at the end of his discussion (p. 170) he prefers to speak of desires which have that in common with ideals that they are "dispositions to action". Since in *MT* he does not make use of 'ideals', I have decided to leave out any discussion of what he means by ideals. See also *HC*, p. 233.

V. Conditional Representation

5.0 Introduction

Many an obstacle must be overcome before a person can be said to have performed his moral reasoning on the critical level. He must ascertain the facts of the situation, disregard his moral intuitions, and exercise whatever imaginative power he possesses. All of this, as Hare is very well aware (cf. *HC*, p. 216), is bound up with difficulties. In the preceding chapters we considered at some length Hare's views on the logical properties of moral terms. Moreover, I outlined in chapter IV Hare's views on rationality and on relevance, and gave a preliminary statement of how to understand *T1*.

In this closing chapter, I will focus on a further central thesis, viz., what I will call Hare's *Principle of Conditional Preference Representation*[1] (the *CPR*-principle, for short), which concerns the process of representation which, according to Hare, a rational moral agent must perform. By combining the *CPR*-principle with his theses about the logical properties of moral words, he sets out in *MT* to show that a rational person will reach impartial conclusions that are the same as those of a preference-utilitarian.

Hare's account of what I call the *CPR*-principle requires some unravelling, which I will be concerned with in section 5.1. Related to Hare's claim that we should take facts into consideration - especially facts about the motivational state of those persons who we believe are likely to be affected by our actions - are some complicated problems that concern intersubjective comparisons. There is to begin with the classical problem of in what sense we can know that there are other beings with the same kind

1. This idea has attracted attention from various writers who have given it different names. Gibbbard (1988), for instance, who originally gave it a name, speaks of the Conditional Reflection Principle, by which he refers to the quoted passage from *MT*. Persson (1989, p. 161) refers to Hare's "principle of Hypothetical Self-Endorsement", and, finally, Rabinowicz (1989, p. 148) ascribes to Hare the following thesis "Knowledge of preferences entails preferences". See also Hare's reply to Rabinowicz, in which he mentions that these writers are concerned with the same idea (Hare 1989d, p. 154). Also of interest here is Hare 1976a, which can be said to contain an early version of the argument in *MT*, in which the idea behind the important *CPR*-principle is not formulated. The need for such a principle was pointed out by Reed (1979), who speaks of a "Desire-appropriation principle". Relevant here is also Locke 1981, pp. 538–539.

of mental set-up as ourselves. Hare's way of dealing with epistemological scepticism will be made the topic of section 5.2. Here I will also consider the somewhat less fundamental but nonetheless difficult question of whether we can acquire any justified belief of other people's preferences. In addition, I present Hare's reply to an objection put forward notably by C. C. W. Taylor, who maintains that the process of imagination required by Hare involves not only practical but also logical problems. In 5.3 I turn to Hare's views on the meaning of the word 'I', which plays an important role in the understanding of Hare's views on representation.

Hare's account of critical thinking has been interpreted in various ways. In section 5.4 I consider some recent interpretations that deviate from my own. Moreover, I devote sections 5.5–5.7 to a discussion of three restrictions which Hare introduces in *MT* with regard to the kind of desires and preferences which a critical moral thinker should pay attention to. The ensuing sections will all be concerned with objections that can be directed against the *CPR*-principle. This principle has been the subject of dicussion and criticism ranging under as different headings as logic, epistemology, and metaphysics. I do not claim to cover all the problematic sides of Hare's idea. However, I hope to show that there are reasons why the principle should be questioned. Thus, 5.8 is devoted to a discussion between Hare and some of his critics, concerning whether critical thinking will, as Hare claims, lead in multilateral cases to conclusions that correspond to the ones of an act-utilitarian. As will become clear, I part company with Hare on this issue. In section 5.9–5.9.1 I look more closely into his claim that the *CPR*-principle is a conceptual truth. My main objection here will be that the notion of prescriptivity on which the *CPR*-principle rests is clearly inconsistent with how he has used 'prescriptive term' in earlier works.

5.1 Representation

The general idea of the *CPR*-principle, namely, that knowledge of one's own desires entails desires, is central to Hare's claims that (i) interpersonal conflicts of preferences will in critical thinking be transformed to an intrapersonal conflict of preferences, and by extension that (ii) the conclusions of a perfectly conducted critical thinking will be impartial, and (iii) correspond to the conclusions of a utilitarian. In the following passage Hare offers a general statement of the idea:

> I cannot know the extent and quality of others' sufferings and, in general,

> motivations and preferences without having equal motivations with regard to what should happen to me, were I in their places, with their motivations and preferences (*MT*, p. 99).

The precise nature of the *CPR*-principle is brought out by his comment on the relation which he thinks holds between the following two sentences:

> (1) I now prefer with strength *S* that if I were in that situation x should happen rather than not;
>
> (2) If I were in that situation, I would prefer with strength *S* that x should happen rather than not (*MT* p. 95).

Furthermore, he says:

> I cannot know that (2), and what that would be like, without (1) being true, and that this is a conceptual truth, in the sense of 'know' that moral thinking demands (*MT*, p. 96. Henceforth I will refer to the above two statements as '1' and '2' respectively).

Hare, in other words, treats it as a logically necessary condition for representation that we generate what in the litterature has become known as a "conditional preference" (cf. Gibbard 1988), which is a preference which we actually have now for a hypothetical situation in which we figure. It is to such a preference that (1) above refers. Conditional preferences must not be confused with so-called hypothetical desires or preferences. Sentence (2) above refers, in its turn, to such a hypothetical desire. Such desires are desires that we predict we will have *if* some conditions obtain. I am convinced, for instance, that were my family in danger, I would to begin with act so as to satisfy my desire that they and not myself would be saved. However, were I to be in such a situation it might very well as a matter of fact be that my desire to save my own skin would outweigh my desire to save the family.

Given the truth of the *CPR*-principle, it can be shown, Hare maintains, that interpersonal preference-conflicts will turn into intra-personal. To see this it will be wise to consider a fairly simple bilateral example of a preference conflict, which shows affinity with an example discussed by Hare in *MT*.

A motorist, call him *a*, wants to park his car in a location which is already occupied by a person *b*'s bicycle, who *a* believes does not want to have his vehicle moved. We assume that *a* is rational (complying with *R2*), has a perfect command of the meaning of 'ought', and that he wonders whether he is ready to prescribe 'I ought to remove the bicycle and park my

motorcar'. Since *a* has a perfect knowledge of the meaning of 'ought' he knows that if he assents to this judgement, he must be ready to assent to its universalization. He must in other words be ready to prescribe that if he were in *b*'s position, with *b*'s motivational set-up, he too ought to have his bicycle removed in exactly similar circumstances (with regard to universal properties). Moreover, since this judgement will be prescriptive, he must, if sincere, actually desire, *in propria persona*, that the bicycle should be moved.

Now rationality requires of *a* that he takes into consideration relevant available facts. Given the truth of the *CPR*-principle, facts about *b*'s motivational state will be relevant, which can be seen by the following line of reasoning.

Suppose *a* has knowledge of what it will be like to be *b*, with *b*'s motivational set-up. In such a case, he will have acquired a desire identical in strength and content to that of the bicycle owner, for the hypothetical situation that he was exactly similar to the bicycle owner (with regard to universal properties). Thus, having put himself in the shoes of the bicycle owner, the car owner will have, the argument goes, two preferences: The original one that the bicycle should be moved, and the one generated by conditional representation that the bicycle should not be moved. Moreover, the strongest preference is, if I have understood Hare correctly, by definition the one that wins (cf. *MT*, p. 133).[2] Thus, supposing that it is his conditional desire which is the strongest, then the rational car owner will act in such a way as not to frustrate this latter preference.

It can now also be seen what Hare means by the claim that

> the method of critical thinking which is imposed on us by the logical properties of the moral concepts requires us to pay attention to the satisfaction of the preferences of people (*MT*, p. 91).

Facts about other people's preferences are relevant, since a knowledge of what it is like to be them, will give rise to conditional preferences, which in turn will affect what I, all things considered, will be ready to prescribe and desire (*MT*, p. 215).

2. In contrast to Michael Gorr, who says that Hare "seems to presuppose throughout his writings a Bayesian account of rational behaviour according to which a rational agent always strives to maximize his own expected utility" (Gorr 1988, p. 119), I do not see the need to ascribe such an account to Hare. All that Hare seems to say and needs to argue is that the preference which will result in a decision (to act) is analytically the strongest preference. For a view of strength in a desire in terms of vividness that complements 'the winner is the stronger' view, see Wiggins 1978–79.

Hence, by combining universalizability and prescriptivity with the *CPR*-principle,[3] Hare has launched an argument that aims at showing that impartiality is secured if we think critically about moral matters. By conditional representation, preferences will be generated that correspond in strength and content to the preferences of the people affected by our actions. Once acquired, these conditional preferences will be treated on a par with the agent's original preferences, i.e., in accordance with strength (e.g. *MT*, p. 180). Thus, impartiality is reached, and the moral conclusion arrived at will correspond, Hare claims, to that of a preference act-utilitarian. Henceforth I will allude to this version of utilitarianism simply as utilitarianism.

At this point it is important to steer clear of a possible and fatal misinterpretation, which concerns Hare's view on the very process of imagination or representation. Hare is careful to underline that he is not maintaining that the process of representation is somehow informative.[4] Thus, he says:

> It would be wrong to claim that our imaginations somehow *inform* us of what the experiences of others are like. Imagination is a very common source of error; it can just as well be of experiences and preferences which they do not have as of those which they have. But if we do know what it is like to be the other person in that situation, we shall be (correctly) imagining having those experiences and preferences, in the sense of knowing or representing to ourselves what it would be like to have them; and this, I have been claiming, involves having equal motivations with regard to possible similar situations, were we in them (*MT*, p. 95).

The knowledge we have of other persons, of their minds and experiences must be acquired by other sources. Let me therefore, at this point, turn

---

3. Paden (1988) mistakenly interprets Hare as founding his derivation of utilitarianism solely on the universalizability thesis and prescriptivity-thesis. Williams (1985, p. 85) is another case: "[Hare] believes his argument for utilitarianism to follow strictly from the meaning of moral words" (However, cf. p. 89 where Williams discusses the relation between Hare's sentences (1) and (2).) Robinson (1982) also bases his criticism of Hare's derivation of utilitarianism on the same interpretation. However, his claim concerns the position in Hare 1976a, which he thinks has led Hare to abandon the prescriptivist position of *FR* and lapse into naturalism (cf. Hare's own comments in Hare 1989e, p. 205). As I understand Hare, what he wants to show is that it is the combination of these two theses with what I have called the *CPR*-principle that at the level of critical thinking has a result similar to that of an act-utilitarian.

4. Not all writers have interpreted Hare in this way. Bergström (1982, p. 303), for instance, objects to Hare's use of imagination that "We have no guarantee that our imagination is veridical". As outlined, Hare would not disagree: The purpose of the process of imagination is not to give us knowledge about what the other persons' preferences are. This knowledge must be acquired from other sourses.

briefly to the epistemological discussion which Hare at some length brings up in *MT*.

5.2 Interpersonal Comparison Problems

Under the heading "Interpersonal Comparison" Hare discusses in *MT* some epistemological as well as logical problems that his theory appears to face. With regard to the former problems, he carefully points out that they are not problems that arise within moral epistemology. Rather, they are problems that are part of general epistemology. I have earlier commented briefly on his scepticism about there being any specific moral epistemological questions that are not ultimately reduceable to questions of language, and I will not continue this discussion here. The problems in question concern all our knowledge about our own past and future experiences as well as those of other people.

For a start Hare addresses the sceptic's classical question of whether we can know that there are other beings with the same conscious experiences as we have ourselves. "How do I know there are", he asks, "any conscious experiences but my own and, if so, what they are?" (*MT*, p. 118).

Hare does not purport to have a solution to the sceptic's problem. Rather, he thinks

> we shall be driven back to some form of the old 'argument from analogy'; for, after all, it is tempting to say that we reasonably guess that beings so like us in all other respects are also like us in having similar conscious experiences under similar conditions. I shall not discuss the question further, but shall merely take it that its solution is so crucial to almost all parts of philosophy that it will not be held against my own theory in particular if I just *assume* that it can be solved (*MT*, p. 119).

Since I am in no better position than Hare with regard to this matter, I will not hold it against his method that we are in no possession of a stronger argument than what this analogy can accomplish.

Supposing that scepticism about other minds can be disregarded, Hare maintains further that

> we may perhaps also assume, as part of the solution [of the problem of other minds], and of other unsolved problems raised by philosophical sceptics, that we are entitled sometimes to be confident in representing to ourselves

> the states of mind, including the preferences, of ourselves in the past, ourselves in the future, and other people (*MT*, p. 126–127).

With regard to our past preferences, Hare claims, our memory sees to it that we can make confident representations. In the second case – our future preferences – memory will also play a role. By combining memory and induction we manage to represent experiences that we predict we will have in the future. It will be done by assuming that

> future experiences will be like those we remember having under similar conditions. Or (and this brings us to the third case) we may predict our own experiences on the basis of what others tell us about theirs. But more commonly and more basically we represent to ourselves the experiences of others by analogy with our own, judging that their situations are similar . . ., their physiologies are similar . . ., and that therefore their experiences will be similar (*MT*, p. 127).

Hare's approach to the problem posed by scepticism, seems reasonable from the point of common sense. Whether common sense rather than scepticism eventually is the correct philosophical position to take, is too encompassing a question for me to try to address here.

A further issue that falls under the heading of interpersonal comparison problems deserves to be mentioned here. It concerns the various problems that the process of representation gives rise to, with regard to the issue of personal identity. Can we imagine ourselves having a new set of universal properties? Are there logical obstacles such that there are limits to what I can imagine myself be? Can I, for instance, in any meaningful way be said to imagine myself being in Napoleon's place, with Napoleon's beliefs and desires? Could I, being a man, imagine myself being a woman? Could I, as a member of a certain animal species, imagine that I belonged to another species?

Some writers have maintained that the process of imagination required by Hare's theory involves a logical and not merely, as Hare recognizes, a practical problem.

Locke (1968), for instance, maintains:

> True, 'Wives ought to honor and obey their husbands' would apply to me *if* I were a wife. But in so far as this condition is not one which I could satisfy, this principle is not one which can be applied to me (*in propria persona*), even in a hypothetical situation (Locke 1968, p. 41).

Locke thinks there is a difference here between the principle mentioned above and the following one: "Englishmen ought not to kiss one another

when they score goals". This principle, he says, does not apply to him. He claims, however, "it *could* be applied to me; it would apply to me *if* I were English". The difference which Locke finds here appears to derive from the fact that Locke could not even in a hypothetical situation become a wife, whereas he could become English in a hypothethical situation. Another example is Harry S. Silverstein, who maintains that Hare's universalizability requirement applies also to cases involving our treatment of animals. He states that he has

> considerable sympathy for the contention that one cannot imagine oneself to be an animal (Silverstein 1974, p. 67).

To imagine oneself playing the role of an animal seems, he thinks, as unintelligible as imagining

> oneself to play the role of any person whose background, interests, attitudes, and desires are very different from one's own (Silverstein 1974, p. 67).

It is not entirely clear to me whether Silverstein actually is arguing that there is a logical problem involved here. However, his choice of the term "unintelligible", suggests that he does not conceive of this as merely a practical problem.[5]

C. C. W. Taylor (1965) contains the clearest argument to the effect that Hare's theory requires something of us that we cannot for logical reasons comply with. He discusses, for instance, what is involved in imagining that one is Napoleon. There are, he maintains, three different imaginative states that should not be confused:

> (a) imagining that one is some person other than the person one in fact is, (b) 'putting oneself in someone else's place', *i.e.* imagining what one would oneself feel were one in the same situation as someone else is in, or in a similar situation, and (c) imagining what someone else feels in a particular situation (Taylor 1965, p. 288).

Could there be a situation in which, say, Smith was Napoleon? Since there for logical reasons, Taylor claims, "is nothing which would count as that situation's obtaining, the situation cannot be imagined either, since to imagine it is to imagine it as obtaining" (*op. cit.*, p. 289). Moreover, as

5. In another paper, Silverstein (1972) describes a case that indeed poses a problem for Hare's theory. I will return to this matter when discussing Hare's views on so-called external preferences. See here also Michael H. Robinson (1975) and Whiteley (1966).

Taylor correctly points out, Hare's theory requires more than what is outlined in (b) or (c). He cannot therefore meet the above objection by recurring to (b) and (c), and Hare's "Golden rule argument" cannot for this reason be brought into play.[6]

Hare outlines a reply in both *MT* and *HC* to the above objection, which points at a way of explaining how we could without any logical offence be meaningfully said to imagine having the properties of someone else. According to Hare, this explanation is, moreover, consistent with it being, as Taylor is driving at, self-contradictory to suppose Smith had all of Napoleon's personal characteristics and preferences. Hare's reply centres on a difference between how proper names refer and how a personal pronoun such as 'I' refers. Consider the following passage:

> But can there be any difference between Jones imagining Jones being in Smith's precise situation, and Jones imagining being in Smith's precise situation; or between Jones saying 'If Jones were in Smith's precise situation, he would be suffering horribly', and his saying 'If I were in Smith's precise situation, I should be suffering horribly'? (*MT*, p. 120).

Hare maintains that there is a difference. For Jones to say 'If Jones were in Smith's precise situation . . .' could very well involve him in a self-contradiction, whereas this would not be the case were he to say 'If I were in Smith's precise situation . . .'. Employing 'Jones', Hare says

> involves the ascription to Jones of some properties (not necessarily the same properties for all those who use them to identify Jones); but the bringing into use of the expression 'I' involves the ascription of no essential properties beyond that of being a person (and perhaps not even that, for one could put on one's stove as a warning 'I am hot' and the 'I' be readily understood as referring to the stove) (*MT*, p. 120).

---

6. Cf. Lycan (1985) who sees a difficulty with Hare's Golden Rule argument, when it is directed to the abortion controversy: "The trouble, as will be forseen, is with the hypothesis 'Suppose I were a fetus . . .' If we think of fetuses as little people, the former supposition makes fairly good sense; but if we think of fetuses as warts or protoplasmic eruptions only, the supposition makes no more sense than 'Suppose I were a wart'"(*op. cit.*, p. 149). I do not think that much is solved by regarding fetuses as persons, but I agree with Lycan that the abortion issue raise questions concerning to what extent Hare's Golden Rule argument is applicable to a fetus or a "little person". For one thing, the fetus cannot reasonably be said to have desires other than potential ones. It should be noticed, as Lycan also sees, that Hare (1975) deals with the abortion problem by applying a more traditional Golden Rule argument. For a criticism, see Werner (1976). Relevant here is the qualification which is made in *MT* to the effect that Hare confines attention only to beings "who have self-consciousness" (*MT*, pp. 92–93).

Hare maintains, in other words, that although it may be correct that it is logically impossible for Jones to imagine Jones, *qua* a certain set of properties (viz., that set to which 'Jones' refers), being in Smith's situation, this is not the case when 'I' is involved. There are no essential properties brought into use when using 'I'. The reason for this is the linguistic fact, as Hare puts it in *HC*, that 'I' is an indexical term. That is, two persons could use 'I', he maintains, in order to refer to themselves, and in so doing be using 'I' synonymously.[7] However, they would be using 'I' with different references.

In what follows, I will grant Hare this linguistic claim, viz., that 'I' is an indexical expression, and that it therefore makes sense to say, for instance, 'If I had all the universal or qualitative properties of Napoleon . . .'. However, I do so reluctantly, with the distinct sense of leaving some interesting questions unanswered (but not, *nota bene*, because I adhere to a non-reductionist view on personal identity). Is Hare's reply, for instance, committing him to non-reductionism, with regard to personal identity? As far as I can see it does not. Hare himself denies it. He makes it clear, for instance, in his reply to Vendler (1988) what he does not believe in with regard to "the self":

> I do not want to posit the sort of self that Hume rejected, any more than the sort that Descartes believed in (*HC*, p. 282).[8]

Regrettably, I have no firm belief in what personal identity consists in. This may account for my readiness to accept as a logical possibility that one can indeed imagine oneself as having an entirely different set of universal properties, which is what critical thinking requires. The scope and complexity of the recent discussions of this topic have at least convinced me

---

7. Thus, he says "So when Jones and Smith both use the expression 'I' to refer to themselves, they use it synonymously, though with different references" (*MT*, pp. 120–121). Hare calls 'I' an indexical of self-reference in *HC*. Most relevant for this discussion is Vendler (1988), whose argument Hare claims he is following.

8. What view Hare endorses with regard to the complicated problem of personal identity is unclear to me. Suppose he actually is a reductionist with regard to the "self", as the reply may suggest. An interesting issue to bring up would be how such a view squares with the fact that Hare's theory appeals to what persons are ready to endorse *in propria persona*. Consider here also the continuation of the passage quoted: "But I do think that the other I is distinct from this I; for I care whether this I suffers, but may not care whether that I suffers. That is, I accept different *prescriptions* about the two Is" (*HC*, p. 282). Given a denial of a Cartesian "self", it would have been desirable to know what the rational basis for such a care is.

that I had better postpone a discussion of this to a future work. There are, as I hope to show, enough problems as it is with Hare's *CPR*-principle.

5.3 The Prescriptivity of 'I'

Let us now look further into Hare's claim about what follows from representation of one's own (hypothetical) desires.[9] We saw that representing, correctly or not, what it will be like to be in, e. g., another person's situation, with this person's preferences, will for conceptual reasons lead, according to Hare, to the acquisition of a conditional preference. Following Ingmar Persson I will speak rather of "representations" than beliefs, when discussing conditional representation. One may have several beliefs that are dormant, i.e. "possessed even at times when one fails to call them to mind" (Persson 1983, p. 45). The process involved in conditional representation strongly suggests that the beliefs involved are not dormant ones.

Thus the line of reasoning endorsed by Hare appears to be the following: (i) Rationality requires that before I decide whether or not I ought to do an act *x* to a person *a*, I must take the relevant available facts into consideration. Included among these relevant facts are facts about what it will be like for *a* to have *x* done to him. (ii) To know what it will be like for *a* to have *x* done to him, I must represent to myself what it would be like to be in *a*'s situation. (iii) Representing to myself what it would be like to be in *a*'s situation generates in virtue of the CPR-principle that I acquire a conditional preference.[10]

Suppose I know what the universal properties of a person *a* are at a given time *t*. Knowing this, I will among other things know what *a* prefers at *t*, say, that he with strength *s* desires to listen to Verdi's opera *Rigoletto*. Moreover, I will not know this unless I represent to myself what it is like to be in *a*'s situation with a desire to listen to Verdi's *Rigoletto*. Given that I represent myself in *a*'s position, it will for logical reasons be the case that - as (1) formulates it (p. 179 above) - I now prefer with strength *s* that if I were in

---

9. I am indebted to Ingmar Persson for his many valuable comments on how to interpret Hare's argument.

10. Consider for instance the following passage: "But if we do know what it is like to be the other person in that situation, we shall be (correctly) imagining having those experiences and preferences, in the sense of knowing or representing to ourselves what it would be like to have them; and this, I have been claiming, involves having equal motivations with regard to possible similar situations, were we in them" (*MT*, p. 95).

a's situation *Rigoletto* should be played rather than not. I may in fact utterly dislike opera music. Nevertheless, if I have represented myself in the above situation, I will, given the conceptual truth of the *CPR*-principle, actually have a desire to hear *Rigoletto*, were I in *a*'s situation.

Now, imagining what it would be like to have the universal properties of someone else enables us to compare or weigh the strength of someone else's preferences with our own preferences. Unless I have got Hare all wrong, representing oneself being in this state is not entirely a matter of being in a cognitive state (cf. *HC*, p. 216). Hare explicitly repudiates the idea that what he is doing is going from facts about other people's preferences to prescriptions (preferences) held by the speaker himself (see e.g. *MT*, p. 220). He is, in other words, not giving up his earlier idea that there is no inference from facts to prescriptions.

The generation of the conditional preference rests rather on a different and novel idea of Hare's, viz., that the term 'I' is not (at least not entirely) a descriptive word but carries a prescriptive element in its meaning. Due to this prescriptive meaning, (2) above will not exhaustively be descriptive, and representing oneself being in some situation will not completely be a matter of being in a cognitive state. There will in addition be a conative state involved that derives from the prescriptivity of 'I'.[11] Hare elucidates his idea in the following passage:

> The suggestion is that 'I' is not wholly a descriptive word but in part prescriptive. In identifying myself with some person either actually or hypothetically, I identify with his prescriptions. In plainer terms, to think of the person who is about to go to the dentist as myself is to have now the preference that he should not suffer as I believe he is going to suffer. In so far as I think it will be myself, I now have in anticipation the same aversion as I think he will have (*MT*, pp. 96, 97).

Hare does not deny in *MT* that 'I' generally is used with descriptive meaning. It is likely, he maintains, that 'I' attaches to certain properties such as having a brain, memories, and personalities in varying degrees (see *MT*, p. 98; cf. p. 123). What he claims is that to this descriptive sense of 'I', we must add a further, prescriptive sense:

> As applied in ordinary cases, it is probable that the word 'I' is attached to all the identifying properties above-mentioned in varying degrees. I wish merely to suggest that to these we should add, not another identifying property, but another feature of the word 'I', namely that, by calling some

11. Cf. Millican 1983.

> person 'I', I express at least a considerably greater concern for the satisfaction of his preferences than for those of people whom I do not so designate (*MT*, p. 98).

It is reasonable to understand the concern which Hare speaks of here in terms of having desires that the desires oneself would have, were one in the position of the other, should be satisfied. This idea is supported when he returns once more to this topic in the final chapter of *MT*:

> If it be accepted that 'I' has a prescriptive element in its meaning, then, when I entertain the thought that the person who is actually now suffering might, in a hypothetical case, be myself, I acquire some hypothetical prescriptions by doing this. In general, when I say that somebody who would be in a certain situation would be *myself*, in so saying I express a concern for that person in that hypothetical situation which is normally greater than I feel for *other people* in the same situation. *To recognize that that person would be myself is already to be prescribing that, other things being equal, the preferences and prescriptions of that person should be satisfied.* This is what is involved in 'identifying' with that person (*MT*, p. 221. The latter italics are mine. Cf. Hare 1989e, p. 206).

This passage supports the following formulation of the *CPR*-principle:

CPR-principle: To recognize or imagine that a person *a* would be myself is already to be prescribing that, other things being equal, the preferences and prescriptions of that person should be satisfied.[12]

The argument founded on the *CPR*-principle is complicated, and it will be wise to repeat it once more. Moreover, it will be convenient this time to introduce Hare's idea about the archangel. Such a being is presented in *MT* in order to bring out what critical moral thinking can result in when it is performed optimally (see e.g. *HC*, p. 241 n. 89/16). As may be recalled, this archangel is a non-human being

> with superhuman powers of thought, superhuman knowledge and no human weaknesses (*MT*, p. 44).

Such an archangel will be able, when presented with a novel situation,

12. An alternative formulation would be to say: 'To entertain the thought that a person *a* might, in a hypothetical situation, be myself, is already to be prescribing that, other things being equal, the preferences and prescriptions of that person should be satisfied'.

> at once to scan all its properties, including the consequences of alternative actions, and frame a universal principle (perhaps a highly specific one) which he can accept for action in that situation, no matter what role he himself were to occupy in it (*loc. cit.*).

Being an angel of a utilitarian cut,[13], possessing exhaustive knowledge about the complete consequences of any act, is, of course, a highly desirable trait. It is not clear to me, though, to what extent such a being constitutes a coherent idea.[14] However, since I will focus on the archangels' capabilities to gain full knowledge of a singular individual at a certain time, something which does seem to be possible for such a superbeing, a discussion concerning the details of this omniscient being would be a digression (for a criticism of Hare's use of the idea of the archangel, see Vendler 1988 in *HC* and Hare's comments).[15]

Now suppose that the archangel wonders whether to do a certain act A that he knows will affect a certain person *b*. We may even assume that the archangel has an original preference that is in opposition to *b*'s preference. To comply with the logical universalizability-requirement, the archangel must imagine what it would be like to be in *b*'s place. Rationally imagining this means imagining among other things having *b*'s motivational set-up. Being superhuman, the archangel will have no difficulties in finding out that *b* prefers that A should not be done to him. By conditional representation, the archangel now acquires the conditional preference that if he were in *b*'s place, A should not be done.

---

13. Will not all archangels be utilitarians? Since the devil is said to have a past as an angel, we may perhaps conclude that there is at least one kind of angel who will not be a utilitarian.

14. Assume that there is no end to the chain of consequences. Is it then, for instance, within even an archangel's reach to pass a judgement at a given time, which is supposed to be founded on knowledge that concerns what will never have an end? Some will probably see no problem here. However, I do find this difficult to swallow. We might, of course, consider the chain to be finite or believe that the effects of an act, like ripples on water, gradually vanish, but I do not see why these alternatives are more reasonable than the former one.

15. See also McGray (1986–87, p. 86, note 8) who claims that "One would misinterpret Hare's regular appeal to archangelic knowledge and rationality if one makes the archangel the ultimate moral judge. The archangel has no personal stake in the dispute. It is a completed detached observer". Cf. McDermott (1983), who also argues that the archangel is a detached observer who lacks reasons for making moral judgements. There is no conclusive textual evidence that these interpretations are correct. See the discussion between Hare and Gibbard in *HC* about archangels with idiosyncratic preferences.

An interpersonal conflict has, on this account, turned into an intrapersonal one. The archangel will now be in a position to express a judgement that expresses the archangel's (or in general, the critical thinker's) "preference all in all" (*MT*, e.g. p. 227). This "preference all in all" is the resultant after the archangel has weighed his original preference against his acquired, conditional preference.[16]

Here it is important to pause a moment and ask just what this weighing of preferences consists in. Part of the problem of how to understand the argument which Hare founds on the *CPR*-principle is tightly bound up with the problem of understanding Hare's view of desires and their role in (moral) deliberation. As outlined in chapters I and II, he is far from being detailed on the former topic, which is rather surprising, considering the far-reaching conclusions he draws from our concern to satisfy the preferences of those whom we refer to with 'I'. Moreover, judging from some recent attempts to understand Hare's arguments in *MT*, there is also room for different interpretations of how he regards the role of desires in critical thinking, and particularly what the process of weighing consists in. Let us therefore begin by considering the objections which some writers have recently raised concerning these issues.

5.4 Desires and Deliberation

Pettit and Smith's recent objection to Hare's derivation of utilitarianism sets out from the example of Socrates, and his decision to take the hemlock. Suppose Socrates, they maintain, realized his friends did not want him to carry out his decision. On the face of it, the number of friends (they call them *b* and *c*), and the strength of their preferences would mean that maximizing preference satisfaction here should convince Socrates to pour out the hemlock. However, Socrates believes, *ex hypothesi*, that the option to drink the hemlock is "the only virtuous one". Pettit and Smith suggest now that Socrates, given this belief, will be unmoved by the belief that were he in the situation of his friends he would then desire that the hemlock should not be taken. In Pettit and Smith's own words:

> I will be unmoved by the effect of universalizing and seeing, say, that were I in *b*'s position or *c*'s, I would desire that the agent not drink the hemlock. I will be unmoved at least, so long as the property in view of which I opt for

16. Cf. Gibbard (1988), who speaks of an agent's preference "all told"; McDermott (1983, p. 386) uses the expression "on balance".

> **F**(*a, b, c*) [where this is an action done by me which adversely affects *b* and *c*] is that this is the virtuous option. I will say that I should discount what I would desire were I in one of the other positions because, clearly, I would not be sufficiently alert to the virtuous character of the action, were I situated there; otherwise I would continue to desire that the action be performed (Pettit and Smith 1990, p. 581).

Thus, what Pettit and Smith seem to be saying is that there is no conclusive reason why Socrates should be moved not to drink the hemlock, even if he came to believe that were he in his friend's position he would not endorse the decision to take the hemlock. Believing that their opposition is based on not realizing the virtuous character of the act, Socrates would be quite consistent to carry out his original decision.

However, Pettit and Smith's objection is not convincing. It rests on a mistaken view on Hare's theory of moral reasoning. I suspect, actually, that there is more than a single mistake involved here. McNaughton (1988), in his comment to an earlier version of this argument (Pettit 1987), correctly pointed out, "it only succeeds ... because it makes an assumption which Hare denies" (*loc. cit.*, p. 170); the assumption being that Socrates may on the critical level invoke the moral conviction that the act of taking the hemlock is a *virtuous* act. On the critical level, which is the only level from which Hare thinks utilitarianism can be derived, only non-moral facts may be invoked.

Pettit and Smith appear to commit a further misinterpretation here, viz., that of applying the universalizability-requirement in an indirect way: Socrates begins by making up his mind about whether he ought to take the hemlock or not. Only then, after having decided that taking it is what he ought to do, does he go on to consider what he would desire were he in the place of his friends. Facts about how the act under consideration will affect his friends are on this account not part of what is Socrates' reason for prescribing the taking of the hemlock. But this is clearly not how Hare conceives of what is involved in reasoning on the critical level. If Socrates' decision to drink the hemlock is to be rational, he must take into consideration all relevant, available non-moral facts. Among these are facts about how his action will affect the preferences of others (notably his two young friends). To know this, he must put himself into the shoes of those affected, and given the truth of the *CPR*-principle[17], this will have as a

17. Pettit and Smith (1990) do not have any discussion of anything to the effect of what I have called the *CPR*-principle. Pettit does, however, have a reference at the end of his 1987 paper. However, its role in Hare's argument is not considered. McNaughton (1988) does not have any account of the *CPR*-principle either.

consequence that he will actually acquire a similar, albeit conditional desire as his young friends. The case, in other words, is not, as Pettit and Smith put it, whether the consideration of what he would desire in a hypothetical situation would move him. As we have seen above, Hare maintains that it will be a conceptual truth that Socrates, if he imagines himself in the position of each of his friends, will acquire a conditional preference.[18] What Socrates will prescribe all in all is what will result after his preferences (his original and his conditional) have been balanced against each other. If the steps involved here are carried out properly, and Hare is right about the nature of the *CPR*-principle, then it seems evident that the preferences of the young friends may alter Socrates' original preference. Pettit and Smith do not, as I see it, provide us with any argument to the contrary.

Pettit and Smith also raise a question concerning Hare's view of the role of desires in moral deliberation that must be commented on.

Pettit and Smith consider two views of the role which desires may play in decision making. The *strict foreground view of desires* is contrasted with what they call a *strict background view of desires*. Whenever an agent decides to do an act, a desire will be in the foreground of the decision "if and only if the agent was moved by the belief that a justifying reason for the decision was that the option promised to satisfy that decision" (*loc. cit.*, p. 568). The strict background view contains, on the other hand, the following claim: "a desire is present in the background of an agent's decision if and only if it is part of the motivating reason for it: the rationalizing set of beliefs and desires which produce the decision" (*loc. cit*). The strict background view is arrived at by adding to this passage that there is always a desire in the background in deliberative decision making. Pettit and Smith convincingly argue that of these two views the former is false. In addition they make a claim that is not compelling, viz., that the so-called "strict foreground view of desire" is at the origin of Hare's claim that taking logic and facts into account will maximize preference satisfaction of all affected parties.

Now, the reason Pettit and Smith (1990) offer for ascribing the strict foreground thesis to Hare, is that such a view lends Hare's derivation of utilitarianism some plausibility. But since it was shown earlier that what

18. Hare makes the following remark that is relevant to this discussion: "We can utter the words 'I shall be desiring not to be being whipped'; but we shall not really be *thinking what it would be like* unless there is a desire not to be being whipped if we are in that situation" (*HC*, pp. 216–217). Contrast this with Pettit and Smith's claim that I would be unmoved by the consideration that "were I in *b*'s position or *c*'s, I would desire that the agent not drink the hemlock" (Pettit and Smith 1990, p. 581).

they mean by Hare's derivation of utilitarianism is an incomplete, and therefore erroneous, account, this claim loses all weight. Pettit and Smith fail to make it clear why Hare needs something to the effect of their "foreground view of desires". A central feature of Hare's account of critical thinking is the following idea: since the critical thinker will acquire actual, albeit conditional desires, corresponding to the ones of the persons involved in the situation under consideration, the decision - what he will prescribe all in all after having taken logic and facts into account - will express what is on balance his strongest desire. As far as I can see, there is nothing in this process that requires that the conditional preferences or the strongest preference should be in the foreground.

Hare's view on the role of desires in practical deliberation is also criticized by Pettit (1987). Pettit maintains here that Hare "is much given to a weighing metaphor in characterizing practical reasoning". He thinks that Hare

> interprets this metaphor in a way that insinuates the assumption that an agent's reason for prescribing anything must mention his desire (Pettit 1987, p. 80).

The following passage from *MT* is considered:

> Our other preferences may outweigh this one; what we prefer all in all is determined by the balance of them without external constraint (*MT*, p. 225).

Pettit now claims:

> Under one interpretation, this metaphor is to be understood mechanically; under the other, it is to be taken metrically. The mechanical idea, neither subtle nor insidious, is that the deliberative conflict is like a physical disequilibrium whose resolution depends, at least in part, on which of a number of elements weighs most. As the spinning die tends to come down on its heaviest side, the undecided agent is swayed by his weightiest or strongest desire (Pettit 1987, p. 80).

Such a mechanical view should be contrasted with the metrical interpretation, which Pettit thinks should be ascribed to Hare:

> Here, deliberative conflict is pictured on the model of a competition to determine the heaviest of a set of objects by weighing them in a balance. Desires each get their turn on the scales and victory goes to the one which notches the highest mark. It outweighs all competitors (*loc.cit*).

Pettit rightly claims that the metrical interpretation

> begs a serious issue by imposing the assumption that the resolution of desire-conflict requires an umpire. There has to be someone who sees which desire outweighs competitors; there has to be someone, indeed, to weigh the desires against one another (*loc. cit.*).

Now, Pettit has a point in that Hare in *MT* (and elsewhere too, for that matter) sometimes speaks of the weighing procedure as if it required an umpire who actually performs the weighing.[19] Notwithstanding, in what follows I will understand Hare as if he was not holding something to the effect of the "metrical view" desires. Again, my reason is that it has not been shown that Hare needs something to the effect of the "metrical view" (or for that matter, the "foreground view"). Moreover, since I agree with Pettit that the metrical view is misleading in positing an umpire where there is no need for such, it would be unreasonable to ascribe such a view to Hare.

Consider the following passage, in which Hare maintains that the resultant preference - our preference all in all - will be

> a function of our separate and perhaps conflicting preferences and their respective intensities and of nothing else (*MT*, p. 225).

Given this passage, there seems to be no need to bring in any umpire in the process of weighing, no umpire, that is, who will consider the preferences as data. Our preference all in all will be a *function* of our separate preferences - those we originally had and those we have acquired by conditional representation.

A further interpretation should be mentioned here. Some writers have either supposed or considered the possibility that the strength of preferences, which Hare speaks of, should be understood as "felt strength" or "felt intensity" (McCloskey 1979, p. 71; Griffiths 1983, pp. 503-504; Lennon 1984, p. 618; Griffin 1988, p. 79). Hare did in one place actually

---

19. Besides the term 'weighing' Hare characterizes the process other times as one of 'comparing' or 'balancing' preferences (e.g. *HC*, p. 220). However, the use of these verbs constitutes no conclusive evidence for either of the interpretations under consideration. See also Rabinowicz (1983), who discusses two interpretations of how Hare conceives of an agent's weighing desires against each other. Desires may be regarded as "driving forces" or as "data" (*op. cit.*, p. 146).

suggest in *FR* that 'desire' should be limited to "*any* felt disposition".[20] But in "Wrongness and Harm" (originally published in *MC*), he seems not to follow this line of reasoning:

> The troublesome phrase 'intensity of desire' can perhaps also be eliminated if we speak in terms of prescriptions. A desire is said to be more intense than another concurrent and conflicting desire if and only if it is the former that is acted on – i.e. if the person who has them and is faced with the choice between assenting to, and thus acting on, the prescription which expresses the former desire, and doing the same for the latter desire, chooses the former (*MC*, p. 104).

Moreover, I have found no trace in *MT* and *HC* that he wishes to limit 'desire' or 'preference' in the way of *FR*.

There is no phenomenological evidence, I venture, that all or most desires are accompanied by bodily sensations that somehow can be experienced and felt. And even if I were wrong on this, it is dubious how we on the one hand can speak of desires and preferences in terms of dispositions to act, and on the other regard desires and preferences in terms of say, felt intensity. I am not at all convinced, for instance, that the symptoms of worry that I sometimes experience, and which I believe accompany certain desires of mine, necessarily are marks of the strength of such a desire (cf. Wetterström 1989, pp. 182–183).[21]

I will in what follows keep to the interpretation outlined in chapter II, where desires were understood in terms of dispositions to act.

What, then, does Hare say about the weighing procedure? Since the archangel, by imagining being in *b*'s place, with *b*'s desires, has actually acquired a conditional preference that is as much his preference as his

---

20. Hare says "The wide sense in which we are here using 'desire' is that in which *any* felt disposition to action counts as a desire; there is also a narrower and commoner sense in which desires are contrasted with other dispositions to action, such as a feeling of obligation (which in the wider sense of 'desire' could be called a desire to do what one ought" (*FR*, p. 170). Griffiths claims that "As Hare uses the term ['preference'], I cannot have a preference without being conscious of it or being able to be conscious of it" (*op. cit.*, p. 504). However, he does not give any evidence for this claim. Cf. Locke (1981, p. 531): "I will also follow Hare in using 'want' and 'desire' interchangeably, and both in the broadest possible sense, in which desire is opposed neither to duty nor to reason, nor does it require any specific feeling or sensation, much less passion".

21. Wetterström (1989, p. 192) endorses the following thesis: "The dissatisfaction of one preference could bring more suffering (and less happiness) into existence than the dissatisfaction of another preference in spite of the latter preference's being the strongest one".

original one, he will, Hare maintains, solve this conflict as he solves any intrapersonal conflict of preferences:

> both the conflicting preferences are mine. I shall therefore deal with the conflict in exactly the same way as with that between two original preferences of my own (*MT*, p. 110).[22]

Hare does not further explain what this "way" consists in. His point is rather, I suspect, that since it cannot reasonably be denied that we in our daily life face and solve intrapersonal conflicts of preferences, the above situation will not constitute a particular problem (cf. the passage from *MC* above). The resultant preference - our preference all in all - will, as mentioned above, be a function of our conflicting preferences and their respective intensities and of nothing else.

Now, the theory of critical thinking in *MT* faces some problems that Hare has not yet come to terms with. It redounds to Hare's credit and seriousness as philosopher that he himself has noted that there is an air of unfinished business about some parts of his work (*HC*, p. 230; *MT*, p. 104). As a next step I will outline and comment on those issues that according to Hare still await to be properly dealt with. They concern the following three kinds of desires:

1. Irrational preferences.
2. Imprudent preferences.
3. External preferences, i.e. preferences which are for things other than experiences of the preferrer.

The problems connected with the above kind of preferences are problems for any rationalistic account of moral reasoning that maintains that people's preferences must be taken into account. Hare has expressed his conviction that "a full account of the matter would assign weight to *all* preferences", but adds, "I must confess, however, that such a general theory of preferences is beyond my grasp at the moment" (*MT* p. 104). It is not entirely clear whether we should interpret Hare here as meaning that also irrational preferences (and not only imprudent and external preferences) belong to the group of preferences which he preliminarily excludes from the sphere of critical thinking. On the whole it is difficult to ascertain just what

22. This solution could very well be understood as taking place in the "background", or being in accordance with the mechanical view.

Hare's view is with regard to irrational preferences and their exclusion from critical thinking. Let me therefore begin by considering this matter.

5.5 Irrational Preferences

The idea that not only theoretical questions but also practical ones can be answered in a rational way, is today endorsed by several philosophers. However, the matter of to what extent or in what sense practical reasoning is rational is much debated. The Humean idea of practical reason is often criticized. Generally speaking, to know the answer to 'what shall I do?', which is what practical reason will tell us, the Humean tradition[23] has it that we must begin by finding out what we want. Practical reason is viewed as beginning with pre-rational intrinsic desires, and is considered to supply us with the best means to get these satisfied. Factual beliefs enter into this picture of human deliberation, essentially as beliefs about how to carry those ends and purposes into effect that are specified by the agent's desires.

The claim that practical reason involves pre-rational desires has been questioned by several writers, who instead endorse the idea of desire-independent reasons. Thus, for John Finnis, for instance, practical reason is first of all a thinking by which we "identify the desirable" - where "desirable" does not just stand for that which we happen to desire - but is rather that which appears to the agent as "somehow good to be getting, doing, having, being ..." (Finnis 1983, p. 35). He maintains that it is possible to predicate a want, about a person, by specifying what his reasons are, and this, the argument goes, can be done without referring to a desire.

The two views - that reasons are independent and dependent of desires respectively - are not exhaustive but, I believe, the most frequently discussed ones. Let me just state what I take to be a major problem for each theory. Thus, the former view must, for instance, find a way of dealing with the question 'Why should we want to do what we have reasons to do?', which on this view seems an intelligible thing to ask. Given the idea that wants can be predicted of people by stating reasons which do not refer to wants, Ingmar Persson (1981, p. 89; cf. Williams 1981), has argued convincingly that an answer to the question involves us in an infinite

23. Hume's view on this matter is complex, and I do not pretend here, with this presentation of a simple model, to express his idea. The view is, however, often referred to as "Humean" in the literature on practical reasoning (cf. Williams 1981, p. 102, who prefers the expression "sub-Humean model").

regress.[24] The latter view must, in its turn, come to terms with the problem of *Akrasia* or weakness of will. If there is such a tight bond between reasons and wants, how is it that we sometimes, at least seemingly, act against what we consider are our best reasons for acting? Unless we deny the existence of *akrasia* or explain it as following from mistaken beliefs about what our best reasons are, an answer to this question will apparently involve a third element besides beliefs and desires - which suggests that such a theory of practical reason is at best incomplete (for Hare's view see *LM*, II.2, *FR*, 2.8, *MT*, 3.7).

Now, Hare does not embrace the idea of desire-independent reasons in practical thinking.[25] We can, he argues, speak meaningfully about rational desires and preferences (See e.g. *HC*, p. 202, 1988).

As was outlined earlier, in chapter IV, Hare sometimes restricts the term 'rational preference' to those preferences which survive maximal criticism by facts and logic (e.g. *MT*, p. 214). The idea of there being rational preferences is one endorsed by several writers, notably by Brandt (1979). The view is intuitively appealing. New beliefs often give rise to new desires. Generally speaking we seem disposed to give up a desire if it has been based on false or insufficient beliefs, or if it is inconsistent with some other desire, the fulfilment of which we are not ready to give up in order to fulfil our other desire.

Now, consider the following reply which Hare makes regarding Brandt's question of how a critical thinker should act when he realizes that the person affected by his act has an irrational desire (Brandt 1988, p. 36). Hare comments:

> Brandt has himself given a good account of these [irrational preferences], from which I did not think I was dissenting. In *MT* I specifically mentioned Brandt's 'cognitive psychotherapy' as a means of discarding all but prudent preferences. *Brandt and I both think that the preferences which have to be considered are those which survive this 'cognitive psychotherapy'.* These may include some bizarre ones, but these are not irrational by Brandt's

---

24. There is a vast literature on the subject matter of practical reasoning. Some recent relevant titles are Williams 1981, Hudson 1988 in *HC*; Wallace 1988, which contains a correct, I think, criticism of Finnis's views. I have found much valuable material in Clarke 1985. Schueler 1989 contains a challenge to the Humean tradition. See also Hare 1963b, also in *MC*.

25. See for instance "Wanting: Some Pitfalls" (1968), and his example of the heartless James who desires his Uncle John's fortune (republished in *PI*). For a recent account see also "How to Decide Moral Questions Rationally" in Hare 1988.

> definition, provided that they 'would not extinguish after cognitive psychotherapy' (*HC*, p. 217–218; my italics).

Hare is here evidently suggesting that only the rational preferences of others should be considered. Moreover, he expresses his alliance to Brandt's theory of 'cognitive psychotherapy'. The latter fact considerably complicates an understanding of Hare, since it is unclear to what extent Hare is ready to endorse the different claims which Brandt makes in regard to the performance of what he calls "cognitive psychotherapy". Brandt is a firm believer, for instance, that not only instrumental desires but also, and more importantly, that at least some intrinsic desires can be shown to be irrational by what he calls cognitive psychotherapy (Brandt 1979, e.g. p. 111). I do not think we should take the above as indicating that Hare is ready to agree with Brandt on this issue.[26] The magnitude of such a step would be considerable. This suggests that were Hare to endorse such a line of reasoning, it would be reasonable to expect Hare to have made it explicit. Moreover, such a claim appears to square badly with what Hare maintains elsewhere. The suggestion that Hare parts company with Brandt on this issue gets support at the end in *MT*, for instance, where he says:

> there remains an irreducible and large minimum of sheer autonomous preferences which rational thinking can only accept for what they are, or will be. To that extent Hume was right (*MT*, p. 226).

To avoid confusing what Brandt says with what Hare claims, I suggest that we focus on Hare's own formulations. Consider the following passages:

> when I am considering the desires of others, considering what they would be if those others were perfectly prudent – i.e. desired what they would desire if they were fully informed and unconfused (*ET*, p. 218, 1976a).

> I might be seeking to make my present preference as rational as possible by exposing it, as Professor Brandt bids, to 'cognitive psychotherapy', i.e. to logic and the facts (*MT*, p. 101).

> It is in accordance with our method to assign equal weight, strength for strength, to all preferences alike, provided that they survive exposure to logic and the facts (*MT*, p. 180).

Evidently Hare is maintaining that we should take into consideration only those of a person's desires that this person would have if he were "fully

26. Brandt's claim that intrinsic desires can be irrational has been criticized by Egonsson (1990), who shows that there is an alternative way, consistent with the Humean tradition, of interpreting the examples which Brandt thinks evidences his claim.

informed and unconfused" or which had survived an exposure to 'logic and facts'. It will be convenient to formulate an imperative directed to the critical thinker that expresses this idea:

> *R3*: Pay attention only to those of an agent's desires that are rational, i.e. desires that the agent would have, if he were fully informed and unconfused' (i.e. had exposed himself to logic and facts).

The idea that only an agent's "idealized desires" (i.e. those which he would have, were he in the ideal situation, which in Hare's case is that of being fully informed and unconfused) should be considered, is shared by other utilitarians.[27]

It should be clear that for anyone but an omniscient being, *R3* is unreasonably demanding (cf. Friedman 1989; Egonsson 1990).[28] There would always be more facts to consider. Moreover, from the point of view of the critical thinker who wonders which preference to consider, the situation seems equally hopeless. Since the human critical thinker neither can be said to possess full information, he will in effect not know what is the other person's idealized rational desire. This does not necessarily mean that *R3* is entirely out of place. As we shall see later on, an omniscient being plays a important role in the establishment of what I have referred to as Hare's *T2* thesis. But as an imperative directed to human beings, *R3* appears impossible to comply with.

Should we then entirely put critical thinking aside as a mere possibility for omniscient beings? Although Hare underlines that what we, *qua* less than omniscient beings, can accomplish in the field of critical moral thinking, is not completely satisfactory, he clearly conceives of situations in which human beings should be performing critical thinking. However, in those cases it does not become clear what we should replace *R3* with.

Let me here come with a suggestion: Once we focus on a human critical thinker, it seems natural to adjust *R3* in such a way that it shifts attention

---

27. See e.g. Harsanyi (1982, p. 55), who says that a person's "rational wants are those which are consistent with his true preferences", the latter being explained as follows: "a person's true preferences are the preferences he *would* have if he had all the relevant factual information, always reasoned with the greatest possible care, and were in a state of mind most conducive to rational choice". Contrast Egonsson 1990, p. 93.

28. Friedman maintains that Hare's method of critical thinking as well as Rawls's contractual model "constitute excessively demanding psychological/cognitive feats". She pays most attention, however, to the psychological aspects of his method. Egonsson's (1990) examination of rationality requirements is here relevant. See e.g. pp. 95–99.

from what is rational viewed from the perspective of *b*, to what is rational, relative the critical thinker's full use of rationally supported beliefs, i.e., beliefs supported by logic and relevant available facts. Whether *b*'s preference is rational or not will be judged from the critical thinker's full use of rationally supported beliefs, which includes beliefs about what are *b*'s beliefs and motivational set up. For instance, suppose I know that *b* has an actual desire to spend his winter vacation in the south of Sweden, because he wants to go cross-country skiing. Assume further that he had no other reason for wanting to spend his time in the south of Sweden. Believing that it seldom snows during winter in the south of Sweden, I would be justified in concluding that *b* should, on pain of being irrational, give up his desire.

A possible way of making *R3* a more reasonable demand on less than omniscient beings would perhaps be to adjust *R3* to the following effect:

> *R4*: Pay attention only to those of an agent's desires that you have reason to believe would survive, were your rationally supported beliefs fully used by the agent (cf. Egonsson 1990, p. 99).[29]

With regard to human beings, *R4* would be a reasonable replacement for *R3*.

Now, it is a much discussed issue to what extent, if any at all, we should put restrictions on the kind of desires to consider in moral deliberation. Utilitarianism, of the desire form that Hare thinks critical thinking consists in, has actually been criticized from two sides. On the one hand, it is often held, for instance, that unrestricted utilitarianism is counter-intuitive on the grounds that it is incompatible with demands of justice or equality (Dworkin 1979, Harsanyi 1988).[30] Thus, critics as well as some utilitarians

---

29. Although *R4* appears more reasonable, from a human perspective, it is debatable whether it stands in need of being qualified. For instance, there will be times, I am sure, when I will consider someone as having better founded beliefs than I have. As I understand *R4*, such a case would not necessarily upset it.

30. Hare's attempt to meet the objection from counter-intuitiveness cannot be dealt with here. In fine, Hare tries to do so by invoking his distinction between critical and intuitive thinking (*MT*, chapter 8; cf. Hare, 1976a, p. 216). The reason why utilitarianism may go counter to our intuitions springs, he claims, from not realizing that those moral principles and related intuitions which utilitarianism is said to upset, only are of a *prima facie* nature in that they can be overridden or adjusted by the conclusions of critical thinking. Conducing our thinking on the critical level, there will be no room for a conflict between intuitive thinking and the conclusions of utilitarianism, since there is on this level a logical ban on any appeal to moral intuitions. Moreover, since Hare argues for the epistemologic priority of critical thinking (*MT*, p. 46) he furnishes his critics with the following problem: The reasoning on the critical level cannot, on pain of committing a *petitio principii*, be objected to on grounds that the conclusions of this kind of reasoning conflict with our moral intuitions. Critical thinking is prior in that the

have called upon an exclusion of certain malevolent preferences that have been considered "antisocial", to borrow a term from Harsanyi (1982, p. 56), such as sadism, envy, and resentment.

On the other hand, utilitarianism of the idealized kind has been considered too restrictive in that it excludes people's actual desires.[31]

The idea that the critical thinker should take into consideration only those preferences of other persons which they would have, were they rational, is far from uncontroversial.[32] It has been criticized by various writers. As mentioned earlier, Richard Brandt, for instance, makes the point that "one person may not be interested in the desires of another if he thinks them irrational" (Brandt 1988, p. 36). Peter Sandøe, in his turn, follows up this claim:

> In general, how does Hare avoid the conclusion that we should give people what they *would* prefer if they were rational, even when this because of their irrationality is pointless or even damaging? Brandt points out the same difficulty (*HC*, p. 33 & p. 41), but Hare in his reply simply seems to miss the point (Sandøe 1989b, p. 218).[33]

The situation which Sandøe appears to have in mind here corresponds to the case where a person *b* has an *actual* desire which the critical thinker

---

principles used on the intuitive level should be selected by critical thinking (cf. Frankena 1988, p. 780).

31. Several modern critics of utilitarianism trace the defect of this theory to another feature, viz., the alleged fact that it ignores the separateness of persons. See e.g. Rawls (1971), Williams (1973), and Hart (1979). Hardin (1988) has a recent and thorough account of the objections made against utilitarianism. Also relevant is Trigg (1988).

32. J. W. McGray has recently objected that rationality and logic do not necessarily dictate in a multilateral preference-conflict, as Hare says, only one universal prescription. A moral agent might "prescribe partially as a utilitarian and partially as concerned that no person should suffer significant and avoidable frustration" (McGray, 1986–87, p. 84 s.f.). This follows, he thinks, once we understand that there is nothing irrational about restricting utilitarianism to the effect that it is something irreducibly bad or wrong to force individuals to "make substantial, uncompensated sacrifices"(*loc. cit.*). McGray may be right in this. However, he gives us no other reason than his contention that it is not irrational. To admit considerations of the above kind into critical thinking, in order to show that there is more than one option open to a rational moral agent, seems nevertheless to beg the question. Such a person is reasoning not on the critical but on the intuitive level, and on the latter level Hare has never argued that the conclusions are the same as is required by utilitarianism. Regarding the objections put forward by Richards (1988), I believe these are met by Hare's comments in *HC*. Cf. *MT*, p. 49.

33. See here also Sen & Williams (1982, p. 10); Narveson in Potter & Timmons (1985, p. 32); Griffin (1986, p. 11), and Egonsson (1990, p. 100).

considers to be irrational. That is, the critical thinker believes that if *b* complied with *R2* he would, on pain of being irrational, acquire what we may call his *ideal* desire. Sandøe mentions the example of someone who is afraid of getting AIDS, and therefore does not want to give blood. If such a person knew under what conditions blood is given in Denmark, he would, Sandøe is convinced, no longer resist giving blood (for those reasons, at least). However, the problem is that "some of these people are so ignorant and unintelligent that there is no hope that they ever *will* be able to know the relevant facts and draw the relevant inferences from them" (*op. cit.*, p. 218).

The situation to which Sandøe draws attention to here is, in other words, the following: The critical thinker believes that the person, *b*, who has an actual irrational desire, will not ever acquire - because of *b*'s irrationality - what the critical thinker considers to be *b*'s rational ideal desire. Moreover, since *b* will not have this rational preference there seems no point in satisfying it.

Sandøe raises here an interesting issue. Moreover, I am inclined to agree with Sandøe that it would be objectionable in at least some cases to disregard an actual, albeit irrational desire, and consider instead a theoretical, albeit rational desire.[34] On the face of it, a critical thinker would in such a case do well to take the irrational desire into consideration.

However, there is little point in pursuing this line of criticism here. Hare's view on the matter of irrational preferences does not enable us to say what view he would take about a case such as Sandøe outlines. It will here be illuminating to consider a passage from *HC*, in which Hare discusses *ideal preferences*:

> Brandt's point in parentheses, that my argument to act-utilitarianism requires that we have to form a conditional preference corresponding to the *ideal*, not the *actual* preference of the other person, cannot be dealt with fully in the space available. Briefly, the answer is that act-utilitarianism seeks to maximize utility over time for the person affected. If, therefore a person's present actual preference, if realized, would lead to the frustration of a greater future preference, it is not prudent or ideal. I therefore excluded it by my 'requirement of prudence' (*HC*, p. 217, note 33/23).

For the moment let met put Hare's requirement of prudence aside - it will occupy us in the next section. Hare is here referring to a case in which

34. See Keekok Lee (1985, pp. 65-66) for an example that raises much the same questions as Sandøe's example.

the critical thinker is convinced that a person's actual preference is in conflict with a *future* preference of *b*, i.e., a preference which the critical thinker believes *b* will have some time later on. In such a case it does not seem unreasonable to put the actual desire aside, and instead take into consideration the future desire.[35]

Some lines further on Hare addresses the issue of irrational desires. The preferences which he thinks we should consider are the preferences which survive 'cognitive psychotherapy'. Moreover, he claims with regard to these preferences that have survived the following:

> These may include some bizarre ones, but these are not irrational by Brandt's definition, provided that they 'would not extinguish after cognitive psychotherapy' (1979:113), so he cannot be thinking of them here. In my view it is morally right to give equal weight to such bizarre preferences, because, however bizarre, they are important to the people who have them (*HC*, p. 218).

Hare's treatment of this issue is, as mentioned, brief. There is nothing in the above passages that suggests that Hare would actually go along with Sandøe's assumption that what the critical thinker should do is to satisfy the ideal rational desire.[36]

The question whether we should take into account preferences that we consider are irrational is not one that in the long run ought to be circumvented by Hare. Moreover, since I think it is not much of an exaggeration, nor too disheartening to say that irrational preferences abound that will not be adjusted or changed by presenting (further) facts, and since I think people in general show a tendency to consider their moral

---

35. In *MT* (pp. 177ff.) Hare discusses a case which at one point seems to concern the possibility of a doctor who realizes that he ought to give up a certain preference in the light of the facts but who nevertheless is unable to do so. In such a case, Hare says, "It may be that by going against his conviction in this case, albeit at the cost of great suffering, he may make it easier for himself and others to overcome it in future cases, so that in the end the attitude will get generally abandonded" (*MT* p. 181). However, it is not clear whether we should, *qua* critical thinkers, disregard the preference at issue. With regard to the case in question, see Wetterström (1989).

36. Hare does not avail himself of the distinction between intrinsic and instrumental desires of a person, which seems to me valuable when considering what preferences to take into consideration. See Egonsson (1990), for instance, who questions whether it is of any "intrinsic importance" whether the instrumental desire of someone is rational or irrational. His conclusion is that we, as utilitarians, should try to maximize people's intrinsic desires (Egonsson 1990, pp. 106–112). Supposing this is a correct conclusion to draw, we seem to be in a position to formulate a criterion, namely: Take only those actual preferences into account that will not frustrate the maximization of the person's intrinsic desires.

opponents as being irrational (that is, when there is seemingly no way to solve their disagreement), this problem must be dealt with by any theory that claims to yield a method for how to solve moral problems. In the present state of his theory the problem does not get the attention it deserves.

5.6 Imprudent Preferences

The second restriction of a provisional nature which Hare makes, concerns a certain kind of imprudent behaviour. Thus, Hare in *MT* brings forward the following requirement of prudence:

> we should always have a dominant or overriding preference now that the satisfaction of our now-for-now and then-for-then preferences should be maximized (*MT*, p. 105).

The former kind of preferences are preferences for states of affairs that would obtain at the time of having the preference. If the time is not the present but some future (or past) moment we may speak of then-for-then preferences. Let me give an example. I would, at the moment of writing these words, prefer to lean back and smoke a cigarette. Having smoked my cigarette, I no longer have a now-for-now desire to smoke a cigarette. However, I know all too well from experience that in one hour or so, I will then want a new cigarette. Such now-for-now and then-for-then preferences must in their turn be distinguished from now-for-then preferences that are for states of affairs that will take place at a time which differs from the time at which we actually have the preference. Thus, I prefer now, say, that at some later time (then) some event should take place rather than not, such as that I would have the will power to not give in to my desire to smoke. I can entertain such a preference and nevertheless believe with great certainty that I, then, will have at that future moment a then-for-then preference that is inconsistent with my now-for-then preference.

The requirement of prudence demands that we should not satisfy a now-for-then preference at the cost of a then-for-then preference. This would be, according to Hare, to show imprudence in one of its senses.

Hare puts forward the requirement of prudence as a "simplifying assumption" (*MT*, p. 105) in order not to have to deal with complications which Brandt has brought out, namely whether (and how) weight should be

assigned to the duration of preferences.[37] The requirement in question has been a subject of lively discussion (see here Hare 1984c, where he comments on the criticism put forward by Persson 1983, Schueler 1984, and Feldman 1984; see also Rabinowicz 1989, and Hare's comment in Hare 1989d). However, since Hare more than once has expressed his conviction (e.g. in *HC* p. 217) that the requirement of prudence can be dispensed with, I will here confine myself to bringing attention to one argument of Persson's that Hare does not comment on in his reply.

Persson points out that when an agent realizes that there is a difference between his now-for-then preference and his then-for-then preference, it is likely to be due to changes in those of the agent's beliefs that are present in his mind - what Persson calls the agent's representations. Whether he should for prudential reasons let his then-for-then preference override his now-for-then preference depends, Persson argues, on whether the change is due to a narrowing or extending of the scope of his representations. Suppose that I believe that the reason why I, then, will want X, is that I will hold a set of representations that has deteriorated in comparison to my present representations. Persson maintains convincingly that, in such a case,

> complying with Hare's requirement of prudence seems the very opposite of prudence: surely, it is imprudent to decide to act on preferences one believes to be based on confused or incomplete representations. For the prudential person, the then-for-then preference must here give way (Persson 1983, p. 46)[38]

On the other hand, Hare might reply that if it is the case that the then-for-then preference is based on confused or incomplete representations, then it is excluded in the first place as being irrational. The requirement would be valid in all but those cases when the then-for-then preference

37. Thus, Brandt asks "Does the length of time a person entertains a desire make a difference, so that a sadistic wish for an angry hour counts less than a wish entertained for a whole month (say only 1/720 as much)?" (1979, p. 250).

38. In his reply to Persson (1983), Schueler (1984) and Feldman (1984), Hare dissociates himself from what he calls the veto-theory, by which he understands a theory embracing the following claim "A fully rational critical thinker could never endorse any ought-statement that prescribed an action that anyone preferred not to occur" (*ET*, p. 248, 1984c; Hare quotes here Feldman 1984, p. 277). The objections of the above-mentioned writers derive, at least in part, he thinks, from interpreting him as holding such a veto-theory. Notwithstanding, I do not see how a rejection of the veto-theory somehow should counter the point made by Persson to which I bring attention. For a writer who clearly seems to interpret Hare as holding the so-called veto-theory, see Roxbee Cox (1986, p. 12).

would be irrational, i.e., when it would not survive cognitive psychotherapy. However, such a reply does not get to the bottom of the problem. As we saw, the notion of rationality which Hare endorses, admits that a person *a* is rational when he has a certain desire, whereas another person *b* is irrational when he has the same desire, namely, in the case that *a*'s but not *b*'s preference would survive cognitive psychotherapy (an exposure to facts and logic would make *b* but not *a* withdraw his preference). As a consequence of this we cannot rule out the possibility that although a certain preference that I have now would not now survive cognitive psychotherapy, it would do so at a future time (for instance, if I came to acquire *b*'s universal properties). In such a case, assuming that we believe that the then-for-then preference is founded on confused or insufficient beliefs, it still seems imprudent not to act on a now-for-then preference that runs counter to the then-for-then preference.

5.7 External Preferences

The third[39] group of preferences which Hare temporarily excludes is what he calls, following (Dworkin 1979), a person's "external preferences". These preferences, Hare explains, consist of "those preferences which are for things other than experiences of the preferrer" (Hare 1989c, p. 177).[40] For instance, my desire that people should not fare badly is an external preference, whereas my desire that I *know* whether my neighbour has had an accident or not is what we may call a *personal preference*, i.e., a preference for the experience of the preferrer. The person who on his deathbed desires that once dead, he should be buried at such-and-such a place, has an external desire (which is also a now-for-then preference).

In a much commented article, Gibbard argues that if we allow weight to external preferences in a utilitarian balance, then two rational critical thinkers, in contrast to what Hare has maintained, may come to disagree.[41]

---

39. There is actually a fourth group of preferences that he excludes, namely past preferences that we do not hold now. See here Rabinowicz (1989) and Hare's comments in Hare 1989d.

40. Hare points out in *HC* that his usage of "external preference" "has somewhat the same meaning" as Dworkin's "external preferences" (*HC*, p. 265). Cf. Rabinowicz (1989).

41. Actually Gibbard uses a different terminology. My example above of an external preference would be according to Gibbard a universal preference tendency that is not rationally required. He speaks of a further kind of preferences, which are for something other than the experience of the person entertaining the preference (Gibbard 1988, p.

To avoid this result, one option suggested by Gibbard would be to exclude from moral thinking external preferences as being irrational. However, as Hare has pointed out, this has the drawback that he cannot then rely on his argument against the fanatic.[42] Excluding external preferences also leaves him with a weaker argument for why deathbed promises should be kept (cf. *HC*, pp. 232–233).[43]

The discussion between Gibbard and Hare underlines the crucial nature of the *CPR*-principle, since the dilemma which Gibbard draws attention to is founded on assuming that the scope of this principle is limited. However, if I understand Hare correctly, his reason for temporarily excluding external preferences does not derive from any doubts on his part that the *CPR*-principle applies to these preferences. His reason is rather that he has not found a good way of "avoiding the counter-intuitiveness of allowing them to count" (*HC*, p. 233 s.f.).[44]

More recently Mane Hajdin (1990) has suggested that the problem of external preferences can "be resolved by taking a more careful look at what is involved in our imaginatively putting ourselves in the shoes of others" (*op. cit.*, p. 306). Hajdin argues that what critical thinking requires of me is not only that I imagine what it is like for the other to have a preference. More importantly, it requires, Hajdin maintains:

---

63), viz., what he calls "self-pertinent but not rationally required preferences".

42. As mentioned, Hare deals with the problem of the fanatic in *MT* by showing either that the external preferences of the fanatic will be based on a logical or factual mistake, or, if they withstand such criticism, by treating them on a par with other preferences in a utilitarian weighing. See also *HC* (p. 246), where Hare says that one could be a fanatic with personal preferences.

43 See Silverstein (1972), who claims that the role-shift technique required of Hare's theory in *FR* cannot be applied to deceptive deathbed promises. Such a technique, he claims, is "logically incoherent", since what the promiser imagines will either be false or incomplete. It will be false if he imagines himself "in the hypothetical role as a dying man, knowing or learning that the promise made to him is deceptive, [since] this falsifies a crucial feature of the actual case. Yet neither can he imagine himself not knowing or learning that the promise is deceptive, for then a crucial feature of the case would be left out." Robinson (1975) comes to Hare's rescue, but his defence rests on a misinterpretation of Hare's universalizability requirement. He claims, for instance, that we do not have to imagine ourselves with the "'inclinations' and 'ideals'" of the person affected by our action (*op. cit.*, p. 579). Silverstein is wrong, though, in that imagining being the dying man would leave out a crucial feature. A complete description of the case would surely describe the situation of the dying man as being one of incomplete information. So what we should do as a critical thinker is to imagine being in such a situation of incomplete information, which would be possible.

44. Sandøe (1989b) interprets Hare as if he "seems to agree that in case of some preferences the principle may not be true" (p. 219).

> my imagining what it is like for him to have these preferences *frustrated* (or satisfied) (Hajdin 1990, p. 306).

Once it is clear that what we must imagine is what it is like for the other person to have his preferences *frustrated* (satisfied), it can be seen, Hajdin claims, that "the very method of critical thinking provides for their [the external preferences'] exclusion" (*op. cit.*, p. 307).

Hajdin considers the following example:

> Suppose that someone has a preference that others do not engage in homosexual acts even if he will not know anything about it (nor be otherwise affected), and that I want to find out what moral consequences this has. According to the above, applying the method of critical moral thinking to this question will require me to imagine what it is like for him to be in his position (which will include my imagining what it is like for him to have this preference) *and* also to imagine what it is like for him to have this preference frustrated. But, *ex hypothesi*, imagining what it is like for him to have his preference frustrated is the same as imagining what it is like for him to have this preference satisfied (Hajdin 1990, p. 307).

Hajdin maintains that if I imagine the two situations (i.e. frustration and satisfaction of the preference), it will be two different representations only because I have "allowed into the thought experiment *my actual* knowledge as to whether homosexual activities are going on" (*op. cit.*, p. 307). But in such a case, I am no longer viewing the situation from the viewpoint of the other. Rather, I am, Hajdin continues, imagining the two situations from "some combination of his viewpoint and mine, something that is forbidden by the method of critical moral thinking" (*loc. cit.*).

Hajdin concludes that Hare does not have to make the *ad hoc* move and exclude external preferences. The reason is that we shall not form any conditional preference that corresponds to the external preference of the other. If we do form a preference it will only be because we have viewed the hypothetical situation from some combination of the other's viewpoint and our own viewpoint.

However, Hajdin's argument is not convincing. From the fact that *felt frustration* of the other's preference is relevant (i.e. has a causal impact on what we can prescribe), *when the frustration occurs*, it does not follow that frustrating the desire without the felt frustration is irrelevant. Recall how the relevancy of (external) preferences is argued for: Suppose a person *b* has an external preference. Given the truth of the *CPR*-principle, I will, by imagining myself as having this external preference, acquire a conditional

preference that corresponds to *b*'s external preference. My preference all in all will, in its turn, be a function of my conditional preference and whatever original preference I had. It is by such a line of reasoning that Hare argues for the relevancy of facts about other persons' preferences. The relevance of the external preferences of other persons is not made irrelevant by showing that in a different situation, namely when *b* actually experiences the frustration of his external preference, the frustration he feels will also be a relevant fact.[45]

There is nothing in the *CPR*-principle itself that suggests that it would not apply to external preferences (cf. Persson 1989). Hare is, I believe, well aware of this.[46] In *HC* he maintains the following, for instance, when he hesitates to exclude external preferences:

> My reason for being averse to doing so [arises] from the facts that my basic theory seems to require me to include *all* preferences, and that to exclude them would still leave me with difficulties stemming from the possibility of fanaticism and 'autofanaticism' (*HC*, p. 246).

It is to be hoped that Hare will present us with a more comprehensive account of critical thinking which shows how, being no peripheral phenomena in our lives, irrational, imprudent, and external preferences should be accounted for. As especially the two former restrictions stand today, they are bound to give rise to complications.[47]

5.8 Critical Thinking in Multilateral Cases

Suppose the *CPR*-principle rests firmly on a logical ground. Then it does seem, granted the exclusion of the above kind of preferences, as if the

---

45. This point can be considered from a different angle too. How great changes from the actual situation should we admit? Should we stop at considering the felt frustration it would mean to know that one's neighbour was homosexual? The frustration will presumably be greater in some cases for the person who has such an external preference, depending on who, where, and when the homosexual act is performed. I do not see how universalizability, prescriptivity, and the *CPR*-principle can logically require us to consider anything but a hypothetical situation that is exactly similar in universal respects to the actual situation.

46. See here Hare's note on Gibbard's claim that whether there are rational preferences to which the *CPR*-principle does not apply is an empirical claim (*HC*, p. 61). Hare makes it clear that for him it is an analytical matter (*HC*, p. 234).

47. Besides the ones already mentioned, there is another reason why they will complicate the issue, viz., that Hare must show why his view on rationality and/or prudence is the better one. See here for instance Norman O. Dahl (1986).

critical thinker will in a bilateral situation come to conclusions that correspond to those of a utilitarian. But will this claim extend to even multilateral cases as well? Hare, on his part, is convinced that it will. Brandt (*HC*, pp. 37–39), McDermott (1983), and Persson (1989), in contrast, have each argued that multilateral cases constitute serious threats to Hare's *T2*-claim. Since Hare has given a detailed reply to Persson's objections, it will be convenient to focus attention on their discussion of this issue.

Consider the following case: I, *a*, want to park my car on a certain spot, on which two other persons, *b* and *c*, want to park their respective bicycles. The desires of *b* and *c* are equally strong. Moreover, *a*'s desire is stronger than *b*'s and *c*'s separate desires, though *a*'s desire is not twice as strong as the either of them.

It follows now from the *CPR*-principle (which corresponds to what Persson calls the principle of hypothetical self-endorsement, or *PHS* for short) that if *a* imagines what it will be like to be *b* in the situation described above, call it S_1, he will acquire a conditional preference corresponding to *b*'s actual preference. The same holds true, *mutatis mutandis*, when he puts himself in the shoes of *c*.

As Persson correctly points out, when *a* has to find out what universal desire to form

> I must balance my preference to park my car against my preferences, with respect to hypothetical situations in which I am a bicyclist, that my bike here be parked. If I abide by PHS, I will in the actual situation S_1 have: (a) as regards S_1, a desire to park my car, (b) as regards a hypothetical situation S_2 in which I am one cyclist, a desire to park my bike, and (c) as regards another hypothetical situation S_3 in which I am the other cyclist, a desire to park my bike (Persson 1989, p. 165).

Persson's point is now that in order to form a universal desire with regard to S_1, I first weigh my desire (a) against my conditional desire (b). The result of this weighing will be that I form a universal desire to park the car, since, *ex hypothesi*, my desire (a) was stronger than my desire (b). The same outcome will follow when I weigh my desire (a) against my desire (c). Having performed these two operations I will be able to deduce the universal desire that I park my car in the location and prevent the two

bicyclists from parking their bikes. However, such a result is not what a utilitarian would require of me to arrive at.

In order to show that critical thinking even in multilateral cases will lead to conclusions endorsed by a utilitarian, the critical thinker must weigh his (a), (b), and (c) desire together. However, what is characteristic of this case is that the desires concern states of affairs in different possible worlds. I imagine that I, in one possible situation, S_2, have *b*'s preference, and in another world, S_3, that I have *c*'s preference. To add strength of preferences, the preferences must be jointly satisfiable. But this, Persson, rightly points out, only applies to preferences that concern the same possible world. This is not the case with the preferences involved here:

> I prefer(b) with strength S that if I were in the logical possible world S_2 the bicycle should be parked in R, where 'S_2' refers to the logical possible world in which I, *a*, had all of *b*'s universal properties, and there were two other persons with *a*'s and *c*'s universal properties respectively.
>
> I prefer(c) with strength S that if I were in the logical possible world S_3 the bicycle should be parked in R, where 'S_3' refers to the logical possible world in which I, *a*, had all of *c*'s universal properties, and there were two other persons with *a*'s and *b*'s universal properties respectively.
>
> I prefer(a) with strength S that in the actual world S_1 that the car should be parked in R.

Preferences (b) and (c) are not co-satisfiable, since they concern different possible worlds. Therefore, if Hare thinks the critical thinker must add this together, he must give us a reason why we should add these together.

In his reply to Persson's objection, Hare suggests that the solution to this problem depends on whether we think that there can be a possible world in which I both can have *b*'s and *c*'s universal properties "not of course at the same time in the same place, but at different times and perhaps also at different places" (Hare 1989c, p. 171).[48]

Hare replies to the effect that it is possible to imagine that in one and the same logical possible world the car owner occupies *b*'s and *c*'s roles. Hare maintains that "it is possible that the same situation, as regards its universal

48. That multilateral cases pose a problem for the account given in *FR* has been argued by some writers. Madell (1965), for instance, objects to Hare's claim that bilateral cases can be generalized to multilateral ones. Madell thinks that Hare supposes that the moral agent puts himself imaginatively in the place of all the parties "*at once*" (*op. cit.*, p. 39). Whether or not this was the case in *FR*, which I doubt, the passage quoted makes it redundantly clear that this is not the view he endorses in Hare 1989c. Another example is Harold J. White (1969).

properties, should recur many times in the same possible world" (*op. cit.*, pp 171-172). Moreover, he claims that "I have to suppose myself occupying, successively, universally identical situations but in different roles. There is nothing to prevent these identical situations succeeding one another in one possible world" (*op. cit.*, p. 172).

But, as Hare concedes a page further on, this reply does not in effect answer Persson's objection:

> However, this may be thought not to get to the root of Persson's objection. The objection was that there is no possible world in which I could occupy all three positions: the motorist's and the two cyclists'. For, it might be objected, it is logically impossible that any one person could occupy, even successively, different spatio-temporal positions, if these were separated in such a way as to break the continuity, mental and/or physical, which is thought to be a necessary condition of personal identity (Hare 1989c, p. 173).

Hare outlines an answer to this objection. He asks us to suppose that "certain episodes in people's lives recur after intervals of time in which the individuals who initially experienced them have ceased to exist" (Hare 1989c, p. 174). This would enable us to "have qualitatively identical situations recurring in all cycles" (*loc. cit.*). Moreover, if this were possible:

> we could ask someone to imagine that in one or more of these qualitatively identical situations he himself occupies different individual roles from that which he occupies in his actual time and place; and we could ask him to prescribe universally for such situations. The objection will then be that he cannot imagine this, because there will be an interruption of bodily and mental continuity between the cycles, and therefore the person in the other cycles could not be identical with the person in the present time and place. . . .To this I would reply that in spite of the interruption, it *is* possible to imagine being in all the roles successively (Hare 1989c, p. 174).

Thus, Hare counters by maintaining that it would in effect be possible for the motorist to imagine one possible world in which he would occupy successively the positions of *b* and *c*.

However, Hare's reply is not very compelling. To begin with, even if it was granted that we could imagine a logical possible world of the above character, Hare's argument would at most be effective for those who shared the view on personal identity which underlies Hare's answer, and which asserts that personal identity over time does not just consist in physical and/or psychological continuity. The fact that Hare's theory about critical thinking requires, with regard to multilateral cases, a controversial view on

personal identity, must be considered as a shortcoming for his attempt to derive utilitarianism from critical thinking.

A graver point is that Hare's reply does not answer the criticism that these preferences are not co-satisfiable. Thus, although it was granted that in one and the same possible world I could occupy the role of *b* and *c* *successively*, my *b* and *c* preferences would still not be co-satisfiable at a certain time in that possible world. But if they are not co-satisfiable, then we are considering a possible world that differs from the actual world, in which they are co-satisfiable at a certain time. The possible world which Hare is considering is not identical to the actual world, with regard to universal non-moral properties. The possible world which Hare is construing contains two successive situations, viz., one in which *a* has *b*'s preferences, and one in which he has *c*'s preferences. But the situation which Persson discusses is not the one of the motorist who asks whether on two successive occasions he ought to put his car on a certain spot, on which *b* and *c* wants to park their bikes. The discussion concerns rather whether *a* on a given occasion ought to park his car, viz., when *b* and *c* each prefers to park their bikes.[49]

Although Hare may be right that in a possible world we could imagine *a* occupying two successive roles, each with different preferences, these preferences would concern different states of affairs, and could therefore not be said to be conflicting.[50]

I conclude that Hare's argument for utilitarianism cannot rely only on[51] universalizability, prescriptivity, and the *CPR*-principle. Something more

---

49. In one place Hare misrepresents the situation which Persson discusses. Thus, he says "it seems clear that the motorist, who is asking whether he ought to move the bicycles to park his car on the two occasions, could be asked to suppose that he might be *going* to dream that he was in the successive situations" (Hare 1989c, pp. 175–176). But the case concerns not what the motorist should do on two occasions, but what he should do on one occasion, when two preferences are conflicting with his own one.

50. McDermott's point is precisely this that in a multilateral case the conditional preferences that I gain, all concern different cases, and therefore are not conflicting. McDermott (1983, p. 388).

51. According to Persson (1989, p. 166), a solution would be to require of the agent that when forming his opinion about what he ought to do, he should "*exclude knowledge of the numerical identities of the particulars involved*". In fine, by disregarding my knowledge that I am the car owner, I will face the option of a small probability of obtaining a greater satisfaction and a double probability (there are two bicyclists) of getting a somewhat smaller satisfaction. Persson thinks it would be irrational to choose the larger but less probable satisfaction.

needs to be brought in to account for multilateral cases of preference conflict.

5.9 In Criticism of the Prescriptivity of 'I'

In the preceding section I assumed the validity of the *CPR*-principle. Let us now examine it in more detail. It rests, we saw, on the idea that the word 'I' was prescriptive. That is to say, Hare attempts to account for the alleged analytical nature of the *CPR*-principle by suggesting that the word 'I', in addition to whatever descriptive meaning it has, also carries a prescriptive meaning element. As I hope to show in this section, this is a puzzling claim given what he says elsewhere about prescriptive word-meaning.

To begin with, allow me to repeat what was referred to above as sentences (1) and (2)

> (1) I now prefer with strength S that if I were in that situation x should happen rather than not.
>
> (2) If I were in that situation, I would prefer with strength S that x should happen rather than not (*MT*, p. 95).

Hare claims, as may be recalled, that a person cannot know that (2), without (1) being true, and that this is a logical truth.

Now, the point that I wish to argue is that the claim that 'I' is prescriptive is inconsistent with Hare's explanation from earlier works of what is involved in performing a prescriptive speech act. Moreover, since Hare does not further elucidate in *MT* what he means by 'prescribing' in the above context, he has not shown that it is due to the meaning element in 'I' that knowing what it is like to be someone else conceptually involves a conative state. Therefore, I will conclude that Hare has not given an account that unravels why the *CPR*-principle is a conceptual truth. All that remains is Hare's contention that (1) and (2) above are logically related in the way he thinks they are.

Allow me to recapitulate some central theses from chapter II to substantiate this objection. There I maintained that Hare endorses the following two theses (the third one is not relevant at the moment), with regard to what is involved in performing the speech act of prescribing.

> **(P)** A speaker *a* performs a prescriptive speech act in uttering the sentence (s), only if *a* by uttering (s) intends to tell the addressee(s) to make something

the case – where this something is determined by (s) and where the term "tell" should be understood in its generical sense.[52]

(**SC**) A speaker *a* prescribes sincerely (only if (**P**) holds, and) only if he intends that the addressee(s) should do what is prescribed.

Moreover, I showed that in Hare's view only sincere prescribing involves (**SC**). The prescriptive speech act is explainable (at least in principle) without invoking the speaker's intention (or wish) that the addressee should do the act prescribed. Such intentions are only required when the speaker is sincere or subscribing to the speech act. However, it is a necessary condition when a speaker prescribes, that he has the necessary intention to tell the adressee to make something the case (just as it is a necessary condition when a speaker asserts or states that something is the case, that he intends to tell that something is the case).

Furthermore, I drew attention to Hare's idea of how word-meaning and speech acts relate to each other:

> the meaning of a certain word can be explained, or partly explained, by saying that, when incorporated in an appropriate sentence in an appropriate place, it gives to that whole sentence the property that an utterance of it would be, in the appropriate context, a performance of a certain kind of speech act (*PI*, p. 75).

Thus, word-meaning was explained by Hare in terms of its contribution to the meaning of sentences, the meaning of which, in its turn, is explained at least in part in terms of speech acts.

Now with this and **P** and **SC** in mind, let us examine the claim that 'I' has a prescriptive meaning-element. By recognizing that were person *a* myself in some hypothetical situation, I would, according to Hare, in virtue of the prescriptivity of 'I' have a preference that the desires of myself should be satisfied.

It seems unreasonable to apply what Hare says about meaning and speech acts above to 'I'. For instance, since Hare claims that 'I' has prescriptive meaning, the question is whether we should then conclude that he holds the following: when 'I' is incorporated in an appropriate place in a sentence, it gives to that whole sentence the property that an utterance of it would be, in the appropriate context, a performance of a prescriptive speech act. In the light of the various types of sentences in which 'I' can

52. Recall that (P) may be regarded as being about intended and not actually performed speech acts.

figure, and given that I do not see how we can ascribe logical priority to any of these, it seems hard to uphold the view that the meaning of 'I' is explainable in terms of its contribution to prescriptive speech acts.

There is a more serious objection. Consider the following two sentences, which would be examples of prescriptions that would be expressions of the preference involved here:

(i) Satisfy the preferences I would have in the hypothetical situation S.

(ii) I ought to satisfy the preferences I would have in the hypothetical situation S.

To utter (i) would be an example of a prescription with the content required by what Hare says in relation to the prescriptivty of 'I'. Given that what Hare says about 'ought', uttering (ii) would also be a prescriptive speech act. However, both of these sentences are of a form such that in entertaining them we cannot be said to be imagining anything. Imagining seems necessarily to involve a hypothetical sentence such as (2) below:

(2) If I were in that situation, I would prefer with strength S that x should happen rather than not (*MT*, p. 95).

In contrast to what Hare in *MT* seems to be saying, I would say that we are clearly not in entertaining (2) telling ourselves to make something the case; neither (**P**) nor (**SC**) must be met in order to hold (2) sincerely, and I cannot therefore see how (2) could be prescriptive.[53]

Compare this with the following sentences:

(3) If I ought to pay my taxes, I will pay them.

(4) I ought to pay my taxes (and I will pay them).

Whereas Hare has argued that an utterance of type (4) will constitute the performance of a prescriptive speech act, I cannot find any support in his work for the idea that (3) also should have a prescriptive illocutionary force. Quite the reverse, Hare has maintained that certain sentence-types are such that they cannot be used to make prescriptive speech acts, although they contain a prescriptive word such as, for instance, 'ought'. Recollect the

53. There is the further problem, as argued earlier, of how to understand first-person sentences in the light of **P** and **SC**. However, the point that prescribing and issuing hypothetical judgements logically are separated, is sufficient for my purposes here.

argument between Searle and Hare, discussed in chapter II. Searle objected to Hare's idea that moral terms have prescriptive meaning, by arguing that occurrences of these words in hypothetical and interrogative sentences constitute a reductio of Hare's claim. In brief, although a sentence such as 'This is good', characteristically is used in order to prescribe (commend), Searle argued, a sentence such as 'If this is good, then so-and-so' is not. Words, Searle continued, have the same meaning whether they occur in, say, a hypothetical sentence or in a categorical affirmative.

Hare, as I have shown, agrees with Searle that in the case of hypothetical and interrogative sentences, the speech act of prescribing is not actually performed. However, he denied that such an admission means that we must give up the idea of explaining the meaning of, say, 'good' in terms of speech acts. According to Hare, what happens when prescriptive terms occur in a hypothetical sentence, is that the speech act is embedded or, as he sometimes puts it, is "in the offing" (*PI*, p. 79).[54]

Thus, what is noteworthy in Hare's reply to Searle is that he agrees that entertaining a hypothetical sentence is not a case of prescribing (cf. *PI*, pp. 86–89, 1970).[55]

Analyses of hypothetical sentences are complicated. What holds for one kind of hypothetical sentence does not have to be true of another kind. For instance, in the case of (2) above, 'I' figures in two places: in the *conditioning* clause "If I were in that situation," and in the *conditioned* one "I would prefer ...". Now, it may well be that when a prescriptive expression figures in the conditioned clause – as in 'If it rains, bring an umbrella/you ought to bring an umbrella', the prescriptive speech act is not in the offing.[56] Correspondingly, it could perhaps be argued that in (2) above it is the 'I' of the conditioned clause that makes (2) carry prescriptive meaning. However, such an interpretation is not convincing. The presence

54. As may be recalled, Hare accounts for the meaning of 'good' in a hypothetical sentence, by maintaining that the analysis of 'good' in a categorical affirmative is logically prior to the one in a hypothetical affirmative.

55. In "Meaning and Speech Acts" (*PI*, 1970) there is ample evidence for the view that I ascribe to Hare. Thus, in one place he says "categorical commendations do get transformed into something else when they get inserted into conditional clauses" (*PI*, p. 88; cf. p. 93). Notice that in his comment to Nagel, Hare claims that he holds the following thesis: "To prescribe that something be done is to express the desire that it be done" (*HC*, p. 250). The point which I am making is simply that issuing a hypothetical utterance cannot be a case of "prescribing that something be done".

56. Should we say the same about, e.g., 'If I want/ought to promote greatest possible preference satisfaction, I ought to promote greatest possible preference satisfaction'?

of 'would' in "I would prefer ..." strongly suggests that this conditioned clause should not be analysed as a sentence that we standardly utter with the intention to tell some one to make something the case.[57]

In the light of the above, it becomes hard to understand just what Hare means by saying that 'I' is prescriptive, or that "To recognize that that person would be myself is already to be prescribing" (*MT* p. 221).[58]

The kind of representation involved in critical thinking will, as far as I can see, necessarily involve a hypothetical sentence such as (2) above. Hence, I think Hare is stranded with the following problem: he cannot maintain that we are imagining and prescribing at the same time, without violating what he elsewhere has argued is involved in prescribing. Therefore, I conclude that Hare has not shown in what way 'I' can be a prescriptive term in the hypothetical context. The alleged analyticity of the *CPR*-principle accordingly remains to be accounted for.

It might be objected that although Hare sometimes characterizes what is involved in prescribing in terms of **P** and **SC**, there figures another characterization which may account for the prescriptivity of 'I'. This characterization is brought out by the following passage from *MT*:

> We say something prescriptive if and only if, for some act *A*, some situation *S* and some person *P*, if *P* were to assent (orally) to what we say, and not in *S*, do *A*, he logically must be assenting insincerely (*MT*, p. 21).

I have already criticized this account of prescribing, and will not repeat my objections to it here (see chapter II). However, assume that with the proper adjustments the above would explain how 'I' in the hypothetical sentence had a prescriptive meaning that was not in the offing. It would then make sense to say that the meaning of 'I' was such that in order to use it correctly (in at least some cases) we had to be in a certain motivational or conative state. We might even suppose that it was only this that Hare meant by his claim that 'I' is a prescriptive word. The *CPR*-principle would still be vulnerable to a different kind of objection that does not rest on the **P-SC** analysis of what is involved in prescribing. Suppose that 'I' had, as Hare claims, a prescriptive as well as a descriptive meaning-element, and that *eo*

---

57. Nor would a sentence containing 'would' make a suitable answer to the question "what shall I do?"

58. Among other things, it remains unclear what is the sentence-form, such that when 'I' is incorporated in it, an utterance of this sentence would be, in the appropriate context, a performance of a certain speech act, viz., prescribing.

ipso knowing what it is like to have in a hypothetical case the universal properties of someone else, conceptually involves not only being in a cognitive state but also involves being in a conative state. What prevents us from separating these meaning-elements from each other to the effect that we get a purely descriptive 'I'?[59]

If it were possible it would then no longer be the case that full representation logically involves a conative state. Moreover, a step such as the above should not appear alien to Hare. He has himself used a similar move in order to argue against attempts by descriptivists to show that applying certain descriptive words commits the speaker logically to certain attitudes or evaluations (e.g. *FR*, pp. 188ff.). We can always, Hare argues, "*alter* our conceptual apparatus - by treating as descriptive a word which used to be evaluative" (*FR*, p. 191).

To meet this argument Hare must show that 'I' is different in this respect. This would, in its turn, add to the conclusion from the former argument, viz., that by saying that 'I' is prescriptive Hare means by 'prescriptive' something that he has not explained or argued for.

59. I am grateful to Ingmar Persson for this idea (see Persson 1988). See also Sandøe (1989a, p. 197), who makes a similar point. As far as I can see, Hare (1989e) does not answer this objection in his reply to Sandøe. Cf. also Griffiths (1983, p. 509) who says about the alleged prescriptive meaning of 'I' that "if there is such an implication of the use of the word 'I', it is surely an eminently withdrawable one".

Concluding Remarks

Let me now bring this work to a close by summarizing the major points of my examination. In this essay I have striven to reach an understanding of what Hare considers is moral thinking on the fundamental, critical level, and what follows from such reasoning.

In chapters I and II I set out to examine his claim that moral judgements primarily have prescriptive meaning. My aim in chapter I was to consider Hare's argument for why we must analyse moral judgements as entailing at least one imperative, and why moral judgements in their primary sense cannot be purely descriptive. I maintained that a crucial premiss in Hare's argument was what I called (*S3*), namely that *only an imperative can be an answer to the practical question 'What shall I do?'*. Hare backs up his claim with the internalist contention that there is a necessary relation between assenting to a moral judgement and acting on it. A statement of fact cannot, according to Hare, answer the question 'What shall I do?', since such statements can only establish a contingent relation between assenting and acting. As was pointed out, it is interesting to note that for Hare only an imperative (speech act) can answer this question; to give an answer in terms of what the person who asked the question has best reason to do, will not answer this question.

Having presented Hare's argument, I suggested that it would actually be possible to retain an internalist position, and at the same time analyse moral judgements as descriptive statements. I proposed to understand these judgements as being statements about the addressee's reasons for acting. This seems promising, I ventured, as an analysis of 'ought'-judgements. As various writers have argued, to accept that one has a stronger reason to do something rather than avoiding it, and not do what one has stronger reason to do, is for logically reasons a sign that the person has changed his mind or been insincere (or suffered from weakness of will). Moreover, such an analysis will in an unqualified form have an advantage over a prescriptivist analysis: It will not, as is the case with a prescriptivist analysis, be confronted with questions of the following form: 'I ought to do *x*, but do I

have stronger reasons for doing *x* than for avoiding *x*?'. On a prescriptivist analysis, this question will be open (cf. Fumerton 1990, pp. 86-89).[1]

In addition, I examined in chapter I two objections that Hare puts forward against descriptive definitions of value terms. The first of these objections centres on the attempt to analyse value-judgements - such as the following: 'An *A* which is *C* is good' - as being equivalent in meaning to a descriptive judgement. Such attempts, the argument goes, fail, since a descriptive indicative sentence cannot be used to guide the actions/choice of the listener. However, I maintained that the argument rests on the claim that sentences such as 'An *A* which is *C* is good' is used to guide the actions/choice of the listener - a claim that in this context is question-begging.

The second argument that I looked into is an attempt by Hare to show that a distinction must be made between evaluative and descriptive meaning. This argument carries weight in that it indeed shows, given the validity of its steps, that the descriptive analysis which Hare considers - viz., 'X is a good wine' means 'X is a wine which has the non-evaluative properties ø' - is not successful. Moreover, it has an advantage over Moore's classical "open-question argument", to which it bears a close resemblance, that it seemingly puts the finger on what is wrong with equating value judgements with descriptive judgements: There is something we cannot do in issuing a descriptive judgement, which we can do by issuing a prescriptive judgement. Notwithstanding, I claimed that it is not sufficient to show that some naturalistic definition is incorrect, if we want to conclude that value terms carry more than descriptive meaning.

Hopefully, it became clear that I am in no disagreement with Hare with regard to the implausibility of certain forms of naturalism - namely, what he himself has called objectivist naturalism/descriptivism. Such forms of naturalism/descriptivism are hardly combinable with an internalist position. However, this does not hold true with regard to various subjectivist forms of naturalism. There may be other reasons why such subjectivist forms should be rejected. Such subjectivist analyses appear to me at least equally well

1. Fumerton actually proposes as a test for the adequacy of a metaethical view whether it considers such a question as open or closed. If the question remains open, the analysis must, he argues, be rejected. Whether or not we should side with Fumerton is not at issue here - I still have some doubts that concern whether and how universalizability enters into this analysis. This does not, from the view point of internalism, prevent this kind of analysis from being a serious alternative to a prescriptivism.

suited as a prescriptivist analysis to account for the internalist feature of moral language.

In chapter II I moved on to consider Hare's characterization of the speech act of prescribing. Central to this chapter is my interpretation of what Hare understands as the generic speech act of prescribing. I argued that we should ascribe to Hare notably two theses, namely that prescribing is, first and foremost, the intentional illocutionary act of telling someone to make something the case; secondly I showed that Hare considers it to be a sincerity condition for such speech acts that the speaker intends the addressee(s) to do what the speaker tells the addressee(s) to do.

The two theses of this interpretation - which I called (**P**) and (**SC**) respectively - are both open to criticism. The sincerity condition, for instance, requires of us to have intentions that we cannot possibly be said to have - on a reasonable interpretation of 'intention'. This holds true for 'ought'-judgements about the past as well as other kinds of 'ought'-judgements.

With regard to the claim that, in making moral judgements, we are performing the intentional act of telling someone to make something the case, I side with those of Hare's critics who claim that first- and third-person 'ought'-judgements are counter-examples to a prescriptivistic analysis. Either Hare must claim that 'ought' has a different meaning in the above mentioned cases, or he must explain how such cases can be examples of "telling someone to make something the case".

Chapter II also contains two suggestions about how Hare could revise his prescriptive analysis. Notably past tense 'ought'-judgements could be regarded - not as prescriptions that require of the speaker to have intentions to change the past, but rather as optative expressions that require of the speaker to wish that what had happened had not happened. Such a step is consistent with Hare's non-descriptivist and non-cognitivist position. However, it is unclear whether Hare would be prepared to take such a step. In the light of the fact that Hare explains word- and sentence-meaning in terms of illocutionary force or act potential, it seems that an effect of such a move would be that Hare would have to say that 'ought' in, say, 'You ought to have come' does not have the same meaning as 'You ought to come'. In the former we perform an optative speech act, whereas in the latter we are telling someone to make something the case.

I considered also briefly the possbility that Hare by "telling someone to make something the case" may have a wider notion in mind than the one

which I tried to capture by **P** and **SC**. This wider notion of prescribing would not claim that, in issuing a third-person 'ought'-judgement we were actually performing the intentional act of telling some absent person to make something the case. Prescribing should rather be viewed as that speech act which requires that we have the desire(/wish) that the absent person should do something.

Chapter III was devoted to Hare's influential universalizability-thesis. On my interpretation, to be universal a judgement or principle must meet two requirements: First, the principle must, if formalized, be governed by a universal quantifier; secondly, it must not contain any ineliminable individual constants. An issue that interested me concerned Hare's claim that moral judgements are universalizable in the same sense as descriptive judgements are. On the basis of more recent works by Hare, I suggested that Hare appears to have revised his claim. On a reasonable interpretation Hare should be understood as claiming that primarily evaluative judgements as well as descriptive judgements are universalizable - not for the same reasons but in the sense that holding either of them logically commits the speaker to a principle which must meet the two requirements mentioned above. This issue was also raised in connection with an account of Hare's views on supervenience, where I showed that Hare recently has made a distinction between trivial supervenience and substantial supervenience. The former relation is what we find in descriptive judgements, according to Hare, whereas the latter one is limited to value (and causal) judgements. In addition I was also interested in whether the account of the universalizability of value judgements given in terms of supervenience was consistent with the account given in terms of descriptive meaning. I maintained that it was.

In chapter III I also addressed the fundamental issue whether we are logically committed to universalize our 'ought' judgements. A way of focusing on this issue - one that Hare himself has used - is to ask whether we logically must have reasons for our 'ought'-judgements that do not contain ineliminable reference to individuals. Hare's claim that we must universalize was contested. 'Ought'-judgements do require reasons. So far my linguistic intuitions correspond to Hare's. However, there is nothing in the notion of having a reason that prevents it from containing a reference to an individual. It makes perfect sense to say, for instance, that his reason for doing so-and-so was that he believed that doing the act led to *his* well-being. Moreover, not only does it seem to be logically conceiveable that we

can have non-universalizable reasons for our 'ought'-judgements but that such non-universalizable ought-judgements are actually employed. Furthermore, I mentioned in this connection that Hare himself appears clearly aware of this possibility. In view of this I suggested that Hare's universalizability thesis should best be viewed as a suggestion aimed at clarifying our language.

Having examined Hare's views on prescriptivity and universalizability, I began in chapter IV my examination of how taking facts and logic into consideration will, according to Hare, constrain our moral reasoning. An issue that has interested various writers including myself is to what extent Hare can do without a substantial criterion of relevance. I gave in this connection an account of Hare's view on rationality that is relevant to an understanding of his view on this matter. In order to make a rational prescription a person should take facts that are relevant and available into account. If he has done this he will have complied with the rationality requirement. This much, I claimed, we may reasonably conclude with regard to Hare's view on this topic. However, in the absence of a more detailed account of 'available' it is difficult to estimate in greater detail just where Hare wants to draw the border between rationality and irrationality in this context.

Next I gave a preliminary account of Hare's views on relevance - showing that he endorses a causal one to the following effect: A property will be relevant for a person if a knowledge of this property will actually affect the persons preferences. This discussion was then followed by an examination of an objection which various writers have voiced against using universalizability-tests as a criterion for knowing what we morally ought to do. To be valid, it has been argued, such tests require some form of standardized act-descriptions. What interested me here was whether such a point of view constituted any serious threat to Hare's position. Hare rests his case not only on moral judgements being universalizable but also on their being prescriptive. My conclusion was that Hare's method does not require any standardized act-descriptions, nor will the fact that a situation can be described in many true ways constitute any threat to Hare's universal prescriptivism.

In the final section of chapter IV, I presented in the form of eight steps, an account of how to understand what critical thinking consists in. Moreover, this account, I stressed, was only preliminary. For the sake of convenience, I omitted from these steps Hare's more recent idea on

conditional representation. This idea is what most clearly marks a difference between Hare's view in *MT* and earlier works such as *FR*.

In chapter V I began by giving an account of what I called Hare's principle of conditional preference representation (*CPR*-principle, for short). This principle states that recognizing or imagining what it would be like to be in some other person's place with that person's universal properties, is already to be prescribing that, other things being equal, the preferences and prescriptions of that person should be satisfied. A central tenet in *MT*, which Hare attempts to establish, is that interpersonal preference-conflicts can be turned into intra-personal conflicts. The line of reasoning, I argued, is the following: Given that a person, say, *a*, has a desire to express a moral 'ought'-judgement (a universal, overriding prescription), *a* is rationally required to imagine what it is like for the person who is affected by the act which he considers doing. To know this *a* must imagine what it would be like to have the universal properties of that other person. Furthermore, given the validity of the *CPR*-principle, imagining this will generate in him a conditional preference that corresponds in strength and content to that of the person affected by the act (alternatively to the preference which *a* believes the other person has). As a result of such conditional representation, *a* will have an actual, acquired, conditional preference. This conditional preference will, together with any other original preference of *a*'s, determine, on balance, what action he is ready to prescribe universally all in all.

The process of weighing or balancing preferences referred to above has been interpreted in different ways. Notably the idea, in brief, that the preferences involved here are treated by the critical thinker as being in the *foreground of the decision* (Pettit and Smith 1990) was considered. This interpretation was contested on the ground that there is nothing in Hare's account of conditional representation that implies that this process requires that the preferences involved should be in the foreground. What the process does require is that the critical thinker forms actual, albeit, conditional preferences that correspond to the ones likely to be affected by the act he is considering doing.

Another exegetical matter concerns Hare's notion of "strength" or "intensity" of desires and preferences. Basing myself on works after *FR*, I maintained in contrast to some writers that there is no need to understand Hare as viewing strength as "felt" or experienced disposition to act.

Moreover, I claimed that there is no phenomenological evidence that all desires are accompanied by such experiences.

As Hare himself at least in part has made clear, three kinds of preferences raise difficulties for his view in *MT*, namely irrational, imprudent, and external preferences. I consider in this connection some objections that try to show either why more work on these matters is needed or that Hare's views stand in need of a revision. The most interesting conclusion that emerges from this discussion is, perhaps, that cogent reasons have been given that Hare's views on prudence require revision.

Assuming that Hare's account of the logic of 'ought' is correct and that the *CPR*-principle is valid, I consider whether critical thinking in multilateral cases poses a problem for Hare's claim that the conclusions of a critical thinker will correspond to those of a preference act-utlitarian. I consider especially the discussion between Hare and Persson (1989). Hare's attempt to meet the criticism is not convincing, and I conclude here with Persson that multilateral cases are indeed a serious problem for Hare's attempt to show that utilitarian conclusions follow from critical thinking in such cases. Moreover, in the light of the fact that bilateral cases are rare, this objection is all the more grave.

The *CPR*-principle raises many questions. In this essay I have focused on what seems to be the fundamental idea on which it rests, namely, that the word 'I' is a prescriptive term. I argue that such an idea seems difficult to square with what Hare elsewhere has to say about word-meaning. More importantly, the idea in question seems clearly inconsistent with what he specifically claims about prescriptive words that figure in conditional clauses of hypothetical sentences.

My conclusion is that further argument is needed to account for the alleged analytical relation which Hare argues holds between knowing or imagining having a hypothetical desire and having an actual, conditional desire. I strongly suspect that there is no "prescriptive" element involved in the meaning of 'I'.

In view of the above I suggest the general conclusion that Hare's argument for utilitarianism is not successful: it is questionable whether moral judgements carry prescriptive meaning in the sense Hare has explained; it is questionable whether 'ought' requires universalizable reasons, and Hare's idea that 'I' is prescriptive needs to be argued for. Thus,

the logical ground on which Hare in part founds his theory is less firm than he seems to think it is.

The discussions in this essay substantiate or are at least relevant for an even more general matter. I have in mind the issue of whether the methodology of building a moral theory on linguistic intuitions (and facts) has an advantage over a method that starts from some fundamental moral principle. Is it, for instance, so obvious that conflicts between fundamental moral principles are more common than linguistic conflicts? "Analytical" philosophers have failed to reach an agreement or consensus with regard to linguistic intuitions, and this fact reflects on the employment of linguistic intuitions in founding a moral theory. This essay has not shown this. However, it has hopefully given reasons why Hare's argument for utilitarianism has not been successful.

Bibliography

Alston, William P. (1963). "Meaning and Use", *Philosophical Quarterly* vol. 13, pp. 107–124. Reprinted with revisions and additions in Parkinson (1968).

Alston, William P. (1964). *Philosophy of Language*, Englewood Cliffs: New York, Prentice-Hall.

Anscombe, G. E. M. (1958). "Modern Moral Philosophy", *Philosophy* 33 reprinted in Hudson (1983).

Armstrong, D. M. (1968). *A Materialist Theory of the Mind*, New York: Routledge & Kegan Paul, New York.

Austin, J. L. (1980). *How to do Things with Words*, J.O. Urmson and Sbisá, Marina (eds.), Oxford: Clarendon Press (1962).

Bach, Kent (1976–77). "When to ask, 'What if everyone did that?'", *Philosophy and Phenomenological Research* 37, pp. 464–481.

Ball, Sthephen W. (1988). "Reductionism in Ethics and Science: A Contemporary Look at G.E. Moore's Open Question Argument", *American Philosophical Quarterly* vol. 25, nr 3.

Beardsmore, R. W. (1969). *Moral Reasoning*, London: Routledge & Kegan Paul.

Bergström, Lars (1964). "Review of *Freedom and Reason*, by R.M. Hare", Theoria 30, pp. 39–49.

Bergström, Lars (1970). "Imperatives and Contradicition", *Mind* vol. LXXIX, no.135, pp. 421–424.

Bergström, Lars (1982). "Interpersonal Utility Comparisons", *Grazer Philosophische Studien*, vol. 16/17 pp. 283–312.

Bergström, Lars (1990). *Grundbok i Värdeteori*, Stockholm: Thales.

Bilbeny, Norbert (1990). *El Discurso de la Ética; Historicidad y Lingüisticidad de la Ética*, Barcelona: Promociones y Publicaciones Universitarias, S. A.

Blackburn, Simon (1985). "Supervenience Revisited" in (ed.) Hacking, I (1985).

Blackburn, Simon (1986). *Spreading the Word; Grounding in the Philosophy of Language*, Oxford: Clarendon Press, Oxford (1984).

Blackstone, W T. (1965). "The Golden Rule: A Defense", *Southern Journal of Philosophy*, winter, pp. 172–77.

Blegvad, Mogens (1959). *Den Naturalistiske Fejlslutning*, København: Gyldendal.

Blose, B. L. (1974). "A Dilemma for Nondescriptivism", *Journal of Philosophy*, LXXXI, no. 20, nov 21. pp. 769–779.

Boatright, John R. (1973). "The Practicality of Moral Judgements", *Philosophical Quartly*, vol. 23, no. 93, pp. 316–334.

Bosque, Ignacio (1980). "Retrospective Imperatives", *Lingustic Inquiry* 11, p. 415.

Braithwaite, R. B. (1954). "The Language of Morals"; *Mind* vol. LXIII, no. 250 April, pp. 249–262.

Brandt, R. B. (1979). *A Theory of the Good and the Right*, Oxford: Clarendon Press.

Brandt, R. B. (1988). "Act-Utilitarianism and Metaethics" in (eds.) Seanor, D & Fotion, N (1988), pp. 27–41.

Brink, David O. (1989). *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press.

Broadie, Alexander (1972)., "Imperatives", *Mind* vol. LXXXI, no. 322 April, pp. 179–190.

Brunton, J. A. (1966). "Restricted Moralities", *Philosophy*, vol. XLI, no. 156, pp. 113–126.

Carr, C. R. (1977). "Speaker Meaning and Illocutionary Acts", *Philosophical Studies* 34, pp. 281–291.

Carter, David Kent (1982). *Hare's Theory of Morals*, Ann Arbor, Michigan: Dr. diss., Yale University.

Castañeda, Hector-Neri & Nakhnikian, George (1963)., (eds.) *Morality and the Language of Conduct*, Detroit: Wayne State University Press.
Castañeda, Hector-Neri (1955). "A Note on Imperative Logic", *Philosophical Studies*. vol. VI, no. 1, pp. 1–4.
Castañeda, Hector-Neri (1963). "Imperatives, Decisions and 'oughts': a Logico-Metaphysical Investigation"; in (eds.), Castañeda, H-C and Naknikian, G. (1963).
Caton, Charles E. (1963). "In what sense and why 'Ought'-Judgements are Universalizable", *Philosophical Quarterly* 13, pp. 48–55.
Charlton, William (1988). *Weakness of Will*, Oxford, Basil Blackwell.
Chellas, Brian F. (1971). "Imperatives", *Theoria* 37, pp. 114–129.
Clarke Jr, D. S. (1985). *Practical Inferences*, London, Boston: Routledge & Kegan Paul.
Cohen, Brenda (1967). "An Ethical Paradox", *Mind* no. 76, pp. 250–259.
Cohen, L. Jonathan (1964). "Do Illocutionary Forces exist?", *Philosophical Quarterly*, 14 pp. 118–137.
Cohen, L. Jonathan (1970). "Searle's Theory of Speech Acts", *Philosphical Review*, pp. 545–570.
Conway, D. (1972). "Descriptions and Evaluations", *Mind* 81, 590–594.
Cooper, Neil (1973). "Language, Purpose and Morality", *Philosophical Quarterly* vol. 23, no. 92 July, pp. 230–240.
Cottingham, John (1983). "Ethics and Partiality", *Philosophical Studies* vol. 43, no. 1 jan, pp. 83–99.
Crowell, Edward F. (1974). "Hare on Desire and Moral Reasoning", *Theoria* vol. XL, part 3, pp. 152–162.
Curtler, Hugh Mercer (1971). "What Kant might say to Hare", *Mind* 80, pp. 295–97.
D'Alessio, J. C. (1967). "Dispositions, Reduction Sentences and Causal Conditions", *Critica*, 1,3.
Dahl, Norman O. (1987). "A Prognosis for Universal Prescriptivism", *Philosophical Studies* 51, pp. 383–424.
Dale, A. J. (1985). "Hare on Supervenience", *Mind* vol. XCIV, no. 376, pp. 599–600.
Dancy, J, Moravcsik, J. M. E., and Taylor C. C. W, eds., (1988) *Human Agency; Language, Duty, and Value* Stanford, California: Stanford University Press.
Daniels, Charles (1970). "Hare on the Meaning of 'good'", *Mind*, vol. LXXIX, nr 13, pp. 139–141.
Devitt, Michael & Sterelny, Kim (1987). *Language and Reality; an Introduction to the Philosophy of Language*, Oxford: Blackwell.
Donagan, Alan (1965). "Mr Hare and the Conscientious Nazi", *Philosophical Studies* 16, no. 1–2, pp. 8–11.
Donagan, Alan (1977). *The Theory of Morality*, Chicago: The University of Chicago Press.
Dunn, Robert (1987). *The Possibility of Weakness of Will*, Indianapolis: Hackett publishing co. Durrant, Michael (1973). "Numerical Identity", *Mind*, vol. LXXXII, no. 325, p. 95–101.
Durrant, Michael & Attfield, Robin (1981). "Prescriptivity and Justification", *Philosophical Papers* 10, pp. 16–23.
Dworkin, R. (1977). *Taking Rights Seriously*, London: Duckworth.
Dworkin, R. (1979). "We do not have a Right to Liberty", from (ed.) Cunningham, R *Liberty and the Rule of Law*, College Station: Texas A & M University Press.
Edel, Abraham (1963). *Method in Ethical Theory*, Indianapolis, New York: Bobbs-Merrill.
Egonsson, Dan (1990). *Intersts, Utilitarianism and Moral Standing*, Dr diss, Lund: Lund University Press.
Emmet, Dorothy (1963). "Universalizability and Moral Judgements", *Philosophical Quarterly* 13, pp. 214–28.
Falk, W. D. (1947–48). "'Ought' and Motivation", *Proceedings of the Aristotelian Society*, 48. Reprinted in Falk (1986).
Falk, W. D. (1953). "Goading and Guiding", *Mind* vol. LXII, pp. 145–171.
Falk, W. D. (1986). *Ought, Reasons, and Morality*, London: Cornell University Press, Ithaca.

Feldman, F. (1984). "Hare's Proof", *Philosophical Studies* 45, pp. 269–283.
Finnis, J. (1983). *Fundamentals of Ethics*, Oxford: Oxford University Press.
Føllesdal, D. & Hilpinen, R. (1970). "Deontic Logic: An introduction" in Hilpinen, R. (ed.) (1971).
Foot, Philippa (1958). "Moral Beliefs", Aristotelian Society 59, republished in Foot (1978).
Foot, Philippa (1977). *Theories of Ethics*, Oxford, Oxford University Press.
Foot, Philippa (1978). *Virtues and Vices*, Oxford: Blackwell.
Forrester, Mary (1974). "An Argument for Descriptivism", *Journal of Philosophy*, LXXI, no. 20, nov 21 pp. 759–769.
Fotion, Nicholas (1983). "Hare, R M: Moral Thinking its Levels, Method and Point", *Ethics* 93, no. 4 july, pp. 800–801.
Frankena, William K. (1939). "The Naturalistic Fallacy", *Mind* vol. 48, pp. 464–477.
Frankena, William K. (1988). "Hare on the Levels of Moral Thinking" in (eds.), Seanor, D & Fotion, N (1988).
Frankena, William (1988). "Hare on Moral Weakness and the Definition of Morality", Ethics vol. 98, july no. 4 pp. 779–792.
Franklin, R. L. (1973). "Recent Work on Ethical Naturalism", *American Philosphical Quarterly*, monograph series, no. 7 pp. 55–83.
Friedman, M. (1989). "The Impracticality of Impartiality", *Journal of Philosophy*, vol. LXXXVI, november no. 11, p. 645.
Fullinwider, R. K. (1977). "Fanaticism and Hare's Moral Theory", *Ethics* 87, no. 2, pp. 165–181.
Fumerton, Richard A. (1990). *Reason and Morality; a Defense of the Egocentric Perspective*, Ithaca and London: Cornell University Press.
Furberg, Mats (1963). *Locutionary and Illocutionary Acts*, Dr Diss. Gothenburg Studies in Philosophy.
Furberg, Mats (1971). *Saying and Meaning*, Oxford: Basil Blackwell.
Geach, P. T. (1960). "Ascriptivism", *Philosophical Review*, vol. 69, pp. 221–226, also in (ed.) R. M. Rorty (1992). *The Linguistic Turn. Essays in Philosophical Method With Two Retrospective Essays*, Chicago and London: The University of Chicago Press.
Geach, P. T. (1977). "Again the logic of 'ought'", *Philosophy* vol. 52.
Geach, P. T. (1981). "On a Claim of Hare's", *Analysis* vol. 41, no. 1 p. 63.
Gellner, Ernest (1956). "Morality and *Je Ne Sais Quoi* Concepts", *Analysis* vol. 16, no. 5 pp. 97–103.
Gettner, A. (1972). "Hare And Fanaticism", *Ethics* 87, no. 2, pp. 165–181.
Gewirth, A. (1978). "The Golden Rule Rationalized", *Midwest Studies in Philosophy* III, pp. 133–47.
Gibbard, Allan (1988). "Hare's Analysis of 'Ought' and its Implications" in (eds), Seanor, D & Fotion, N (1988).
Gilbert, Joseph (1972). "Neutrality and Universalizability", *Personalist*, vol. 53, pp. 438–441.
Gilbert, Margaret (1972). "The Abilities of Prescriptivism", *Analysis* vol. 32, no. 4, pp. 141–144.
Godwin, W. (1793). *An Enquiry concerning Political Justice*.
Goldman, Alan H. (1987). "Red and Right", *Journal of Philosophy*, vol. LXXXIV, no. 7, july pp. 349–362.
Gorr, Michael (1885). "Reason, Impartiality, Utilitarianism" in (eds.) Potter, T & Timmons, M. (1985).
Gowans, Cristopher W. (1989). "Moral Dilemmas and Prescriptivism", *American Philosophical Quarterly* 26, pp. 187–197.
Grice, G. R. (1967). "Motive and Reason", Cambridge University Press, reprinted in ed., Raz, J. (1978).
Grice, H. P. (1968). "Utterer's meaning, Sentence-meaning, and Word-meaning", *Foundations of Language*, 4, pp. 1–18, reprinted in Searle, J. R. (1971).
Griffin, James (1986). *Well-Being; its Meaning, Measurement and Moral Importance*, Oxford: Clarendon Paperbacks.
Griffiths, A. Phillips (1983). "Hare's Moral Thinking", *Philosophy* 58, pp. 497–511

Guisán, Esperanza (1986). *Razón y Pasión en Ética; Los dilemas de la ética contemporánea*, Barcelona: Anthropos.
Hacking, I. (1985). (ed)., *Exercises in Analysis; Essays by students of Cassimir Levy*, Cambridge University Press.
Hajdin, Mane (1990). "External and Now-For-Then Preferences in Hare's Theory", *Dialogue* XXIX, pp. 305–310.
Hamblin, C. L. (1987). *Imperatives*, Oxford: Basil Blackwell.
Hancock, Roger (1963). "A Note on Hare's Language of Morals", *Philosophical Quarterly* 13, pp. 56–63.
Hardin, Russell (1988). *Morality within the Limits of Reason*, Chicago, London: The University of Chicago Press.
Hare, R M
LM = Hare (1952). *The Language of Morals*, Oxford, Oxford University Press.
FR = Hare (1963a). *Freedom and Reason*, Oxford, Oxford University Press.
PI = Hare (1971a). *Practical Inferences*, London and Basingstoke: MacMillan.
MC = Hare (1972a). *Essays on the Moral Concepts*, Berkeley and Los Angeles: University of California Press.
MP = Hare (1972b). *Applications of Moral Philosophy*, London and Basingstoke: MacMillan.
MT = Hare (1981). *Moral Thinking: Its Levels, Method, and Point*, Oxford: Clarendon Press.
ET = Hare (1989). *Essays in Ethical Theory*, Oxford: Clarendon Press.
PM = Hare (1989). *Essays on Political Morality*, Oxford: Clarendon Press.
Hare, R. M. (1949). "Imperative Sentences", *Mind*. Reprinted in *PI*.
Hare, R. M. (1952). *The Language of Morals*, Oxford, Oxford University Press.
Hare, R. M. (1954). "What is Value? – An essay in Philosophical Analysis by Everett W Hall", *Mind* vol. LXIII, no. 250, pp. 262–269.
Hare, R. M. (1955). "Universalisability", *Aristotelian society* 55, reprinted in Hare (1972).
Hare, R. M. (1957). "Geach: Good and Evil", *Analysis*, XVII, reprinted in *MC*.
Hare, R. M. (1960a). "Ethics" in ed. Urmson, J. O. *Encyclopedia of Western Philosophy and Philosophers*, (1960, Hutchinson).
Hare, R. M. (1960b). "Philosophical Discoveries", *Mind*, vol. LXIX, no. 274 pp. 145–162.
Hare, R. M. (1962). "Critical study of Generalizastion in Ethics", *Philosophical Quarterly* 12, pp. 351–355.
Hare, R. M. (1963a). *Freedom and Reason*, Oxford, Oxford University Press.
Hare, R. M. (1963b). "Descriptivism", British Academy 49. Reprinted in Hudson (1969) and in *MC*.
Hare, R. M. (1964a). "Pain and Evil", *Aristotelian. Society*. suppl. 38. Reprinted in *MC*.
Hare, R. M. (1964b). "The Promising Game", *Revue Internationale de Philosophie*. 70. Reprinted in Foot (1967) and in *ET*.
Hare, R. M. (1967). "Some Alleged Differences between Imperatives and Indicatives", *Mind* 76. Reprinted in *PI*.
Hare, R. M. (1968a). "Review of *Contemporary Moral Philosophy*, by G J. Warnock", *Mind* 77, pp. 436–440.
Hare, R. M. (1968b). "Wanting: Some Pitfalls", first publ. in *Agent, Action and Reason*, Proceedings of Western Ontarion Colloquium, (ed.), R. Binkley. Reprinted in *PI*.
Hare, R. M. (1969). "Practical Inferences" in (ed.) Kruse, V. (1969). *Festskrift till Alf Ross*. Reprinted in *PI*.
Hare, R. M. (1970). "Meaning and Speech Acts", *Philosophical Review* LXXIX. Reprinted in *PI*.
Hare, R. M. (1971a). *Practical Inferences*, London and Basingstoke: MacMillan.
Hare, R. M. (1971b). "Appendix: A Reply to Mr G. J. Warnock", *Philosophical Review* LXXX.

Hare, R. M. (1971c). "Austin's Distinction between Locutionary and Illocutionary Acts", originally published in *PI*.
Hare, R. M. (1972a). *Essays on the Moral Concepts*, Berkeley and Los Angeles: University of California Press.
Hare, R. M. (1972b). *Applications of Moral Philosophy*, London and Basingstoke: MacMillan.
Hare, R. M. (1972c). "Principles", *Aristotelian society* 72. Reprinted in *ET*.
Hare, R. M. (1972d). "Wrongness and Harm", originally published in *MC*.
Hare, R. M. (1972e). "Reasons of State", originally published in *MP*.
Hare, R. M. (1972f). "'Nothing Matters'", originally published in *MP*.
Hare, R. M. (1975). "Abortion and the Golden Rule", *Philosophy and Public Affairs*, 4.
Hare, R. M. (1976a). "Ethical Theory and Utilitarianism" in Contemporary British Philosophy 4, ed., H.D. Lewis (Allen and Unwin). Reprinted in *ET*.
Hare, R. M. (1976b). "Some Confusions about Subjectivity" in *Freedom and Morality*, ed. J. Bricke, Lindley Lectures, U. of Kansas. Reprinted in *ET*.
Hare, R. M. (1978b). "Relevance" in *Values and Morals*, (eds.), Goldman, A. I. and Kim, J. (Reidel). Reprinted in *ET*.
Hare, R. M. (1979a). "Universal and Past-Tense Prescriptions: A Reply to Mr Ibberson", *Analysis* oct.
Hare, R. M. (1979b). "Utilitarianism and Vicarious Affects" in *The Philosophy of Nicholas Rescher*, (ed.) Sosa, E. (Reidel).
Hare, R. M. (1979c). "What Makes Choices Rational", *Review of Metaphysics* 32; reprinted in *ET*.
Hare, R. M. (1981a). *Moral Thinking: Its Levels, Method, and Point*, Oxford: Clarendon Press.
Hare, R. M. (1981b). "On a Misunderstanding of Geach's", *Analysis* vol. 41, no. 1 p. 64.
Hare, R. M. (1984a). "Rights, Utility and Universalization", originally. in (ed.) R. Frey (1984). *Utility and Rights* Minnesota University Press. Republished in Hare *PM*.
Hare, R. M. (1984b). "Supervenience", *Aristotelian. Society*. supp. 58. Reprinted in *ET*.
Hare, R. M. (1984c). "Some Reasoning about Preferences: A Response to Essays by Persson, Feldman, and Schueler", *Ehtics* 95. Reprinted in *ET*.
Hare, R. M. (1985). "Ontology in Ethics" in (ed.), Honderich, T (1985). *Morality and Objectivity: Essays in Memory of John Mackie*: Routledge.
Hare, R. M. (1986). *A Reductio ad Absurdum of Descriptivism* in Philosophy in Britain Today, (ed.) S. Shanker (Croom Helm). Reprinted in *ET*.
Hare, R. M. (1987a). "Review of *The Philosophy of Rights and Wrong*, by B. Mayo", *Philosophical Quarterly*, 37.
Hare, R. M. (1987b). "Why Moral Language?" in P. Pettit *et al* (1987).
Hare, R. M. (1987c). "How to decide moral questiones rationally", *Critica* 18, also in *ET*.
Hare, R. M. (1988). "Comments on Gibbard" in *HC*.
Hare, R. M. (1988a). "Comments" in Seanor & Fotion (1988).
Hare, R. M. (1988b). "Some Sub-atomic Particles of Logic", *Mind*, vol. XCVIII, no. 389 Januari.
Hare, R. M. (1989a). *Essays in Ethical Theory*, Oxford: Clarendon Press.
Hare, R. M. (1989b). *Essays on Political Morality*, Oxford: Clarendon Press.
Hare, R. M. (1989c). "Universalizability and the Summing of Desires: Reply to Ingmar Persson", *Theoria*, vol. LV, part 3, pp. 171–177.
Hare, R. M. (1989d). "Prudence and Past Preferences: Reply to Wlodizimierz Rabinowicz", *Theoria*, vol. LV, part 3, pp. 153–158.
Hare, R. M. (1989e). "Amoralism: Reply to Peter Sandøe", *Theoria*, vol. LV, part 3, pp. 205–210.
Hare, R. M. (1989f). "Fanaticism: Reply to Thomas Wetterström", *Theoria*, vol. LV, part 3, pp. 186–190.
Hargrove, Eugene C. (1985). "The Role of Rules in Ethical Decision Making", *Inquiry* vol. 28, no. 1 pp. 3–42.
Harman, Gilbert (1977). *The Nature of Morality: An Introduction to ethics*, New York: Oxford University Press.

Harsanyi, J. C. (1982). "Morality and the Theory of Rational Behaviour" in Amartya Sen & Bernard Williams (eds.): *Utilitarianism and Beyond*, Cambridge: Cambridge University Press pp. 39–62.
Harsanyi, John C. (1988). "Problems with Act-utilitarianism and with Malevolent Preferences", see *HC* pp. 89–99.
Hart, H. L. A. (1979). "Between Utility and Rights", Columbia Law Review, 79, pp. 828–846.
Hartland-Swann, J. (1960). *An Analysis of Morals*, London: George Allen & Unwin LTD.
HC (1988). *Hare and Critics; Essays on Moral Thinking*, (eds.), Seanor, D & Fotion, N, Oxford: Clarendon Press.
Hermerén, Göran (1972). *Värdering och Objektivitet*, Lund: Studentlitteratur.
Hermerén, Göran (1988). *The Nature of Aesthetic Qualities*, Lund: Lund University Press
Hermerén, Göran (1992). "Allusions and Intentions" in (ed.) G. Iseminger *Intention and Interpretation*, Philadelphia: Temple University Press.
Hill, James Franklin (1974). *Universalizability and Prescriptivity: A Critical Evaluation of the Moral Philosophy of R. M. Hare*, Dr. diss., University of Georgia.
Hilpinen, Risto (1971). *Deontic Logic: Introductory and Systematic Readings*, Dordrecht-Holland: D. Reidel Publ. Company.
Hoche, H-U. (1983). "Zur logischen Struktur von 'Goldene-Regel' Argumenten im Sinne Hares", *Kantstudien* 74, pp. 453–478.
Holdcroft, David (1964). "Meaning and Illocutionary Acts" in (ed.) Parkinson, H. H. R. (1982).
Holmes, Robert L. (1966). "Descriptivism, Supervenience, and Universalizability", *Journal of Philosophy*, 63, pp. 113–119.
Hornsby, Jennifer (1988). "Things done with Words" in Dancy, Jonathan *et al* (1988). *Language, Duty, And Value*, Stanford California: Stanford University Press.
Hudson, W. D. (1974). *New Studies in Ethics* (2 vols.;), London: MacMillan.
Hudson, W. D. (1978). *Modern Moral Philosophy*, London: MacMillan Press Ltd.
Hudson, W. D. (1983)., (ed.) *The Is/Ought Question*, London: MacMillan.
Hudson, W. D. (1988). "The Development of Hare's Moral Philosophy" in *HC*.
Hume, David (1896). *A Treatise of Human Nature*, Oxford: Clarendon Press.
Huntley, Martin (1980). "Propositions and the Imperative", *Synthese* vol. 45, no. 2, pp. 281–293.
Ibberson, John (1979). "A Doubt about Universal Prescriptivism", *Analysis* vol. 39, no. 3, pp. 153–158.
Ibberson, John (1986). *The Language of Decision*, Houndmills, Basingstoke: MacMillan Press Ltd.
James, E W. (1972). "Working in and working to Principles: Penn's Lie and Hare's Myth of Universalizability", *Ethics* 83, pp. 51–57.
Johanson, A. (1988). "Imperative logic as based on a Galois connection", *Theoria*, vol. LIV, part 1, pp. 1–24.
Jost, Lawrence J. (1976). "On A Prescriptivist Dichotomy", *Mind* vol. LXXXV, no. 338, pp. 258–261.
Kambartel, Friedrich (1976). "Symbolic Acts; Remarks on the Foundations of a Pragmatic Theory of Language" in (ed.) Ryle, G. (1976).
Kant, I (1785). *Groundwork of the Metaphysic of Morals,* trans. H.J. Paton, New York, 1953.
Kenny, A. (1963). *Action, Emotion and Will,* London: Routledge & Kegan Paul.
Kerner, George C. (1965). "Hare's New Thoughts on Reason in Ethics", *Theoria*, vol. XXXI pp. 125–132.
Kerner, George C. (1966). *The Revolution in Ethical Theory*, New York: Oxford University Press.
Kese, Ralf (1989). "Zur Methodologie und Logik von 'Goldene-Regel'-Argumenten; Eine Gegenposition zu Hoches Hare-Exhaustion", *Kanstudien* 81, heft 1, pp. 89–98.
Knox, Alison (1970). "The Polemics of 'Descriptive Meaning'", *The Review of Metaphysics*, vol. XXIV, no. 2, Issue no. 94, pp. 245–275.

Kolenda, Konstantin (1975). "Moral Conflicts and Universalizabilty", *Philosophy* vol. 50, no. 194 October, pp. 460–465.
Kordig. Carl R. (1973). "Moral Weakness and Self-Reference", *Analysis* vol. 32, no. 1, pp. 11–12.
Kripke, S. (1980). *Naming and Necessity*, Cambridge, Mass.: Harward.
Kurtz, Paul (1970). "Has Ethical Natualism Been Refuted?", *Journal of Value Inquiry* 4, pp. 161–71
Lange, John F. (1966). "R. M. Hare's Reformulation of the Open Question", *Mind* 75, pp. 244–247.
Laymon, Ronald E. & Machamer, Peter K. (1970). "Personal Decisions and Universalizability", *Mind* vol. LXXIX, july no. 315, pp. 425–236.
Lee, Keekok (1985). *A new Basis for Moral Philosophy*, London, Boston, Melbourne and Henley: Routledge& Kegan Paul.
Lennon, K. (1984). "Moral Thinking. By R. M. Hare", *Mind* vol. XCIII, no. 372, oct pp. 617–619.
Levin, Michael E. (1979). "The Universalizability of Moral Judgements Revisited", *Mind* vol. LXXXVIII, jan no. 349, pp. 115–119.
Lewis, H. D. (1976)., ed., *Contemporary British Philosophy*, Allan & Unwin.
Locke, Don (1968). "The Trivializability of Universalizability", *The Philosophical Review*, vol. LXXVII, no. 1 pp. 25–44.
Locke, Don (1981). "The Principle of Equal Interests", *The Philosophical Review*, XC, no. 4, October, pp. 531–559.
Luper-Foy, S. (1990). "The Anatomy of Aggression", *American Philosophical Quarterly*, vol. 27, no. 3, pp. 213–224.
Lycan, W. Gregory (1969). "Hare, Singer and Gewirth on Universalizability", *Philosophical Quarterly*. vol. 19, pp. 135–144.
Lycan, William G. (1985). "Abortion and Civil Rights of Machines" in Potter and Timmons (eds.) (1985, pp. 139–156).
MacIntyre, Alasdair (1957). "What Morality is not", *Philosophia* vol. XXXII, no. 123, pp. 325–35, also in (ed.) Wallace & Walker (1970).
Mackay, A. F. (1986). "Extended Sympathy and Interpersonal Utility Comparisons", *Journal of Philosophy*, vol. LXXXIII, no. 6, june pp. 305–322.
Mackie, J. L. (1981). *Ethics: Inventing Right and Wrong*, Middlesex: Penguin books, Ltd. (originally. 1977).
Mackie, J. L. (1985). *Persons and Values: Selected Papers, Volume II*, Oxford: Clarendon Press.
MacLean, A. (1984). "What Morality is", *Philosophy* 59, no. 227 pp. 21–37.
MacNiven, C. D. (1972). "Strong and Weak Descriptivism in Ethics", *Mind*, vol. LXXXI, no. 322, pp. 161–178.
Madell, Geoffrey (1965). "Hare's Prescriptivism", *Analysis*, vol. 26, no. 2, pp. 37–41.
Magee, Bryan (1978). *Men of Ideas; Some creators of contemporary philosoph*, London: British Broadcasting Corporation (includes longer interview with R M Hare).
Margolis, Joseph (1971). *Values and Conduct*, Oxford: Clarendon Press.
Mayo, Bernard (1986). *The Philosophy of Right and Wrong*, London, New York: Routledge and Keagan Paul.
McCloskey, H. J. (1979). "Universalized Prescriptivism and Utilitarianism: Hare's Attempted Forced Marriage", *The Journal of Value Inquiry*, 13, pp. 63–76.
McDermott, Michael (1983). "Hare's Argument for Utilitarianism", *The Philosophical Quarterly* vol. 33, no. 133, pp. 386–391.
McDowell, John (1985). "Values and Secondary Qualities" in T. Honderich (ed.) *Morality and Objectivity*, London Boston: Routledge & Kegan Paul.
McDowell, John (1986). "Singular Thought and the Extent of inner Space" in Pettit, P & McDowell, J (1986). *Subject, Thought, And Context*, Oxford: Clarendon Press.
McGray, James W. (1986–87). "From Universal Prescriptivism To Utilitarianism", *Philosophy Reaserch. Archives*. vol. XII, pp. 85–87.
McGuire, M. C. (1961). "Can I do what I think I ought not to do? Where has Hare gone wrong?", *Mind*, pp. 400–404.

McNaughton, David (1988). *Moral Vision. An Introduction to Ethics*, Oxford: Basil Blackwell.
McNeilly, F. S. (1972). "Promises De-Moralized", *Philosophical Review*, vol. LXXXI, pp. 63–81.
Milkman, Kenneth Alan (1982). "Hare, Universalizability, and the Problem of Relevant Descriptions", *Canadian Journal of Philosophy*, vol. XII, no. 1, pp. 19–32.
Millican, Peter (1983). "Moral Thinking. By R. M. Hare", *The Philosophical Quarterly* vol. 33, no. 131 pp. 207–211.
Monro, D. H. (1967). *Empiricism and Ethics*, Cambridge: Cambridge University Press.
Morris, Edward Wm. (1966). "Professor Sen and Hare's Rule", *Philosophy* 41, pp. 357–358.
Moore, G. E. (1903). *Principia Ethica*, Cambridge University Press.
Mortimore, Geoffrey (1971). *Weakness of Will*, London and Basingstoke: MacMillan ST Martin's Press.
Nagel, Thomas (1988). "The Foundations of Impartiality" in Seanor, Douglas & Fotion, N. (1988).
Najder, Z. (1975). *Values and Evaluatives*, Oxford: Clarendon Press.
Nakhnikian, G. (1963). "On the Naturalistic Fallacy", see Castañeda, H-N & Nakhnikian, G (1963).
Narveson, Jan (1978). "Liberalism, Utilitarianism and Fanaticism: R. M. Hare Defended", *Ethics* 88.
Narveson, Jan (1985). "How and Why of Universalizability" in (ed.) Potter & Timmons (1985).
Nesbitt, Winston (1973). "Value-Judgements, Prescriptive Language and Imperatives", *Philosophical Quarterly*, vol. 23, no. 92.
Nickel, James W. (1973). "Functional Evaluations", *Philosophical studies* 24, pp. 57–61.
Nielsen, Kai (1985). "Universalizability and Impartiality" in ed. Potter & Timmons (1985).
Nordenstam, Tore (1966). "On Austin's theory of Speech Acts", *Mind* 75, pp. 141–43.
Norton, David L. (1977). "Can Fanaticism be distinguished from Moral Idealism", *Review of Metaphysic* 30, pp. 497–507.
Norton, David L. (1980). "On An Internal Disparity In Universalizability-Criterion Formulations", *Review of Metaphysics*, vol. XXXIII, no. 3, March, pp. 519–526.
Nowell-Smith, P. H. (1954). *Ethics*, London: Pelican books.
O'Neill, Onora (1989). *Constructions of Reason; Explorations of Kant's Practical Philosophy*, Cambridge University Press.
O'Neill, Onora (1985). "Consistency in Action" in Potter & Timmons (1985).
Paden, Roger (1988). "Hare's Ethical Formalism", *Journal of Value Inquiry* 22, pp. 223–233.
Parfit, Derek (1984). *Reasons and Persons*, Oxford: Clarendon Press.
Parkinson, H. H. R. (1982). *Theory of Meaning*, Oxford: Oxford University Press 1968.
Paul, Ellen Frankel; Miller, Fred D. & Paul, Jeffrey (1990). *Foundations of Moral and Political Philosophy*, Oxford: Blackwell.
Peetz, Vera (1972). "'Propositional Content'", *Analysis* vol. 32, no. 6. pp. 183–186.
Peetz, Vera (1981). "Imperative inference – An Addendum", *Analysis* vol. 41, no. 1 Jan pp. 54–55.
Perry, Reginald C. (1987). "Some Flaws in Hare's Moral Theory", *Journal of Value Inquiry* 21, pp. 161–167.
Perry, Thomas D. (1976). *Moral Reasoning and Truth; An essay in philosophy and Jurisprudence*, Oxford: Clarendon Press.

Persson, Ingmar (1980). *Reasons and Reason-Governed Actions*, Doc dissertation, University of Lund.
Persson, Ingmar (1983). "Hare on Universal Prescriptivism and Utilitarianism", *Analysis*, vol. 43, no. 1, pp. 43–49.
Persson, Ingmar (1988). "Rationality and Maximization of Satisfaction", *Noûs* 22, pp. 537–554.
Persson, Ingmar (1989). "Universalizability and the Summing of Desires", *Theoria* vol. LV, part 3, pp. 159–170.
Pettit, Philip and McDowell, John, (eds.), (1986). *Subject, Thought, and Context*, Oxford: Clarendon Press.
Pettit, Philip (1987). "Universalizability without Utilitarianism", *Mind*, 96, pp. 74–82.
Pettit, Philip & Smith, M (1990). "Backgrounding Desire", *The Philosophical Review*, oct, vol. XCIX, no. 4, October, pp. 565–592.
Pettit, Philip; Sylvan, Richard & Norman, Jean (1987)., *Metaphysics & Morality* Oxford: Basil Blackwell.
Pollock, Lansing (1972). "Formal Moral Arguments", *The Personalist* vol. 53, pp. 25–42.
Pollock, Lansing (1973). "Freedom and Universalizability", *Mind* vol. LXXXII.
Popper, K. (1980). *The Logic of Scientific Discovery*, London, Melbourne: Hutchinson (originally, 1959).
Potter, Nelson T. and Timmons, Mark (1985). *Morality and Universality*, Dordrecht/Boston: D. Reidel Publishing company.
Putnam, Hilary (1978). "Meaning and Reference" in Irving M Copi & James A Gould (eds.): *Contemporary Philosophical Logic*, New York, St. Martin's Press.
Putnam, Hilary (1981). *Reasons, Truth and History*, Cambridge: Cambridge University Press.
Rabinowicz, Wlodzimierz (1979). *Universalizability. A study in Morals and Metaphysics*, Dordrecht/Boston/ London: D.Reidel Publishing Company.
Rabinowicz, Wlodzimierz (1985). "The universalizability Dilemma" in Potter, N & Timmons, M, (eds.), (1985), pp. 75–90, D. Reidel Publishing Company.
Rabinowicz, Wlodzimierz (1989). "Hare on Prudence", *Theoria*, vol. LV, part 3, pp. 145–152.
Railton, Peter (1990). "Naturalism and Prescriptivity" in Paul *et al* (1990).
Rawls, John (1971). *A theory of Justice*, Oxford, Melbourne, Cape Town: Oxford University Press.
Raz, J. (1978)., ed., *Practical Reasoning*, Oxford Readings in Philosophy, Oxford University Press.
Reed, T. M. (1969). "The Implications of Prescriptivism", *Philosophical Quarterly* 19, pp. 348–351.
Reed, T. M. (1979). "The Poverty of Prescriptivism", *The Personalist*, july, pp. 243–252.
Rees, D. A. (1970–71). "Some Problems of Universalisation", *Aristotelian Society* 71, pp. 243–57.
Rennie, M. K. (1968). "On Hare's 'Better'", *Noûs*, vol. II, no. 1, pp. 75–79.
Richards, D. A. J. (1988). "Prescriptivism, Constructivism, and Rights" in *HC*, pp. 113–128.
Robinson, H. M. (1982). "Is Hare a Naturalist?", *The Philosophical Review*, vol. XCI, no. 1, jan, pp. 73–86.
Robinson, Michael H. (1975). "Hare's Golden Rule Argument: A Reply to Silverstein", *Mind* 81, pp. 578–581.
Rohatyn, Dennis (1987). *The Reluctant Naturalist; A study of G.E. Moore's Principia Ethica*, Lanham, New York: University Press of America.
Ross, Alf (1964). "On Moral Reasoning (In criticism of R.M. Hare)", *Danish Year book of Philosophy*, Copenhagen: Munksgaard, vol. 1, pp. 120–132.
Roxbee Cox, J. W. (1986). "From Universal Prescriptivism to Utilitarianism", *The Philosophical Quarterly*. vol. 36, jan no. 142, pp. 1–15.
Russel, Bertrand (1905). "On Denoting", *Mind* 14, pp. 479–493.
Ryding, Erik (1975). *Påståenden och Uppmaningar*, Lund: Doxa Småskrifter.
Ryding, Erik (1988). "Review of C.L. Hamblin: Imperatives", *Theoria* vol. LIV, part I.
Ryle, Gilbert (1949). *The Concept of Mind*, New York: Barnes & Noble Books.

Ryle, Gilbert (1961). "Use, Usage and Meaning", *Proceedings of the Aristotelian Society*, Supp. vol. 35, pp. 223–242. Reprinted in (ed.) Parkinson (1968).
Ryle, Gilbert (ed.) (1976). *Contemporary Aspects of Philosophy*, Stocksfield London: Oriel Press.
Rynin, David (1962). "Non-cognitive Synonymy and the Definability of 'good'" in, (eds.), B. H. Kazemier/ D. Vuysje *Logic and Language; Studies dedicated to professor Rudolf Carnap on the occasion of his seventieth Birthday*, Dordrecht-Holland: D. Reidel Publ. Company.
Sandøe, Peter (1988). *Moral Knowledge – Assessment of a Perceptual Paradigm* D.Phil thesis, Oxford University).
Sandøe, Peter (1989a). "Amoralism – on the Limits of Moral Thinking", *Theoria* vol. LV, part 3, pp. 191–204.
Sandøe, Peter (1989b). "Review of Douglas Seanor & N. Fotion (eds.): *Hare and Critics, Essays on Moral Thinking, Theoria*", vol. LV, part 3, pp. 211–223.
Schiffer, S. R. (1972). *Meaning*, Oxford: Clarendon Press.
Schrag, Brian Eugene (1975). *Universalizability and the concept of morality*, (Dr dissertation) University Microfilms International, Ann Arbor, Michigan.
Schueler, G. J. (1984). "Some Reasoning about Preferences", *Ethics* 95.
Schüller, B. (1986). *Wholly Human; Essays on the Language of Morality*, Dublin Georgetown University Press, Washington D. C.: Gill and MacMillan.
Seanor, Douglas & Fotion, N. (1988). *Hare and Critics; Essays on Moral Thinking*, Clarendon Press, Oxford.
Searle, John R. (1969). *Speech Acts; An Essay in the Philosophy of Language* Cambridge, London: Cambridge University Press.
Searle, J. R., (ed.), (1971). *The Philosophy of Language*, Oxford readings in philosophy.
Searle, John R. (1979). *Expression and Meaning; Studies in the Theory of Speech acts*, Cambridge, London: Cambridge University Press.
Searle, John R. (1983). "Deriving 'ought' from 'is': Objections and Replies" in Hudson (1983).
Sellars, W. (1963). "Imperatives, Intentions, and the Logic of 'ought'" in Castañeda *et al* (1963).
Sen, Amartya (1966). "Hume's Law and Hare's Rule", *Philosophy* 41, pp. 75–79
Sen, Amartya & Williams, Bernard (1982). *Utilitarianism and Beyond*, Cambridge, London: Cambridge University Press.
Shadish, E. (1982). "Universal Prescriptivism and the Spirit of Morality", *Journal of Value Inquiry*, vol. 16, pp. 223–234.
Shwayder, D. S. (1982). "Uses of Language and Uses of Words" in Parkinson (1982).
Sidgwick, H. (1901). *The Methods of Ethics*, Hackett Publ. Company, seventh edition.
Silverstein, Harry S. (1972). "A Note on Hare on Imagining Oneself in the place of others", *Mind* 81, pp. 448–50.
Silverstein, Harry S. (1974). "Universality and treating Persons as Persons", *The Journal of Philososphy*, vol. LXXI, no. 3, feb 14 pp. 57–71.
Silverstein, Harry S. (1983). "Assenting to 'Ought' Judgements", *Noûs*, vol. XVII, no. 2, pp. 159–182.
Singer, Marcus G. (1961). *Generalization in Ethics*, New York: Alfred A. Knopf.
Singer, Marcus G. (1963). "The Golden Rule", *Philosophy* 38, 293–314.
Singer, Marcus G. (1970). "Freedom from Reason", *Philosophical Review* pp. 253–259.
Singer, Marcus G. (1985). "Universalizability and the Generalization Principle" in (eds.) Potter & Timmons (1985).
Singer, Peter (1974). *Democracy and Disobedience*, New York and London: Oxford University Press.
Singer, Peter (1988). "Reasoning towards Utilitarianism" in (eds.) Seanor and Fotion (1988).
Sloman, Aaron (1969). "Transformation of Illocutionary Acts", *Analysis* 30 pp. 56–59.
Slote, Michael (1983). *Goods and Virtues*, New York: Oxford University Press.
Smart, J. J. C. (1984). *Ethics, Persuasion and Truth*, London, Boston: Routledge & Kegan Paul.
Snare Frank (1972). "Wants and Reasons", *The Personalist*, aut 1972, pp. 395–407.

Sosa, Ernest (1966). "On Practical Inferences and the Logic of Imperatives", *Theoria* vol. XXXII, part 3, pp. 211–223.
Spector, Horacio (1987). "Dale on Supervenience", *Mind* vol. XCVI, no. 381 pp. 93–94.
Spitzley, Thomas (1990). "Davidson and Hare on Evaluations", *Ratio* III, pp. 48–63.
Spitzley, Thomas (1989). "Prima Facie versus Critical Moral Principles. Critical Comments on R. M. Hare's 'Moral Thinking'", *Ratio* (New Series) II June pp. 63–74.
Stalley, R. F. (1972). "Intentions, Beliefs, and Imperative Logic", *Mind* vol. LXXX, no. 321 Januari.
Stenius, E. (1960). *Wittgenstein's Tractatus: A critical Exposition of its main Lines of Thought*, Ithaca, Cornell University Press.
Stevenson, C. L. (1944). *Ethics and Language*, Yale University Press.
Steward, H. (1990). "Identity Statements and the Necessary a Posteriori", *Journal of Philosophy* vol. LXXXVII, no. 8, pp. 385–389.
Strawson, P. F. (1964). "Intention and Convention in Speech Acts", *Philosophical Review*, vol. LXXIII, no. 4, pp. 439–60, repr. in Searle (1971).
Sumner, L. W. (1967). "Hare's Argument against Ethical Naturalism", *The Journal of Philosophy*, vol. LXIV, no. 23, pp. 779–791.
Taylor, C. C. W. (1965). "Critical Notice: *Freedom And Reason* by R. M. Hare", *Mind*, no. 74 pp. 280–298.
Taylor, C. C. W. (1980). "Plato, Hare and Davidson on Akrasia", *Mind* vol. LXXXIX, pp. 499–518.
Thau, Stewart (1972). "The Distinction between Rhetic and Illocutionary acts", *Analysis*, vol. 32, No. 6, pp. 177–183.
Thomson, Garreth (1987). *Needs*, London & New York, Routledge & Kegan Paul.
Thomson, Judith Jarvis (1990). "The no good Reason Thesis" in Paul, E. F., Miller, F. D., Jr. and Paul, J. (eds.) (1990).
Torrance, Stephen B. (1981). "Prescriptivism and Incompleteness", *Mind* vol. XC, pp. 580–585.
Toulmin, S. E.(1950). *An Examination of the Place of Reason in Ethics*, Cambridge.
Trigg, Roger (1973). *Reason and Commitment*, Cambridge University Press.
Turnbull, R. G. (1954). "A note on Mr Hares's Logic of Imperatives", *Philosophical Studies* vol. V, no. 3.
Tännsjö, Torbjörn (1974). *The Relevance of Metaethics to Ethics*, Doctoral thesis, Stockholms Universitet.
Urmson, J. O. (1968). *The Emotive Theory of Ethics*, Hutchinson University Library, London.
Vendler, Zeno (1976). "A Note to the Paralogisms" in Ryle, G. (ed.) (1976).
Vendler, Zeno (1988). "Changing Places" in (ed.) Seanor, D & Fotion, N (1988), pp. 171–183.
Versényi, Laszlo (1973). "Prescription and Universalizability", *Journal of Value Inquiry* 6, vol. 32, no. 4, pp. 22–36.
Vesey, G. N. A. (1965). "Hare on Descriptive Meaning-Rules", *Mind*, no. 296 vol. LXXIV, pp. 590–591.
von Wright, George Henrik (1977). *Norm and Action; A Logical Enquiry*, London, Henley: Routledge & Kegan Paul (1963).
Walker, Jeremy (1973). "A Naturalist Reply to Hare", *Philsophical Studies*, 24, pp. 45–51.
Wallace, G. J. & Walker, A. D. M. (1970)., (eds.) *The Defintion of Morality* London: Methuen & Co Ltd.
Wallace, James D. (1988). *Moral Relevance and Moral Conflict*, Ithaca and London: Cornell University Press.
Ward, Andrew (1973). "Morality and the Thesis of Universalizability", *Mind*, vol. LXXXII, pp. 289–291.
Warnock, G. J. (1967). *Contemporary Moral Philosophy*, London: MacMillan.
Warnock, G. J. (1974). "Contemporary Moral philosophy" in Hudson (1974).
Warnock, G. J. (1976). "Imperatives and Meaning" in Lewis (1976).
Wedar, Sven (1985). *Realism and Validity; Studies in the legal theory of Alf Ross*, Lund monographs in Practical Philosophy, (ed.) Hermerén, G.

Wellman, C. (1961). *The Language of Ethics*, Cambridge Massachusetts: Harward University Press.
Werner, Richard (1976). "Hare on Abortion", *Analysis*, vol. 36, no. 4, pp. 177–181.
Wetterström, Thomas (1977). *Intention and Communication*, Lund: Doxa.
Wetterström, Thomas (1986). *Towards a Theory of Basic Ethics*, Oxford: Doxa.
Wetterström, Thomas (1989). "On intuitions", *Theoria* vol. LV, part 3, pp. 178–185.
White, Harold J. (1969). "An Analysis of Hare's Application of the Thesis of universalizability in his moral Arguments", *Australian Journal of Philosophy*, vol. 47, no. 2 aug.
Whiteley, C. H. (1966). "Universalizability", *Analysis* vol. 27, no. 2 pp. 45–49.
Wiggings, D. (1978/79). "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire", *Proceedings of the Aristotelian Society* vol. 79.
Wiggins, D. (1987). *Needs, Values, Truth*, Oxford: Aristotelian Society Series, vol. 6, Basil Blackwell.
Williams, B. (1980). "Internal and External Reasons", Ross Harrison (ed.) *Rational Action*, Cambridge: Cambridge University Press. Also in Williams (1981).
Williams, B. (1981). *Moral Luck*, Cambridge, London: Cambridge University Press.
Williams, B. (1985). *Ethics and the Limits of Philosophy*, London: Fontana Press/ Collons.
Williams, B. A. O. (1973). in *Utilitarianism: For and against*, ed., Williams, B A O & Smart, J. J. C, Cambridge University Press.
Winch, Peter (1965). "The Universalizability of Moral Judgements", *Monist*, vol. 49 (also in Winch 1972).
Winch, Peter (1972). *Ethics and Action*, London: Routledge & Kegan Paul.
Wolff, Jonathan (1989). "Morality For Archangels", *Philosophical Books*, vol. 30, oct no. 4, pp. 201–209.
Zimmerman, David (1980). "Force and Sense", *Mind* 89, pp. 214–233.

STUDIES IN PHILOSOPHY
Editors: Bengt Hansson and Göran Hermerén

1 Per Bauhn, *Ethical Aspects of Political Terrorism. The Sacrificing of the Innocent*, 1989, 184 pp.
2 Thomas Anderberg, *Suicide: Definitions, Causes, and Values*, 1989, 321 pp.
3 Ingrid Petersson, *Four Theories of Responsibility*, 1990, 107 pp.
4 Dan Egonsson, *Interests, Utilitarianism and Moral Standing*, 1990, 206 pp.
5 Urban Forell, *Gerechtigkeit, Pflicht und etische Beurteilung*, 1993, 122 pp.
6 Toni Rønnow-Rasmussen, *Logic, Facts and Representation. An Examination of R. M. Hare's Moral Philosophy*, 1993, 247 pp.

Toni Rønnow-Rasmussen

Logic, Facts and Representation

An Examination of R. M. Hare's Moral Philosophy

The English philosopher Richard M. Hare – one of the leading moral philosophers of the twentieth century – develops in his most recent major work an ingenious theory of moral reasoning. He maintains that there are rules of fundamental moral thinking which are determined by the concepts involved in moral thinking. Moreover, he argues that the evaluative conclusions reached by taking facts into account and by reasoning in accordance with these rules, correspond to those reached by a utilitarian.

The author of this work sets out to examine this "derivation" of utilitarianism. Thus, a large part of this essay examines Hare's view on the meaning of 'ought'. Are evaluative 'ought'-judgements, as Hare claims, universal(izable) prescriptions? Another central topic is Hare's view on what follows from imagining being in someone else's shoes.

Hare's claim that his theory rests firmly on a logical ground is questioned. His view on prescriptive meaning, it is argued, stands in need of revision; the universalizability-thesis is not a logical thesis, and there are, it is maintained, unanswered questions concerning his view on representation.

Lund University Press
Box 141, S-221 00 Lund, Sweden
Art nr 20283
ISSN 1100-4290
ISBN 91-7966-228-5

Chartwell-Bratt Ltd
Old Orchard, Bickley Road,
Bromley, Kent BR1 2NE,
England
ISBN 0-86238-324-4