RATIONAL CHOICE, DELIBERATION, AND DETERMINISM

RATIONAL CHOICE, DELIBERATION, AND DETERMINISM

By

ROBERT L. GRAY, JR., B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

November, 1973

MASter of arts (1973)
(Philosophy)

McMASTER UNIVERSITY
Hamilton, Ontario.

TITLE: Rational Choice, Deliberation, and Determinism

AUTHOR: Robert L. Gray, Jr., B.A. (Marietta College)

SUPERVISOR: Professor S. Najm

NUMBER OF PAGES: v, 152

CONTENTS: In this thesis I attempt to show that, contrary to the arguments offered by Carl Ginet and Richard Taylor, there is nothing in the concepts of rational choice and deliberation which would make them logically inconsistent with determinism. To this end, I begin by offering my own analyses of these and related concepts. Except for the use of criteria and standards in place of wants and desires, my analysis of rational choice is quite similar to the standard conception. In chapter three I attempt to refute Hume's theory of causation and offer my own alternative, the principle of causal entailment. In chapter four, I attempt to point out some of the implications of this view of causation for determinism in human affairs. In chapter five, I attempt to refute the arguments offered by Ginet and Taylor.

ACKNOWLEDGEMENTS

For their suggestions and criticisms, but most of all for their time, I should like to acknowledge the help and assistance of the following persons: Dr. S. Najm, Dr. E. Simpson, and Dr. M. Radner, my first, second, and third readers, respectively.

They will judge whether or not their time and effort have been well spent.

Thanks, too, to my wife, Bonnie, and my son, John, but especially to John--who thought my time ill spent.

CONTENTS

INTRODUCTION

Whether or not men have free will is, as it stands, a fundamentally unimportant and obscure question. What is a will? Do men have will, free or otherwise? What is a free will and what, for that matter, is an unfree will? And, when you come right down to it, what does it matter? Our common morality is not based upon any abstract philosophic notion such as that of the will. But our common morality is based, I think, in a large part upon the notion of choice. We might, then, want to ask: Do men have free choice? This question is redundant. If one does not have free choice, one does not have freedom to choose and, if one does not have freedom to choose, one does not choose. Freedom without choice is meaningless; choice without freedom is impossible. So the question whether or not the concepts of choice and determinism are logically consistent is really the old question whether or not freedom and determinism are consistent. It is with this old question in its new dress that I am concerned in this thesis.

The "new dress" is that which has recently been given to the question by Carl Ginet and Richard Taylor. These men attempt to show that the concepts of choice and determin-

ism are logically inconsistent. Accordingly, I shall attempt to show in this thesis that, contrary to their arguments, there is nothing in the concepts of rational choice and deliberation which would make them logically inconsistent with determinism. But how does one show this? Certainly, it is not sufficient just to refute their arguments. One must show, too, that their conclusions are wrong. In this thesis, I have attempted to do both. Whether or not I have succeeded in either endeavor is not for me to say.

The thesis may be divided for purposes of discussion into three parts. In the first two parts I have attempted to provide some general understanding of the key concepts which the arguments of Vinet and Taylor involve: these are the concepts of "choice" and "determinism". In chapters one and two I discuss the concept of choice, but, since most philosophers appear willing to grant that men's irrational choices may be determined, I have directed my attention only to the concept of rational choice and have attempted to avoid the more fundamental and, perhaps, more difficult question of what it is to make a choice. My emphasis in this chapter is on the concept of a reason for choosing: What is a reason and what is it to choose in accordance with reasons? This discussion is carried over into chapter two where I discuss one of the two components of reasons (criteria) in more detail. In chapter two I identify the process of rational evaluation with that of rational choice. If, then, rational choice is consistent

with determinism, so is rational evaluation. Since many, if not most of the objections to determinism have had a moral or ethical basis, I have attempted in this chapter to point out some of the implications of this view of rational choice and evaluation for ethics and theory of value.

In chapters three and four, I discuss the concepts of causation and determinism. In chapter three I attempt to show that our traditional conception of causation (taking Hume's view as a paradigm of that tradition) is inadequate and I suggest a different conception which I call the "principle of causal entailment." This chapter is primarily concerned to offer a refutation of Hume and to elucidate some of the implications of the principle of causal entailment. In chapter four I take up several different, although related matters. First, I examine the implications of the principle of causal entailment for the view that human activities are determined. In the second part of this chapter I attempt, first, to expand what I think is the major argument for the view that human behavior is determined. This I shall call the "statistical argument." Next, I try to provide an account, based essentially on the principles of operant psychology, of the form which such determination might take, attempting along the way to answer some of the more obvious objections to determinism. Finally, I attempt to point out and clarify what I think has been and continues to be a major misconception of what determinism asserts on the part of those who

argue against it.

These four chapters cover a lot of ground and it is difficult to bring them together. This I attempt primarily in the course of chapter five. Here I offer refutations of Ginet's and Taylor's arguments. Taylor has offered other arguments beyond the fact that choice and determinism are inconsistent, and I attempt to show that, given my conception of rational choice and determinism, these arguments need not be accepted.

The reader will be disappointed if he expects to find any discussion of freedom in this thesis. I am concerned only to show that rational choice is consistent with determinism (although implicit in my arguments is, I think, the view that rational choice entails determinism). If men do make rational choices, then it is obvious that, at least to the extent they do, they are free. I do not, however, pretend to know that they do make rational choices or, for that matter, that determinism is true, but I should like to point out at this time the fact that the opposite of determinism is not freedom: it is indeterminism. And it would require some argument to show that, if determinism were true, a man might choose rationally or act freely.

I

RATIONAL CHOICE

About what it is to make a rational choice, there would seem to be little general disagreement. A rational choice is a choice made on the basis of or in accordance with reasons. What this means, however, is not altogether clear, since it would seem that an irrational choice could also be made in accordance with reasons. It might, therefore, be better to say that a rational choice is a choice made in accordance with the weight of reasons. For instance, where the reasons in favor of an alternative A outweigh those in favor of B, the rational choice would be A. To choose B, despite the fact that there are reasons for so choosing, would be irrational. A rational choice, then, is a choice made in accordance with the weight of reasons; an irrational choice is a choice which is contrary to the weight of reasons. A non-rational choice, then, is what is left, either a choice made in the absence of reasons or a choice made where the reasons favoring one alternative are no stronger than those favoring another. For instance, Buridan's ass, if forced to choose between hay and water when he is equally hungry and thirsty, could not make a rational choice between the two. The reasons in favor of the one alternative are equally as strong as those in favor of the other. It would be irrational,

however, for him not to choose one. In a similar situation, a man might choose by tossing a coin. Choosing arbitrarily in this way would, given the situation, be the rational thing to do, but the choice executed in this manner would be non-rational.

REASONABLE AND RATIONAL CHOICES

At this point it would be well to distinguish "reasonable" and "rational" choices. This distinction relies, I think, upon another distinction in the senses of the word "choice". Sometimes we use the word "choice" to refer to or to describe the object of choice; on the other hand, we sometimes use it to refer to the act or process of choosing. Now, bearing this in mind, it is clear that when we say "rational choice" we may mean two different things. On the one hand, we may mean that the object chosen is rational, and on the other, we may mean that the act or process of choosing is rational. Now, where "rational choice" refers to the object, I shall use "reasonable choice"; where it refers to the act or the process of choosing, "rational choice."[1] A "reasonable choice" may be defined as that object which a rational person would choose if he possessed complete and accurate knowledge

---

1 David Gauthier (Practical Reasoning. London: Oxford University Press, 1963, p. 2n) suggests "reasonable" and "reasoned". I prefer "rational" for "reasoned" although neither term is, probably, better than the other.

about the objects of choice, the situation, and himself. In other words, where a person possessed complete and accurate knowledge of these things, his rational choices would all be reasonable choices. The process of rational choice, that is, of choosing rationally, may be viewed, then, as an attempt to make a reasonable choice. The attempt may or may not be successful. Since "reasonable" or "unreasonable" describe the object (as an object which it is reasonable or unreasonable to choose), "rational choice" describes the process. Accordingly, whether or not a choice is rational will depend on how the choice is made, and we may say that a choice is rational if it is made on the basis of or in accordance with reasons (bearing in mind that this means "in accordance with the weight of reasons").

The above definition, however, is not altogether clear. To make it clear we must know (1) what constitutes a reason and (2) what it is to make a choice in accordance with reasons. I shall take up (1) first.

WHAT IS A REASON?

A reason for or against a given alternative has, in the context of rational choice, two components, a fact and a feeling (e.g. a want, a desire, etc.). A reason must express some fact or set of facts about the object such as that it is such and such an object or that it has such and such properties. That this is so is too obvious to require argument. Were

It not so, the choice would not be between the objects of choice. That an object is such and such an object or that it has such and such properies is not, by itself, a reason for or against choosing it. This is obvious from the following example. Suppose I am to choose between two purses, one containing ten dollars and the other containing one thousand dollars. I can have either one, no strings attached, whichever I choose. It seems to be obvious that I should choose the purse containing one thousand dollars. Why? because it contains the more dollars. But, suppose I do not know what dollars are or what they are for or that they can be of any use to me. Then, knowing that the one purse contains one thousand and the other ten does not tell me which to choose. Suppose I am informed that with one thousand dollars I could buy a car. What is a car, I ask? But suppose it is then patiently explained to me that dollars are a means of exchange , that with dollars I can obtain whatever I want, and the more dollars I have, the more I can obtain. If I understand all this, I will choose the purse containing one thousand dollars without further ado. I understand that dollars can get me things that I <u>want</u> or <u>like</u>. If I did not want or like dollars or if I could not use them to obtain things that I wanted or liked, the mere fact that this purse contains one thousand dollars would not be a reason for choosing it. A reason, then, for choosing a particular object is a belief that it is an

object of a certain sort or that it has some property or set of properties together with the fact that I want or like such an object.

Now, this analysis of "reasons" is, I think, sound enough as far as it goes and is, furthermore, the generally-accepted view of what constitutes a reason for or against a particular choice.[2] However, it does not, in my opinion, go far enough. The concepts of "wanting", "desiring", and "liking" are just too narrow to extend this view of reasons to all cases without a lot of unnecessary trouble.[3] Accordingly, I shall employ a different notion, that of a criterion or standard of choice. Discussion of the nature of criteria and standards will be, for the most part, postponed until chapter two. For the moment I shall only offer a few examples to indicate what I mean.

CRITERIA AND STANDARDS

Suppose I wish to buy an automobile, but that I have never bought one before and do not know anything about them. However, I want to buy a good one. Unfortunately, I do not know how to tell a good one from a bad one. I ask a friend,

---

2 For instance, see Davidson, Donald. "Actions, Reasons, and Causes", Journal of Philosophy, LX, No. 23 (Nov. 7, 1963) pp. 685-700.

3 A similar opinion has been express a bit more strongly by Stuart Hampshire (Thought and Action. London: Chatto and Windus, 1959, pp. 167-8). It should be noted that nearly all writers who accept this view of reasons use fairly simple examples. Again, Davidson (ibid.) is a good example.

who is an auto enthusiast, for advice. He tells me that a good automobile has such and such propeties, e.g., a solid body, a comfortable "ride", an engine of a certain kind, and so on. Assuming I respect his opinion, I may adopt these things as my criteria, and when I go about choosing an automobile I will try to choose that automobile which best fits them. What I want, however, is not "a solid body," but, rather, a good car. My criterion tells me that a good car is one with a solid body, among other things. Of course, in this case, I do, in a sense, have a "pro-attitude" toward my criterion.[4] I say "in a sense" because my pro-attitude is not actually toward my criterion or, for that matter, toward that which fits my criterion. For instance, an apple sorter in a produce market sorts apples according to criteria or standards of what constitutes a good apple and what constitutes a bad. If the apples were to be sorted into three categories such as "eating," "cooking," and "inedible", the sorter would have to choose with regard to each apple between three alternatives and would have to choose according to at least two sets of criteria (e.g. "eating" and "cooking"). Such a choice would, of course, become automatic and, perhaps, even unconscious

---

[4] In recognition of the fact that the "want-desire" vocabulary is too narrow to cover all choice and the kinds or degrees of feelings involved, P.H. Nowell-Smith (Ethics. Penguin Books, 1954, pp. 111-12) has suggested the terms "pro" and "con" attitude. There is, obviously, no universally accepted terminology.

after a while; nevertheless, it would be obvious from an examination of the apples after they had been sorted that they had been sorted according to some fairly specific standard or criterion. The apple sorter, however, might not want cooking apples or eating apples; he might hate apples; he might just not care. His pro-attitude or his want or his feeling that, together with the facts about the apples, constitutes a reason for sorting them in this way enters at an entirely different level--his desire, for instance, to keep his job. Here, if the apple sorter is asked why he put the apples he did in the eating category, he will reply, perhaps, that all those apples were large, round, and uniformly red. If asked why these facts about the apples constituted a reason for placing them in this category, he will reply that the criterion or standard he is using for eating apples specifies that they should have these properties. If asked why he is using that criterion, why he attaches any importance to it, he may reply, in short, that he wants to keep his job.

Given the notion of a criterion or standard of choice, a reason, in the context of rational choice, may be defined as a fact about the object of choice seen in light of some standard or criterion, where it is understood that some reason underlies the adoption and use of this criterion and that this reason will involve a pro-attitude of some kind. (A crucial exception to this rule will be taken up in chapter two).

AN EXAMPLE

It might be well to pause here to consider an example. Suppose a financier has received requests for sizeable loans from two entrepreneurs, call them X and Y. Now, let us suppose that he has decided to invest in one of them but his capital is such that he cannot invest in both. He must, then, choose between them. He considers the two schemes in an attempt to discover the facts about them. He discovers that if he invests in X's scheme the risk will be very great, the returns small; if he invests in Y's scheme the risk will be less and the returns greater.[5] All other things, he finds, are equal. Now, in which of these two schemes ought he to invest? What ought to be noted here is simply that the answer to this question has not been given. Knowing the relative properties of these two schemes does not by itself constitute a reason for choosing between them. If the investor were also a philanthropist it is conceivable that the properties mentioned might be totally irrelevant to him. However, if he has "minimum risk for maximum returns" as a criterion, then the relative properties of these two schemes, in light of that criterion, are reasons for choosing Y. Both the criterion and a fact fitting the criterion are necessary if the financier is to have a reason for choosing one of the investments over the other.

---

5 I am using "risk" here in a neutral, descriptive sense to mean simply the probability that money put into this venture will not be returned.

But suppose the financier had chosen by flipping a coin. To choose in this way would, of course, be irrational. But why? The answer, I think, is fairly obvious. Choosing in this way is irrational because the reasons for choosing between the two alternatives are ignored. Underlying the financier's criterion is, say, a desire for profit, for more money than he has. His criterion serves the function of allowing him to determine, on the basis of the facts available to him, which of the two alternatives is most likely to fulfill this desire (or purpose). The criterion specifies that, in choosing, he is to concern himself with such things as risk and returns, or gains and losses. Where these things differ among the alternatives, his criterion tells him that there is a reason for choosing one of the alternatives rather than the others. Since we may assume that whether the coin falls heads or tails has nothing to do with those properties of the objects which are specified by the financier's criterion to be relevant to his choice, for him to choose by flipping a coing would be irrational.

Again, if, after a careful examination of the alternatives, the financier found that each fit his criterion equally well, then, other things being equal, to choose by flipping a coin would be not irrational, but non-rational. On the other hand, if he were concerned about fairness, then, rather than arbitrarily choosing X or Y, he might decide to give each an equal chance by flipping a coin. Since there is

here a criterion (fairness) and a fact which fits it (neither investment is preferable to the other), there is a reason for choosing by flipping a coin. To choose in this or a similar way would, then, be rational.

THE COMPONENTS OF REASONS

As was stated earlier, in the context of rational choice a reason for choosing a particular alternative has two components, a fact and a standard or criterion which the fact fits. However, a chooser may not always know the facts, and he may not employ the "proper" criteria. From this, of course, it does not follow that he cannot make rational choices. Since it will help to understand the concept of a rational choice, this point should be made clear.

THE COMPONENTS OF REASONS: FACTS

We may begin by noting that a reasonable choice is defined as that which a rational person would make if he possessed complete and accurate knowledge of the alternatives, the situation, and himself. A rational choice is an attempt to make a reasonable choice. It is, then, an attempt to choose in accordance with the facts and with one's standards or criteria. To make a rational choice, then, a chooser must examine the alternatives and the situation to determine, so far as he can ascertain, what the facts are. On the basis of the information at his disposal, he will come to some beliefs

about what those facts are. These beliefs, if the choice is to be rational, must be warranted. This, I think, is fairly obvious. If a man were asked why he made a particular choice and replied "Oh, no particular reason," we would not call his choice rational. If, however, he said, "Because it was X" (where X is some fact about the object chosen) we might ask why he believed this. If, again, he said "Oh, no particular reason," we should conclude that he did not choose rationally. He did not attempt to make a reasonable choice.

A possible objection here might be that, since we can only show that a belief is warranted by giving reasons for believing it, the view that the beliefs upon which a rational choice is based must all be warranted leads to an infinite regress, in which case it would not be possible to make a rational choice. This objection, however, rests upon a false assumption. It is not necessary that I be able to give reasons for all my warranted beliefs (It would, however, be necessary for me to give reasons for all my rational beliefs). As an example, I believe I am writing at a brown desk. This belief is warranted. My reason for it is that I see a brown desk before me. However, if I am asked to justify this belief, I cannot. I can only say I am so constituted that, other things being equal, I cannot help believing that what I see before me is, in fact, before me. This fundamental belief is, we should say, non-rational.[6]

---

6 Of course, this is not to say naive realism is true;

Of course, it is possible for persons to disagree about the facts and for such disagreement to be rational. Given that this is so, it is possible for different people in the same situation to make different rational choices. However, the reasonable choice for each, if their criteria were the same, would be the same. Thus, if a person A were to choose X believing warrantedly that X had properties a, b, and c when X did in fact have those properties and that X, therefore, fit his criterion, his choice would be both rational and reasonable. If B, however, were to choose Y believing warrantedly that Y had properties a, b, and c and that X did not, his choice would be rational but not reasonable.

THE COMPONENTS OF REASONS: CRITERIA

Regarding criteria two points need be made: (1) For a choice to be rational the criteria in light of which it is made must be rationally chosen if they are chosen, and (2) it

---

rather, it is to say we all believe naive realism to be true. Russell's dictum that naive realism leads to physics which, in turn, proves naive realism false, is not completely true (Bertrand Russell. An Inquiry into Meaning and Truth. Middlesex: Penguin Books, 1962, p. 13). No naive realist, at least to my knowledge, would claim that "What I see before me is, in fact, before me" is categorically true. All would admit at least the possibility that I might be looking in a mirror. As an epistemological doctrine, naive realism should be construed as claiming "other things being equal, what I see before me is, in fact, before me." What physics does is show that other things are not equal, that there are other facts which may be perceived, perhaps only with the aid of very delicate instruments, to show that our first impression was wrong. Physics disproves the particular beliefs of naive realism, but it does not, indeed, it cannot disprove naive realism, for that is the fundamental and non-rational condi- of all rational human beliefs, including the physicist's.

is not possible that all of our criteria should be rationally chosen. The first of these two points shall be discussed in chapter two. I shall discuss the second here.

It would seem to be fairly obvious that the standards and criteria by which rational choices are made may, themselves, be chosen. I do not mean that they may be chosen in the sense that, given a choice situation, one must determine which criteria apply (although it is true that they may be chosen in this sense), but, rather, in the sense that we speak of their being adopted. If, for instance, I have decided to purchase a new automobile, I must choose one from among the many models available, but, from the very fact that there are so many varied models on the market, it is obvious that different criteria are available for choosing between them. If, as the consequence of a rational choice, I buy a Volkswagon whereas my neighbor buys a Cadillac, then, other things being equal, this difference can be explained only by a difference in the criteria in light of which our choices are made. In other words, the standards or criteria we have adopted differ.

While it may throw little light on the subject, it is interesting to note that, as a matter of linguistic convention, we do not often speak of standards or criteria as being chosen; indeed, it sounds odd to speak of choosing a choice-criterion (and this is so, I think, not simply because it sounds redundant). Instead, we speak of "adopting" such standards or criteria. This linguistic distinction may be due just to the

fact that the objects differ. When we choose between standards and criteria, we do not do anything (at least, nothing obvious) but the choice affects what we do or choose to do. When we choose a standard or criterion, we do not do it, make it, take it, or anything else of the kind; what we do is "make it our own." When we choose between objects, on the other hand (again, I am using "objects" in a very broad sense to include such things as actions or courses of action) we do one of these things; we act upon our choice. We do not act upon a criterion or standard, nor do we act upon the choice of one. The choice of a standard may affect our future choices but not by itself; it cannot effect them. For instance, if, in buying a car, I adopt the standard of "economy," the choice of this standard over, say, "luxury" would not, by itself, lead to any action to which the choice of "luxury" would not have led. In both cases, something more is needed, and that something is the examination of the available alternatives, the available models, in light of the standard, but when this is done, i.e., when it has been determined which object best fits the standard, the choice is made and this choice determines directly which action I shall perform, that is, which car I shall buy. Once the choice has been made, alternative actions are no longer available (not in a categorical sense, but in the sense that all others have been rejected). A standard, however, simply would not be a standard of choice unless alternatives were available.

When a rational choice is made between objects, the objects are considered in the light of criteria or standards, but in what light are the objects of choice to be considered when the objects are, themselves, standards or criteria? Of course, if the choice of a standard is to be rationally made, it must be made by considering the standard in the light of some other standard, but if this choice is rationally made, we may ask still how the other standard was chosen. Obviously, this will lead to an infinite regress of an impossible sort. In the process of justification it may be possible always to go on giving further reasons and more basic criteria to support the choice of some standard or criterion, but this is because justification applies to reasonable choice. Justification is the process of demonstrating that a choice is reasonable, that it is a choice for which reasons may be given. In the actual process of choice, that is, of choosing, this would not be possible, or the act of choice would never be made and the process of choosing never begun. We are faced, then, with a dilemma: either we admit that ultimately the choice of a standard is non-rational or we admit that ultimately our standards are not chosen. No doubt, there are some (e.g., the existentialists) who would hold the former of these two views, but I know of no arguments sufficient to prove them wrong, and I hold the latter. In either case, it would follow that the complete rational justification of a choice is impossible. In the first case, we must come eventually to a step in the process of choice which is non-rational and

which, consequently, admits of no further justification and, in the latter, to a point at which choice begins with a given, an unchosen criterion which cannot, qua given, be justified. Of course, it is possible in justification to get off the horns of the dilemma, but then the only alternative is an infinite regress, in which case the justification could not end and could not be completed.

From this it is clear that in a situation of choice all the factors relevant to the final outcome, to the determination of the choice, are, in a sense, given. The properties of the objects of choice are given, in the sense that the chooser does not choose what they shall be; they are matters of determinate, empirical fact and it is the chooser's function to ascertain, insofar as he is able, what those facts are, to formulate, on the basis of his warranted beliefs and his examination of the objects and the situation, some warranted beliefs about them. And, of course, in a rational choice, the criteria are also given, in the sense that, if they are rationally chosen, they follow from the chooser's basic, unchosen criteria and the chooser's beliefs.

THE ACT OF CHOICE

As was made clear earlier, an act of choice is rational if the choice is made because of or in accordance with reasons and a reason is defined as a statement of the properties of an object together with the fact that those

properties fit the criterion being used. In a situation of choice, then, when the properties of the objects have been determined, all that remains to complete the process of rational choice is the determination of which of the alternatives best fits the criteria. When this is accomplished the process is completed and the choice is made. The actual choice may or may not be a separate step in the process. We may wish to accept something like Hobbes' definition of the will as the last appetite in deliberation,[7] in which case the actual act of choice would not be a separate step at all but simply the end of the examination of the objects, or we may wish to say that the act is separate, that we first consider the objects in the light of our criteria and then choose, but, even if we say that the act is separate in this sense, we cannot say that it is independent. If it is a separate or a distinct step, it is so only in the sense that the conclusion to a valid argument is separate from the argument, separate but not independent.

This allusion to deductive argument is one which, in this case, should not, perhaps, be made. It runs a very great risk of being misleading. Nevertheless, it is instructive and, while I must stress the fact that the process of rational choice is not identical with that of deductive argument, there are analogies which, if pointed out, may serve to make the

---

7 Hobbes, Thomas, Leviathan, chapt. six.

process clearer. An argument could be set up in this way:

(1) I will choose from the available alternatives that object which best fits criterion K.

(1a) An object will fit criterion K if it has properties a, b, c, d, and not e.

(1b) An object will not fit criterion K if it has property e, regardless of its other properties.

(1c) An object best fits a criterion if more of its properties fit the criterion more extensively than those of any other object.

(2) The available alternatives are objects X, Y, Z.

(2a) Object X has properties a, b, c, and not e.

(2b) Object Y has properties a, b, and not e.

(2c) Object Z has properties a, b, c, d, and e.

(3) Object Z does not fit the criterion.

(4) More of the properties of object X fit the criterion than the properties of Y.

---

(5) Therefore, object X best fits the criterion.

---

(6) Therefore, I will choose X.

Since a choice is reasonable if it is in accordance with the facts and one's criteria and since a rational choice is defined as an attempt to make a reasonable choice, it is clear that the complete statement of the process of rational choice would include (1) an enumeration of all the properties of the objects of choice which were considered (premise two),

(2) an enumeration of the criteria according to which those properties were considered (premiss one), and (3) a statement of how those properties fit the criteria (premisses three and four).[8] When these three steps are completed, the conclusion, in this case that X best fits the criterion, is given, that is to say, the conclusion is entailed and to state it is simply to make explicit what has already been said in a different and perhaps less clear way. The same is true of the final conclusion which serves merely to make explicit the implications of what has already been said. Again, it must be stressed that the actual process is not identical with a deductive argument. It is doubtful that, even at our rational best, we ever take the pains to make all of the steps in the process explicit, but neither is it necessary that we should. All that is necessary for a choice to be rationally made is that it be the result of a consideration of the objects of choice in the light of some criteria. The important point of the analogy, however, should not be missed, and that is, given the premisses, the conclusion follows, or, in a situation of choice, given the criteria and the objects, the choice follows. The conclusion, while it is stated separately, is entailed by the premisses; the choice, while it is made separately (if it is) is a consequence of

---

8 The importance of (1c) should not be overlooked; this premiss, a part of the criterion, serves to specify what conditions must be met in order for the choice to be made in accordance with the weight of reasons.

the objects and the criteria, in the sense that it is the result of a consideration of the objects and their properties in light of the criteria. All of the possible conclusions which follow from a valid argument may, in principle, be determined by and examination of the premisses; in the same way, all the possible choices which may be rationally made in a given situation may, in principle, be determined by an examination of the criteria and the properties of the objects to be chosen from.

CHOICE AS DISCOVERY

From these remarks it should be clear that the process of rational choice is not so much a matter of deciding as of discovering. Given the criteria according to which a choice is to be made and the objects from which the choice is to be made, all that remains to complete the process of rational choice is that the two should be put together, that is, that the objects should be considered in light of the criteria, but this process is simply that of discovering--by observation, analysis, or whatever--what the properties of the objects are and how those properies fit the criteria. Given a choice between two alternative courses of action, if the choice is rationally made, the process of deciding which course to take is not so much a process of "making up one's mind" about what one will do as it is of discovering, in that situation, what one has, in effect, already chosen to do (by

adopting certain criteria) or, to put it differently, of discovering what it is, in that situation, that one really wants to do. Of course, it is not to be denied that men can, and often do act irrationally; we act sometimes impulsively, or, like the gambler, make our choices arbitrarily. Nor is it to be denied that the distinction between rational and irrational choice is sometimes blurred and obscure; was Mary Todd's choice of Abraham Lincoln from among her many suitors a rational choice? Certainly, it was a choice which could have been given, at a later date, more than ample justification. She chose that man who, in her opinion, was most likely to succeed, and she chose rightly, but did she choose rationally? Did the young lawyer, Lincoln, actually fit her criterion better than her other suitors? It would be difficult, if not impossible to say and, probably, more difficult examples than this could be found. Nevertheless, it is clear that, despite the obscurity of the distinction in the middle, the distinction is valid at the extremes. There are choices which fit, in their general outlines, the process which I have described, e.g., the investments of the careful businessman as opposed to those of the compulsive gambler. These choices are rational: that is, they are rationally made and the process by which they are made is, in essence, one of discovery. This conclusion can be denied but only be taking the other horn of the dilemma and the consequent view that even the banker's criteria are non-rationally chosen. However, even if we take this

side of the dilemma and argue that ultimately the banker's criteria are in some way chosen, it cannot be denied that those criteria may function as criteria for his subsequent rational choices, and in this case it would be even clearer that the process of rational choice would be a process of discovering what he had, in effect, already chosen to do.

This conclusion, that the process of rational choice is a process of discovery, is one which follows explicitly from my analysis of rational choice. If the reader would deny this conclusion, he would have to admit either that men do not make rational choices, a view which, in light of the facts, it would be difficult and probably impossible to maintain, or he would have to admit that they do not make them in the manner that I have described, that my analysis is wrong. But, if this were his view, if he would argue that my analysis of the process of rational choice is wrong, then I would ask simply: What else could it be?

II

CRITERIA AND STANDARDS OF RATIONAL CHOICE AND EVALUATION

Rational choice I have defined, roughly, as a choice made in accordance with some standard or criterion or, alternatively, as a choice made because of or in accordance with reasons, where a reason is understood as a property of an object of choice seen in the light of some standard or criterion. However, I have to this point said little about the nature of these standards and criteria. I have, of course, given examples, but I have not attempted a definition. Quite frankly, I do not believe it possible to give a definition which would not appear to be circular. This is so because a standard or criterion is a standard or criterion of choice in virtue solely of what it does. The definition, then, must be functional. For instance, it would have to be something like, "A choice criterion is that in terms of which a fact about an object of choice becomes a reason for or against choosing it." This is the closest I shall come, in this chapter, to defining "standards" and "criteria". In the following pages I shall try, however, to indicate some of the properties of standards and criteria.

THE NATURE OF CRITERIA

As was pointed out in the first chapter, some criteria may be rationally chosen, but it would not be possible for all criteria to be so chosen since this would involve an infinite regress. If I am to defend my choice of criteria as being rationally made, I must again present my reasons for choosing them. Let us suppose that throughout this process, I have got the facts straight. Then the argument will center around my criteria: were the criteria in light of which these criteria were chosen themselves rationally chosen? This process cannot go on forever. At some point, I must come to criteria which I did not choose or, at the very least, which I did not choose rationally. These fundamental criteria which underlie and guide all our subsequent rational choices, I take to be our basic likes and dislikes.

In order to show that these basic likes and dislikes do serve as our fundamental choice criteria, we may begin by making a fairly straightforward assumption (an assumption which, I think, needs little argument) that, in the absence of any other considerations, it is rational to choose that which we like and irrational to choose that which we dislike. I wish to stress the qualifier here: I am not saying that it is always rational to choose that which we like but that this is so in the absence of any other considerations. I shall not attempt to argue directly for the truth of this assumption:

it seems to be obvious that it is just part of what we mean by being rational. If, however, the assumption is accepted, it would seem to be clearly the case that, in the absence of any other considerations, our criteria are what we like and dislike.

Now, assuming this to be so, let us ask what these other considerations, if there were any, could be. First, it is obvious that they must be reasons. That this is so follows from the fact that, if it is rational to choose that which we like, there must be a reason for so choosing (in this case, the fact that it is such and such an object together with the the fact that one likes it). The "other considerations" then, would have to make such a choice either non-rational or irrational. Since there is, in this case, a reason for so choosing, to make the choice non-rational, it would have to be the case that there were some other reasons of equal weight for not so choosing, and, to make the choice irrational, reasons of greater weight for not so choosing. All this is obvious enough and rather trivial. Now, I ask, what could these other reasons be. They cannot be just facts about the object. As I have tried to show, such a fact is not, by itself, a reason for or against choosing the object. It must be a fact seen in the light of some standard or criterion of choice. This standard or criterion must, in its turn, be such that in the absence of other considerations, it would be rational to choose in accordance with it (otherwise it would not make the

fact about the object a reason for or against choosing it). But since, in the absence of any other considerations, it is rational to choose what we like and irrational to choose what we dislike, it would seem reasonable to assume that these other criteria must be other likes and dislikes or must reduce to our likes and dislikes. I assume, therefore, that our fundamental, underived criteria are our basic likes and dislikes. With this in mind, let us look at the following example.

I am to make a choice between two chairs, one of which I shall buy. The chair which I buy is to be used only by me and I want it for no special purpose except to sit in when I feel like sitting. Suppose that the chairs cost the same, are roughly the same size, and are the same color. I have no preference for a particular style. My criterion is comfort. I sit in each of the two chairs and discover that one is much more comfortable than the other; as a matter of fact, I find that, while one is quite comfortable, the other is downright uncomfortable. Which should I choose? This is not like the example given earlier in which a financier was supposed to have noted that the alternatives open to him had such and such properties. In this example, the answer to the question has been given. If I am to choose rationally, then I should choose the more comfortable of the two chairs. In the absence of any other considerations, knowing that one investment presents a greater risk for smaller returns than another does not tell

me which to choose.[1] However, in the absence of any other considerations, knowing that one chair is more comfortable than another does tell me which to choose. To say that the probability that a particular investment will result in the loss of X dollars is m and the probability that it will result in a gain of Y dollars is n is to state a simple fact; it is simply to describe the investment in terms of its probable outcomes. But to say that a particular chair is more comfortable than another is not to state just such a simple fact. To say that a chair is comfortable is not just to say something about the chair, e.g., that it has some specific set of properties. It is to say that the chair has some, unspecified properties such that it is a comfortable chair in which to sit. It is to say that the properties of the chair are such that, other things being equal, one enjoys sitting in it, that one enjoys sitting in it more than in a chair which is uncomfortable, that, given a choice, one would choose to sit in just such a chair. But, suppose that after I choose the more comfortable of the two chairs, I am asked to account for my choice, to show that the choice of this chair over the other was rational. I would, of course, reply that I had chosen this particular chair because it was the more comfortable of the two. My questioner may wonder why this was a sufficient reason. I tell him that I was in the

---

1 See note 5, p. 12.

market for a comfortable chair, that is, that "comfort" was my criterion. "But why comfort?" he may ask. Now, he may be asking one of two different questions here. He may want to know (1) why comfort rather than some other criterion, or he may want to know (2) why I should be concerned with comfort at all. To the latter question I can give no answer. I can only assume that he does not know what "comfort" means. To say that I like being comfortable would be, in a sense, redundant. "Comfortable" is a word which applies to a particularly pleasant or, at least, not unpleaseant bodily state, and to say that I am comfortable is, in effect, to say that I am in a particular kind of bodily state which, other things being equal, I like being in. The only way that I can make his question meaningful, then, is to take it in the other sense. In this sense, to reply that I like being comfortable would be to make a negative statement to the effect that there were no other considerations which would have outweighed my liking for comfort. In the same way, to say that I disliked comfort would be to contradict myself unless I were to qualify my statement with a statement like, "Asceticism is good for the soul" or "Sitting in a comfortable chair gives me back pains." Taking the question in the former sense, then, I can justify the adoption of "comfort" as my criterion only by reference to other possible criteria. For instance, I may say that I am not an ascetic: I do not believe I shall experience any adverse consequences from leading the

"easy life". I may point out that this comfortable chair provides adequate postural support so that the possibility of back pains is not a consideration. I may say that the only other applicable criteria are such things as color, cost, style, size, etc., but that these things are the same for both chairs so that they provide no ground for choice or that, where they differ, I have no preference. Now, it is important to note how this sort of account differs from saying simply that I chose this chair because it was the more comfortable of the two. It cannot be argued that the chair was not the more comfortable. Nor can it be argued that I had a preference where I say I had none--but it can be argued that I should have had a preference. For instance, it could be argued that the style of the chair which I chose would not fit the decor of my apartment.

Such an argument would succeed in showing me that my choosing in accordance with the criterion, comfort, rather than other possible criteria was irrational only if it can be shown that (1) there were other applicable criteria which, if used, would have led to a choice more in accordance with (or in accordance with more of) my basic likes and dislikes and that (2) I knew or had reason to believe that there were such applicable criteria but did not take them into account. If only (1) were shown, my choice would be shown to have been unreasonable but not, therefore, irrational. Such an argument, however, could never succeed in showing me that comfort was

not a criterion, only that, in that situation, it was not rational to choose in accordance only with the criterion of comfort. Again, this is so because comfort, as a choice criterion, functions at the level of our basic likes and dislikes. I do not have a reason for liking comfort; I do not choose to like comfort. That I like comfort is just one of the facts about me, and it is, therefore, neither rational nor irrational.

A choice criterion or standard can be rational or irrational only if it is chosen. It will be a rational criterion if it is chosen rationally, i.e., if an attempt is made to chose or adopt a criterion which is in accordance with one's warranted beliefs and with one's basic, unchosen criteria. A choice criterion will be reasonable, on the other hand, if it is in accordance with one's basic criteria and with the facts. A choice made on the basis of criteria which are chosen will be rational only if the criteria are rationally chosen. The rationality of the choice will not depend on the reasonableness of the criteria, nor will the reasonableness of the choice necessarily depend on the reasonableness of the criteria (for much the same reason that the truth of the conclusion to a valid argument does not necessarily depend on the truth of the premisses). We must be careful, however, in speaking of the rationality of criteria. Not all criteria which are not basic are chosen or adopted, although they are derived from our basic criteria

(how will be discussed briefly in chapter four). This fact is not especially important for the purpose of this thesis; nevertheless, it should be noted that a choice made on the basis of criteria which are not basic will be rational only if those criteria were rationally chosen. Again, the reasonableness of the choice will not depend upon the origin of the criteria.

It might be worthwhile, at this point, to consider an example. Consider, again, the financier discussed in the first chapter. His choice criterion is "minimum risk for maximum returns." The financier knows what this means in fairly precise terms; he also knows a lot about business and the factors affecting the outcome of business investments. When he says his criterion is "minimum risk for maximum returns," he gives a summary account of his criterion. As it is stated, the criterion will inform no one how to go about making such a choice. The financier, however, would understand this criterion to mean that in choosing he is to consider such things as demand, competition, profit margins, capital and so on, and that he is to consider all of these things as interrelated and summarizable under the general criterion of "minimum risk for maximum returns." That all of these things are to be considered under this criterion he has learned, perhaps through experience, perhaps through study. They constitute his beliefs about what properties a particular investment should have in order to meet his criterion. His

criterion specifies what factors are to be considered in order for a choice to be made in accordance with it. Now, suppose that the financier's purpose for investing in the first place is to make money. We could say, then, that, in this case, his fundamental criterion is to make money.[2] In some cases, such a criterion might be sufficient by itself, but in most it is not easy to tell which investment will make money or how much money. This being so, the financier might adopt another criterion such as "minimum risk for maximum returns" in order to determine which investment is more likely to make more money. But he might adopt a different criterion; for instance, he might choose on the basis of whether or not he likes the people involved. In the absence of other considerations, such a choice would be irrational unless he believed that a choice which met this criterion would also meet the criterion of making money (The other considerations would, of course, be some other criterion which would outweigh that of making money and which would be met by a choice made in this way, e.g., a criterion like "moral integrity"). The reader should not be misled, however, by my use of such terms as "purpose" and "desire." In some cases, these are our criteria, but not in all. For instance, the apple sorter mentioned in chapter one, if asked why he put some apples in one box and some in another, might reply that

---

2 Such a criterion would not, however, be basic in the sense that I have been using that term. Money is, afterall,

he put eating apples in one, cooking apples in another and unusable apples in another. These are his criteria (of course, he may well have standards for distinguishing eating and cooking apples) but they need not be his purpose. His purpose is to keep his job, and his desire is for his wages.

Summarizing, then, we can say that (1) our basic criteria are our basic likes and dislikes, (2) our basic criteria are non-rational, (3) criteria which are not basic may be rationally chosen, (4) a criterion is rationally chosen if it is chosen in accordance with one's basic criteria and one's warranted beliefs, (5) a criterion is reasonable if it is in accordance with one's basic criteria and the facts, (6) a criterion specifies which aspects of the situation and the alternatives are relevant to the choice, and (7) the fundamental criterion in a given situation may be the chooser's purpose but need not be.

HUME DEFENDED

This view of rational choice also serves to circumvent the objection to the Humean point that "Reason is, and ought only to be the slave of the passions."[3] For instance, an argument which could be made against Hume might be something like this: Suppose a judge had a very strong feeling against sentencing a man who had been found guilty of a crime.

---

the paradigm of extrinsic, or what I would call derived value.

3 L.A. Selby-Bigge, ed., Hume's Treatise of Human Nature. London: Oxford, p. 415.

Reason alone might tell him that he ought to sentence him, but since his reason is the slave of his passions and since he has a passionate feeling against sentencing him, he could not sentence him, nor could reason tell him to. But we may suppose that in such cases judges very often do sentence the guilty party. Therefore, reason is not the slave of the passions. Now, this argument, whether directed against me or against Hume, is based simply upon a misunderstanding. This is clear from my analysis of rational choice. A judge's criterion for sentencing a man, we may say, is that he be found guilty of a crime. His choice (between sentencing him and not sentencing him) has nothing to do with how he feels about sentencing him, but with which act best fits his criterion. His adoption of this criterion follows, in turn, from other criteria which reduce finally to his basic likes and dislikes. In this case, the judge would have adopted such a criterion because that is the law and he, for some other reason, desires to obey the law (e.g., he could not remain a judge otherwise: he could go to jail himself: the law must be respected to preserve the integrity of society; etc.). In the end, the choice reduces to one or more choices made on the basis of one's likes and dislikes (Hume's passions). It is not, in this case, reason which countermands passion, but passion itself.

EVALUATION AND THE GOOD

This view of reason in the context of rational choice has, I think, important implications for ethics and theory of value. Actually, it would appear that the process of rational choice differs from that of evaluation only in the fact that the one involves a choice whereas the other does not.[4] If this is the case, then we could say that rational choice differs from evaluation as a subset from a set. On this view, we could say that a rational choice between objects is a process of evaluating them with a view to choosing between them. It would appear to be intuitively obvious that evaluation could not take place in the absence of some standard of value. We may wish to say that value or the good is simply intuited by some special faculty but, on this view, evaluation would seem to be, at best, a superfluous process. Determining the value of some object would be a matter simply of looking at it (in whatever way one looks at things with the special faculty) and noting, perhaps with some effort, either that it has some value or that it has not. Where the determination of the object's value status required some effort, where it is difficult to perceive, the process would be analogous to that of straining one's eyes. I shall not argue for or against such a view. It seems to me to be false, but I shall simply assume,

---

4 For a similar view, see Oppenheim, Felix E. "Rational Choice" Journal of Philosophy, L, no. 12 (June 4, 1953), pp. 341-350.

as I assumed that men do make rational choices, that men do, at least occasionally, make rational evaluations, and we can say, I think, that a rational evaluation is one made on the basis of reasons, where a reason is some property of the evaluatum seen in the light of some standard of value.[5] Again, the standards here will all reduce ultimately to our likes and dislikes.

Now, this being so, it might be thought that such a view would be open to the same objections as those levelled against the hedonist view that the good is pleasure or the absence of pain.[6] The two definitions differ, however, and the definition I am offering is superior in that it enables us to see both the truth and the error in the hedonist view.

Pleasure is not a good. Certainly it is not the chief good, nor, for that matter, is displeasure an evil. We can show this by assuming that our basic likes and dislikes correspond to our basic pleasures and displeasures. Pleasure,

---

5 I believe that the distinction between the meanings of "value" and "evaluation" or "valuation" (I take these terms to be synonomous) can be made in a manner quite similar to the distinction between the two senses or rational choice. "Value," then, corresponds roughly to the sense in which a choice is reasonable, and "evaluation" to that in which a choice is "rational." "Value," then, refers to that which is valued, to the evaluatum, and "evaluation" to the process by which it comes to be a value or to be valued.

6 Quoting Epicurus: We recognize pleasure as the first and natural good. ("Letter to Menoeceus" Geer, Russell M., trans., Epicurus: Letters, Principle Doctrines, and Vatican Sayings. Indianapolis: Bobbs-Merrill Co., 1964, p. 56.)

then, would be our basic standard. It would not be a good or the good, but the standard of goodness. The good, therefore, does not reduce to pleasure or to a balance of pleasure over displeasure; the standard of goodness or of value reduces to pleasure or to our basic likes and dislikes. Of course, this is not intended, nor should it be taken as a serious objection to the hedonist position. My purpose in making it is just to clarify the notion of goodness.

As was pointed out above, the distinction between evaluation and value is analogous to that between the two senses of choice. We may say, then, that an evaluation may be rational or irrational: a value, reasonable or unreasonable. An evaluation will be rationally made, then, if it is made on the basis of or in accordance with reasons, where a reason is a warranted belief about the object or the evaluatum seen in the light of some standard of evaluation. These standards, if they are rationally chosen, will, again, be chosen in accordance with our warranted beliefs and our basic, unchosen standards. A value, on the other hand, will be reasonable if it is in accordance with the facts and one's basic likes and dislikes. When we say, then, that an object is good or has value we are, in effect, saying two things: (1) It is a true statement that the object is such and such an object or that it has such and such properties and (2) its possessing these properties is pleasurable. The good, then, is not equivalent to pleasure, considered in itself, but to

the pleasurableness of objects, and an object is good if it is a fact that, all things considered and in light of our basic likes and dislikes, the objects is pleasurable and not displeasurable or more pleasurable than displeasurable.

To say that a given object is, in fact, good, then, is not just to say that I like it or that it gives me pleasure, but that there is reason to like it. Thus, even though contemplating an object is pleasurable, it does not follow that the object is, in fact, good, since the facts may be other than I take them to be. What follows is only that I take the object to be good. (Of course, we do sometimes say that a thing is pleasurable but not good: what we mean here, I think, is that it is pleasurable in some respects but not in all. For instance, sex may be pleasurable in the act but displeasurable in its effects and, therefore, bad. Obviously, I do not mean to imply here that any respectable hedonist has ever denied this. However, hedonists have occasionally been misunderstood and have been thought, if not to deny it, at least not to consider it). From this it is obvious that we never have entirely adequate grounds for saying that an object is, in fact, good, since the facts may always be other than we take them to be. It is then, logically impossible to prove that an object is or is not good. At best we can say that, on the basis of some specific, warranted beliefs, it would appear to be good. As with rational choice, however, it does not follow from this that evaluations cannot be rationally

made. In evaluation, I consider the properties of the evaluatum in the situation as I am aware of them, taking all of those factors of which I am aware into account. If, on the basis of these apparent facts, it follows in light of my standards that the object is good, I am justified in calling it good. That is, it would be rational for me to value it and irrational not to. It does not follow, of course, that the object is, in fact, good or that it is reasonable for me to value it. The situation is analogous to that which we face in science. At some point in our investigations, on the basis of our theories, beliefs, and data, we are justified in saying that a given fact is a fact, that it would be contrary to reason to deny it. It does not follow, however, that it is a fact, only that, given the background information, it would be contrary to reason not to believe it is a fact. By the same token, we may be sometimes justified in saying that an object is good, that, on the basis of our background information and beliefs, it would be contrary to reason not to believe it good. The bigot forgets this is only a matter of what it is (conditionally) reasonable to believe.

AN EXAMPLE

To make this point clearer, let us take the following example: Suppose there is a man who is a murderer-rapist, a man whose greatest joy is raping the bodies of women whom he has just murdered (I choose such an example not just for its

shock value but because it is one on which nearly everyone can agree. What this man does is wrong, immoral, bad, and everything else reprehensible). Is this man irrational? Certainly, but no more so, I think, than the man who simply abhors such activities. We need not concern ourselves here with how this man acquired his peculiar tastes. Let us say simply that it is the result of his genetic endowment and his individual experiences. Suppose such a man to be presented with the possibility of committing such an act without fear of being apprehended, without fear of consequences adverse to himself. Should he commit the act or not? Which would be the more rational choice? Let us suppose the man is not compulsive. He is aware of what he is doing, of the implications of what he is contemplating, and he is able to take all these things into account. Except for his strange penchant for committing murder and rape, he is a normal human being. He feels some pity, perhaps, for the girl, her parents, and friends: he attaches some value to human life, but after some reflection he finds he is not much affected by these considerations. What, then, would he choose to do if he were to choose rationally? He would choose, and I think it obvious, to commit the rape. Were this not so there would be no need of laws against rape. We exempt the irrational from punishment: our laws and the jails to enforce them exist only to deter the rational, to provide an additional, hopefully sufficient, aversive factor in order to deter anyone contemplating such

an act from committing it (Actually, it is only supposition to call it aversive. We assume that imprisonment or execution is aversive, but it may not be. All we do is make it a fact that committing such an act will be followed, upon apprehension and conviction, by imprisonment or execution. Whether or not this is aversive is a matter of how the individual concerned regards the fact). We hope that persons contemplating such acts will be rational, that they will be able to consider this fact rationally, that they will dislike it, and that they will choose accordingly. To have the capacity for rational behavior is to be able to choose in accordance with one's likes and dislikes. To choose rationally is to choose in this way. The rapist described here might well be abnormal in the extreme: he might well be perverted, but it does not follow that he must be irrational as well. Of course, this is not to say that every decision to commit such an act is rational. The rapist's dislike for imprisonment might well outweigh his liking for rape: the situation might be such that he would surely be caught and imprisoned. He might be aware of this and still commit the act, but it is in precisely this sort of situation that we would call his behavior irrational and in precisely this sort of situation that we would exonerate him. Why? Because institutions such as prisons do not exist primarily for the sake of inflicting suffering, but for the sake of the threat which they constitute for anyone contemplating a punishable act. But the man I have described is not irra-

tional: furthermore, punishment is not a factor for him. He is sure that he will not be discovered. Let us suppose that he makes the rational choice. He commits the act. Now, we are going to get into difficult terrain, for what I want now to ask is whether what he does is good or bad.[7]

When the girl's body is discovered, society will be agreed that the act was bad, immoral, wrong, but what about the rapist? He, if he is rational, will think it good. Will he be justified in thinking this? That is, does society have any real ground for calling it bad? What makes this problem so difficult, I think, is the fact that both questions may be answered with a "Yes." It is possible for one and the same thing to be both good and bad and in the same respect, depending upon who is calling it good: who, bad. Thus, if we were to put the rapist and a representative member of society together and to ask each to defend his views, to give their arguments for calling it good and bad respectively, we might find that their reasoning is unimpeachable and, further, that each has got the facts straight as far as can be ascertained. In short, we might find that each is correct in his view, that is, that, all things considered, each is a view which is reasonable and neither view is more reasonable or more

---

7 The reader may well object to this usage of "good" and "bad" rather than "right" and "wrong" to describe an action. I do, however, have reasons for this usage which will be given below. I ask the reader to bear with me until such time.

believable than the other. Let us suppose the rapist justifies his action on hedonistic grounds. It was an act, he says, which gave me pleasure and which was not significantly displeasurable. To such a "justification" the representative member of society would, in all likelihood, react with horror. "Suppose," he may reply, "that I were to murder you. Would you like it?" "Of course, not" the rapist will reply. "Then you wouldn't call it good?" "No." "Well, isn't the situation exactly the same for this girl whom you murdered?" No, it is not the same. It is not in the least inconsistent for the rapist to say, on the one hand, that his being murdered is bad and, on the other, that someone else's being murdered is good. The rapist, of course, values his own life, but this does not, by itself, entail that he must value any others. It may be possible, however, to show him that he ought, that is, to show him that, given his basic values, valuing the lives of others is more reasonable for him than not valuing them. This could be done, perhaps, simply by acquainting him with the facts, by pointing out that he is dependent upon the well-being of others for his own well-being. If he is convinced of this fact, then it may be possible to show him that, all things considered, calling such an act good is inconsistent with the value he places on his own continued well-being. However, even though the rapist might agree that such acts are generally bad, he might disagree about this particular case. He may say, for instance, that this girl was not a

productive member of society and that, while he is concerned with the well-being of society, since the well-being of society is not affected by the death of this one non-productive member, he does not feel that what he did was bad. Here, the representative member of society may take several different approaches to the problem. For instance, he may argue that, yes, there is a sense in which society could be said to be better off for her death, but that, nevertheless, society is affected adversely since commission of such acts contributes to the unhappiness of its members with a subsequent reduction in production and so on. If the rapist is convinced by this argument, he may admit that what he did was, on the basis of his own standards, bad, and he may go so far as to adopt the happiness of society as a criterion for choosing his actions in the future. If, however, he is not convinced, then the representative member of society may take a different approach: "Look here," he may say "as a full-fledged member of this society you cannot say that such an act is good. Are you not aware of the example you are setting? Perhaps, there is no good reason that this girl should not be murdered, but if we say that people may be murdered whenever their lives cease to have value for society as a whole, people will live in constant fear of death at the hands of others, the whole structure of society will collapse. You must admit that your act was bad if for no other reason than that it sets an example which you would not want others to

follow and which you could not prevent them from following." Here, the rapist will have to admit that, while he, as an individual, does not think this act bad, as an individual who is a member of, and dependent upon society, he must agree to call it bad.

THE OBJECTIVITY OF "RIGHT" AND "WRONG"

Throughout this discussion I have attempted to avoid using such terms as "right" and "wrong". This is because my concern has been, until now, primarily with what the individual judges to be good or bad. There would, perhaps, be nothing wrong with using such terms: however, I think it would be a more accurate reflection of what we mean by "right" and "wrong" or of the way in which these terms are generally used, if we assume that they refer primarily to what society judges to be good or bad. That is to say, the terms "right" and "wrong" apply, for the most part, where the evaluation made in terms of them is based upon some generally, but not necessarily universally, accepted standard of value. Making this assumption will enable us to encompass the view that there are objective standards of right and wrong (as well as the view, which would seem otherwise to be inconsistent with what I have said thus far, that the "moral point of view" is both impartial and general), "objective" in the sense that they are evaluations about which nearly every rational person amply acquainted with the

facts can be expected to agree. Thus, we may say that certain acts are objectively wrong because they are held by so many people to be subjectively bad (Of course, it must be remembered that the mere fact that something is held by a rational person to be bad does not entail that it is. This depends, at least in part, on his having got the facts straight). Since it is fairly safe to assume that people's basic likes and dislikes, where these are not socially acquired (this qualifier is needed since it is social standards which are being judged), are very similar, being largely a matter of their genetic endowments, that is, of their being members of the species Homo sapiens, we may assume that the standards of right and wrong will be, in this sense, almost entirely objective and will vary only according to the situation and the facts available.

ETHICS AND THE LIMITS OF REASON

Now, suppose that the rapist does not accept the last step in the argument. Suppose, further, that his disagreement does not rest upon a disagreement about the facts: he agrees that committing such an act will have the social consequences specified. Suppose he agrees that they will even be detrimental to his own continued well-being. "But," he says, "I have had a full life and this was one of the few pleasures I had not experienced. Having experienced it, I am content to face the consequences, and I am fully aware of what those conse-

quences are likely to be." At this point, rational argument ceases, but this is not because either of the parties concerned is irrational. It is because they have reached the limits of rationality. Reason begins with our basic likes and dislikes and it ends with them. Given the fact that their basic likes and dislikes differ, it is rational for the parties concerned to hold their different views, and neither view is more reasonable than the other. Given an equivalent knowledge of the facts, it would be rational for the rapist to evaluate the act as good and for the representative member of society to evaluate the act as bad. Given this, it would be rational for the rapist to choose to commit the act when faced with such a choice and it would be irrational for the representative member of society so to choose.

It would not, however, be irrational for each to evaluate the act as wrong. Both the rapist and the member of society could admit that the act was (objectively) wrong and that, to the extent this is true, that the rapist deserves punishment. We have assumed, however, that the problem here is not the result of a rational difference of opinion but of different basic likes and dislikes, and since, in this case, they differ in such an extreme fashion, we do not commit the rapist to imprisonment. We assume, rather, that there is something psychologically wrong with him, that he cannot "tell right from wrong." His basic standards are such that acting in accordance with what is right is not, at least in this case,

for him the rational thing to do. We absolve him, then, of legal responsibility. He did not choose his basic likes and dislikes. Perhaps, something is defective about his chromosomes; perhaps, he had a defective rearing. At any rate, his standards differ and that is all that can be said about it. Fortunately, that is not all that can be done about it. We commit him to a mental institution for the criminally insane where he can do no harm and may, hopefully, be cured. While the rapist and society may disagree about the subjective goodness or badness of this course of action, they will not, or at least need not disagree about its "rightness."

It is important to note, however, that the one view predominates over the other only by force of numbers (this will be so, at least, with rape if not with murder). Were it the case that rape were not among the basic dislikes of, or did not in any way conflict with the basic likes of people, rape would not be wrong. The correctness of the view that rape is wrong depends upon the reasonableness of believing that it is among the basic dislikes of people or that it conflicts in some way with their basic likes. Where the basic likes and dislikes of men are the same, then, ethics is purely a matter of reason (but not a matter of pure reason which doesn't exist except, perhaps, in the empty syllogisms of formal logic).

From this, we can provide solutions to such perennial philosophical questions as the "is-ought" problem. Logically,

an ought follows only from an ought, an evaluative statement only from an evaluative statement, but this is not at all inconsistent with the view that both are derived from facts. To say that one ought to do something is to say simply that it is, given the facts, the most reasonable thing to do. To ask what we ought to like and dislike, what our standards of goodness ought to be, when this is a question about our basic likes and dislikes, is meaningless, because it must be interpreted as "What, in light of our basic likes and dislikes, ought our basic likes and dislikes to be?" To the statement that good is pleasure, the response, "Yes, but is pleasure good?" is meaningless. Pleasure, our basic likes and dislikes, are neither good nor bad, but the source of good and bad. If it were to be said that pleasure is bad, it could only be meant that an indulgence in pleasurable things, for the sake of pleasure, will lead to displeasure. And what this would mean, I think, is just that sometimes an indulgence in unpleasant things is more rewarding, in terms of pleasure, than is avoiding all such unpleasant things. The golden mean is a standard of virtue, if it is, not because of some higher moral law, but because in the long run to act according to it is more rewarding than not to. The important thing to note is that, once we learn what are the basic likes and dislikes of men, the problem for ethics, the philosophy of right and wrong, hence, of <u>general</u> rules of conduct, is one of discovering what are the facts and what, given those facts, is the most

reasonable thing to do--in the sense of reason that I have described here..

III

THE CONCEPT OF DETERMINISM AND THE PRINCIPLE OF CAUSAL ENTAILMENT

The foregoing discussion of evaluation in a thesis on determinism may cause some consternation among my more traditionally minded readers. Since the time of Hume, it has been forcefully maintained by many philosophers that it is not possible logically to derive an evaluative conclusion from premisses which are all descriptive or non-evaluative. I do not intend here to dispute this statement, but only what some may regard as a corollary, namely that, if evaluative statements cannot logically be derived from non-evaluative statements of fact, then evaluative statements and, hence, values are not and cannot be derived from facts at all. There is, I maintain, a sense in which it may be said that evaluations are entailed by facts, although the entailment is not logical but causal.[1] To defend such a view, it will be well, perhaps, to make clear what I mean by "causal entailment." Accordingly, it is to a discussion of

---

1 The discussion of this topic, more commonly known as the "is-ought" problem, will be taken up in the following chapter. I should point out at this time, however, that, if, as I have claimed, the process of rational choice is a process of evaluation, to show that evaluations can be causally entailed would be equivalent to showing that the concepts of rational choice and determinism are not inconsistent.

this concept and the related concepts of "cause" and "determinism" that this chapter and a large portion of the chapter following will be devoted. I should point out at this time, however, that, in view of my remarks in the last chapter concerning the relationship between evaluation and rational choice, and in view of the over-all purpose of this thesis, the question whether evaluations can, in any sense, be derived from facts is an extremely important one.

THE PRINCIPLE OF CAUSAL ENTAILMENT

The principle of causal entailment or (less satisfactory) causal necessity states simply that, given any event A, A is an event such that another event B will follow or occur.[2] This principle is closely connected, but not identical with the determinist thesis which may be stated as: Given any event B, there is an antecedent event A such that, given A, B must follow. The determinist position is sometimes reformulated as simply: Every event has a sufficient cause. The principle of causal entailment, however, works in the opposite direction. It might be reformulated, with some inaccuracy, as: Every event is a sufficient cause. The principle of causal entailment, in other words, is the converse of determinism. It states simply that events have necessary

---

2 This statement of the principle is not, as it stands, precisely accurate, and a large part of this chapter will be devoted to an elucidation of the necessary changes.

consequents and, further, that the consequent is made necessary in virtue of the nature of the event (This last sentence will receive substantial qualification later in this chapter).

Empiricists, at least those in the Humean tradition, may have difficulty accepting this last statement. This difficulty may be partially ameliorated, if not overcome, by the two following arguments.

HUME AND NECESSARY CONNECTION

In the first place, Hume's arguments against the necessary connection of cause and effect, at least so far as I understand them, are not directed so much against causal necessity as against our knowledge of that necessity, so that, if his arguments are accepted, causal necessity is not thereby shown to be impossible but simply unknowable. But if we cannot know that there are necessary connections between some events, neither can we know that there are no necessary connections between events; the two views are, therefore, logically on a par and we may as easily accept the one as the other. There are, however, compelling reasons for accepting the view that there are, in fact, necessary connections between some events and their effects, namely that, if we do not accept the principle of causal necessity in some form, at least hypothetically, the apparent uniformity of nature becomes rationally inexplicable (this is not, of

course, to say that it is rationally explicable).[3] Those persons who are not philosophers, and most of those who are when they are not acting in their capacities as philosophers, do believe that nature is reliably uniform, that events can be explained in terms of causal sequences. Yet, surely, such explanations would be irrational if the principle of causal entailment were held to be completely false. Empiricist philosophers accept this conclusion, but, so far as I know, non-philosophers do not and neither do empiricist philosophers when they are not teaching or writing philosophy. I suggest, therefore, that there is good reason to accept the notion of causal entailment and no good reason to deny it.

The above argument, it should be noted here, is not to be taken as an argument for the truth of determinism or, for that matter, for the truth of the principle of causal entailment. The argument does not prove, and is not intended to prove anything at all regarding the truth of these two notions. It is offered, rather, as a persuasive argument that, given that the doctrine of causal necessity can neither be proved nor disproved and that, from the apparent uniformity of nature, it would appear to be true, it is reasonable to accept the view that it is true, at least as a working hypothesis.

My second argument will be somewhat longer than the

---

3 This subject will be taken up again in chapter five.

first. Hume not only denied that we have any knowledge of necessary connection; he also proposed a theory of causation which denies any connection at all between cause and effect, a theory which, in one form or another, has gained a large measure of support among modern philosophers. Since it is part of my purpose in this present chapter to propose a different theory, it would be well to begin by refuting Hume's theory of causation. That Hume's theory of causation and the relation between cause and effect must be wrong will, I think, be clear from an analysis of that theory in terms of his theories of time and space.

A REFUTATION OF HUME'S THEORY OF CAUSATION

Hume begins his disquisition on the doctrine of the infinite divisibility of time with the observation, "'Tis also obvious, that whatever is capable of being divided in infinitum, must consist of an infinite number of parts, and that 'tis impossible to set any bounds to the number of parts, without setting bounds at the same time to the division."[4] Hume does not attempt to say why this should be so; he simply states it to be obvious (It is not, however, at all obvious,

---

4 L.A. Selby-Bigge, ed., Hume's Treatise of Human Nature. London: Oxford, p. 26. Future references to this work will be abbreviated to THN. The following discussion of Hume's theory of space and time, for all its length, is not, I should note, superfluous but will form the necessary basis for my argument against his theory of causation.

and if, in fact, it appears to be almost trivially true, as it may to some, I would suggest that it seems so in virtue of a confusion between two very different, but closely connected concepts, those of division and separation. However, the soundness of Hume's theory of time is not my primary concern here, and I shall leave this subject for another time.). Much of the subsequent discussion is concerned with the divisibility of ideas, impressions, and so forth, and, while the discussion of these topics is certainly germane to the subject at hand, I shall ignore them. I am concerned here only with Hume's expressed and implied views on the subjects of time and space, not with his views on the nature and origin of our ideas of time and space.

The next important step in Hume's discussion is the determination of whether or not it is "a contradiction to suppose that a finite extension contains an infinite number of parts," since, if it is, then it would plainly follow, on Hume's view, that no finite extension could be infinitely divisible. He argues, then, from an account of his idea of "a part of extension," that "the idea of an infinite number of parts is individually the same idea with that of an infinite number," and, since these are one and the same idea, it follows that "no finite extension is capable of containing an infinite number of parts," and finite extensions, therefore, are not infinitely divisible.[5] There follows here an addi-

---

5 THN pp. 29-30.

tional argument for the same point from the concepts of number and unity. It is not necessary to give the whole of this argument. Its conclusion is, as might be expected, "that unity, which can exist alone, and whose existence is necessary to that of all number...must be perfectly indivisible, and incapable of being resolved into a lesser unity."[6] No more need be said about this argument in the present discussion, except to mention that its conclusion depends, I think, upon the initial assumption that that which is infinitely divisible must consist of an infinite number of parts.

While it might, at first glance, appear that the foregoing discussion of extension has nothing to do with time, Hume points out that "The infinite divisibility of space [space=extension] implies that of time, as is evident from the nature of motion."[7] From this he takes it to follow that, since space is not infinitely divisible, neither is time. That this conclusion follows "from the nature of motion" will become important later on.

The next distinctive property of time which Hume attempts to elucidate is that of succession. He says "'Tis a property inseparable from time, and which in a manner constitutes its essence, that each of its parts succeeds another and that none of them, however contiguous can ever be co-existent.... every moment must be distinct from, and posterior and antecedent to another. 'Tis certain then that

---

6 THN. p. 31. 7 Ibid.

time, as it exists, must be compos'd of indivisible moments."[8]

To this point, Hume has argued that time and space or extension consist of parts which are indivisible, that any finite space or extension consists of a finite number of such parts, that, from the nature of motion, a finite time, too, must consist of a finite number of parts, and that the parts of time do not "co-exist." Up to this point, however, he has not given an explicit statement of what time is. Nor does he, to my knowledge, offer such an explicit statement elsewhere in the Treatise. Nevertheless, it is not at all difficult to formulate such an explicit statement from his more general remarks about the origin of the idea of time and the distinction between time and space.

The idea of time, he says, is "derived from the succession of our perceptions of every kind."[9] Were these perceptions all of them present at the same time, we should derive from them no idea of time, but, if anything, of space. "For that quality of the co-existence of parts belongs to extension and is what distinguishes it from duration."[10] From this it is easy to see that space and time differ in only one major respect, the "manner" of existence of their respective parts. Hume later makes this point even more explicitly when

---

8 Ibid. Chiefly for reasons of my own which will be made clear later, but partly for reasons of grammar and clarity, I would suggest that the word "posterior" in this passage be read as "postcedent" which is, I think, what Hume actually means.

9 THN pp. 34-35.

10 THN p. 36.

he says, "The ideas of space and time are therefore no separate or distinct ideas, but merely those of the manner or order, in which objects exist."[11] Since the only difference in the "manner or order" of their existence is that, in the case of space, the parts are "co-existent" and that, in the case of time, successive, it appears obvious that time is to be identified with succession. Hume acknowledges this identification much later in the Treatise in the following passage: "Space or extension consists of a number of co-existent parts disposed in a certain order, and capable of being at once present to the sight or feeling. On the contrary, time or succession, tho' it consists likewise of parts, never presents to us more than one at once: nor is it possible for any two of them ever to be co-existent."[12]

Having got this far, there remains only one question to be asked about the nature of space and time, and that is roughly "What does it mean to say that space and time consist of indivisible parts?" Clearly, it is to say that space and time consist of parts each of which is the smallest possible of its kind and, further, that each is a separate and distinct entity from every other of its kind. In other words, any finite extension or space consists, for Hume, of a vast but not infinite number of discrete points or parts which, when taken together, constitute that particular space or extension.

---

11 THN pp. 39-40.

12 THN p. 429. Underlining added.

By the same token, any finite period of time consists of a vast but finite number of discrete instants, no one of which is time or any period of time but which, insofar as they do not co-exist but succeed one another, constitute that particular period of time. That this is so, that the instants comprising any period of time are separate and distinct, is clear from the following passage: "For the same reason, that the year 1737 cannot concur with the present year 1738, every moment must be distinct from, and posterior or antecedent to another."[13] Hume here clearly states that they are distinct; that they must also be separate follows from his dictum that "everything, that is distinguishable, may be separated."[14] Space and time, then, for Hume, consist of discrete, separate and distinct, indivisible parts, and any finite space or time consists of a finite number of such parts.

Before we turn to the actual topic of this section, there remains to be discussed one other concept, that of motion. Hume says that the infinite indivisibility of space implies that of time and that this "is evident from the nature of motion." Oddly enough, he does not attempt to tell us precisely what is the nature of motion. Perhaps, he felt that this subject required no discussion, but obviously motion is not and cannot be what we ordinarily conceive it to be, given Hume's understanding of space and time. For instance, motion

---

13 THN p. 31. 14 THN p. 36.

cannot be infinitely divisible, for, if it were, it would not serve as a basis for deriving the indivisibility of the parts of time from that of the parts of space. What then can motion be? Obviously, motion cannot be defined as an object's moving from a place A to a place B. It does not require a David Hume to spot the circularity in that definition. Let us suppose, then, a plane surface such as a table. This surface will constitute a finite extension and will be composed, therefore, of a finite number of indivisible parts. Now, let us draw a line along the surface of this table extending from point A to point B, and let us suppose an object, such as a ball, to be moving along this line. For the purposes of the argument, we may assume, for the moment, that the ball's motion will consist in its moving or going from A to B. It is plain here that the ball's motion cannot be separated from the ball and that it cannot be separated from the line along which it is moving, so that, if the ball's motion were infinitely divisible, it would follow that the line along which it moves is infinitely divisible. Since this is not the case, the ball's motion is not infinitely divisible. Now, time is a succession of indivisible, separate and distinct instants, and it takes some time for the ball to move from point A to point B. Furthermore, the space through which the ball moves is a certain, linear arrangement of separate and distinct, indivisible points. It is plain that the ball's motion can be neither more nor less divisible than the time that it takes

to move. Let us say that the line AB consists of ten points. In its motion from A to B, then, the ball will occupy ten points, but it will not occupy these different points all at once; it will occupy first one, then another until finally it has occupied each of them and has moved from A to B. But, in saying that it will occupy first one and then another, we are saying nothing more than that it successively occupies a series of points and, insofar as its occupation of these different points is successive, it takes time. We are in a position, then, to define motion (at least motion along a straight line) as "the successive occupation of adjacent, linearly ordered points or places."[15] Now, it should be noted that this is very different from the usual conception of motion. The ball's motion between any two adjacent points is not a moving from one to the other. It is simply its being in the one place at one instant and its being in the other place at the next. It is not a continuous motion, but the occupation of discrete places at different times. It is, as it were, a series of instantaneous leaps. This interpretation, if it is an accurate interpretation of Hume's understanding

---

15 While I believe that this or at least a very similar definition of motion follows from what Hume says about space and time, I do not believe that it is at all an adequate definition. For instance, if there are ten points between A and B, each occupied for an indivisible instant of time, how could a ball ever move between those points in a shorter or longer period of time? This, however, is not a problem with which I am concerned at this time. I do not present Hume's theory in order to endorse it.

of motion, will serve, I think, to account for Hume's propensity, at least in the Treatise, for referring to cause and effect as objects rather than events.[16] Events must, on Hume's view of time at least, be essentially static. Since none of the parts of time co-exist, only one, an indivisible instant, ever exists. For instance, given the instants t_1, t_2, and t_3, and an object X, then, if X is changing, we may say that at t_1 it is in a state Xa; at t_2, a state Xb; and at t_3, a state Xc. The change from one state to another will in every case be instantaneous, but, since on Hume's view, the instant and the change will both be very small, it will give the appearance of being gradual and continuous. It is understandable, then, that Hume should refer to objects as opposed to events when he speaks of cause and effect.

I have been concerned so far only to give an outline of Hume's view of the nature of time, space, and motion. My purpose has not been critical. I have been concerned to give only enough detail to provide an adequate basis for understanding Hume's view of causation. I shall now turn to that subject.

Hume, as every undergraduate philosophy student knows, defined causation as the constant conjunction of objects, and, if such a definition were entirely accurate, I too should feel compelled to accept his conclusions regarding necessary

---

16 A propensity which, it would seem, has not been adequately understood by some writers. For instance, G.J. Warnock

connection. It is my belief, however, that such a view of causation as Hume's cannot be accurate, and I shall attempt to show that this conclusion follows both from what Hume said about causation and from what he said about time, space, and, indirectly, motion.

Let us take the following example: If two billiard balls strike each other in a particular manner, they will both come to a complete stop. The ball's striking we call the cause; their stopping, the effect. In Humean terminology, these events would be the conjoined "objects" and, as different objects, they would be discriminable: that is, they would be two separate and distinct things (objects or events) which, when conjoined in a particular manner, constitute this particular cause-effect relation. As Hume would later put this, "Every event is a distinct event from its cause."[17] The difficulty with this view of causation and the relation between cause and effect is that it is impossible. In order to show this, it will be necessary first to give a brief outline of Hume's remarks on causation.

The first step in Hume's argument is to determine the relations between objects that give rise to the idea of

---

says of Hume that he referred to the different kinds of causes "vaguely and undiscriminatingly" as "objects" ("Hume on Causation" in Pears, D.F., ed., David Hume: A Symposium. London: McMillan & Co., 1963, p. 55.).

17 An Inquiry Concerning Human Understanding, Charles W. Hendel, ed., Indianapolis: Bobbs-Merril Co., 1953, p. 44.

causation. The first of these is contiguity. Every cause is contiguous with its effect. Such contiguity must be both spatial and temporal. As Hume puts it "nothing can operate in a time or place, which is ever so little remov'd from those of its existence."[18] Where such contiguity appears to be lacking, we find upon closer examination that they are connected by a chain of causes such that the effect is finally caused by an event which is contiguous with it.[19] In other words, making a distinction between remote and proximate causes, we may say that the remote cause is connected with the effect through a chain of causes, the last one of which is the proximate cause, and that the proximate cause is that event which is contiguous with and which produces or causes the effect. It is interesting to note, however, that immediately subsequent to this argument Hume argues for the further relation that the cause must be prior to its effect. He believes that to do away with the relation of priority of cause to effect would entail "the destruction of that succession of causes, which we observe in the world, and indeed, the utter annihilation of time."[20]. Hume's argument for this point is that, if every cause were "co-temporary" with its effect, there would be no succession "and all objects must be co-existent."[21] The argument here relies heavily upon the view given earlier that time is a succession of indivisible instants, and, I

---

18 THN p. 75.

19 THN pp. 74, 75.

20 THN p. 76.

21 Ibid.

believe, Hume's view of causation will stand or fall with his conception of time.[22]

Having made these two points, that cause and effect must be contiguous in space and time and that the cause must precede the effect, there is only one point which remains to be made to complete the definition. That, of course, is the idea of constant conjunction. But, from Hume's first definition of "cause", "An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedency and contiguity to those objects, that resemble the latter,"[23] it is clear that the notion of constant conjunction is reducible to that of contiguity in space and time. An object is a cause if the relation of contiguity always obtains between that object and its effect (and, of course, if it always precedes its effect).

This, of course, is only the briefest possible outline of Hume's theory of causation. I have intended only to make the following points, that, on Hume's view, cause and effect are separate and distinct events: they are contiguous: and the cause must precede the effect. This, I think, is just sufficient to provide an adequate basis for my argument, to which I shall now proceed.

Consider, for a moment, the motion of a billiard ball

---

22 For instance, if we assume that time is infinitely divisible, then, if the cause is prior to its effect, it could not be contiguous with it.

23 THN p. 170.

along a straight line extending from the initial position A to the final position J during the time period extending from t_1 to t_{10}. As was previously pointed out, a finite period of time on Hume's view consists of a finite number of instants. Let us suppose, then, that the number of instants here is ten. Then the number of points in the line AJ will also be ten and the ball's motion will consist in its being in position A at t_1, B at t_2,...J at t_{10}. Now, if we wish to account for the ball's moving from A to J, we may say that the cause was the ball's being struck by the cue stick. The effect will be, supposedly, its moving from A to J. When the ball is struck, however, at time t_1, it is in position A. It is not until t_{10} that it is in position J. Now, the only object or event which is spatially and temporally contiguous with and prior to the ball's being at J at t_{10} is the ball's being at I at t_9. It is clear, then, that if anything is the cause of its being at J at t_{10}, it is its being at I at t_9 and not its being struck by the cue stick at A at t_1. Furthermore, if we describe the object or event which is the cause of its motion, rather than its being at J at t_{10}, we find that this is not its being struck, but its being in contact with the cue stick at A at t_1. In other words, the cause is the tip of the cue stick's being at a point contiguous with A at t_1. The effect, then, will be, not its motion from A to J, but its being at B at t_2. However, if we consider the effect to be the ball's moving or going from A at t_1 to B at t_2, then the cause would not be

what happens at t_1, since that is a part of the effect, but at t_0. But at t_0 the cue stick had not yet contacted the ball![24]

It should be obvious that this is an impossible view of causation. To repeat some important points. Time is composed of indivisible instants, and a cause occupies the instant immediately preceding the instant occupied by the effect. We may return now to my earlier billiard ball example: Two billiard balls strike each other in such a way that they both come to a complete stop (to assume that the balls are compressible or elastic, while it would almost certainly be a more accurate view, would only serve to complicate, without substantially changing the issue). We may suppose that the posistions of these balls, when they strike, will be A and B, and we may refer to their positions prior to their striking by number subscripts. Now, let us say that at t_1 the balls will be in positions A_2 and B_2: at t_2, in positions A_1 and B_1: and, at t_3, in positions A and B. At t_4 they will, again, be in positions A and B. Now, if their striking is the

---

24 The argument here is very similar to that given, in a different form, by Justus Hartnack ("Some Remarks on Causality" Journal of Philosophy. vol. 50, 1953, pp. 466-71) as a refutation of Hume's theory of causation. Hartnack's argument fails to take account of Hume's theory of space and time, however, and, with that taken into account, as it is here, I take it to be obvious that Hume's theory of causation has not yet been refuted (since Hume need only reply that it is not necessary that the cue contact the ball. The cause is what happens the instant before, its approaching the ball.), although it has been made to appear very odd.

cause of their stopping and if their striking is defined as their being in contact, that is, as their occupying positions such that parts of each ball are spatially contiguous, then their striking occurs at t_3. But their stopping also occurs at t_3. It is at this point that their respective motions, that is, their successive occupation of different, linearly ordered points or places, ends. And, if the cause in this case is their striking and the effect their stopping, the two "events" are one and the same event occupying one and the same time. Note that we could not describe the cause as their moving and the effect as their being at rest. If the cause is their moving, the cause is not, then, contiguous with its effect, since their motion consists at least of their occupation of different places at successive instants prior to their stopping. In other words, their moving will consist of their occupying A_1 and B_1 at t_2, followed by their occupying A and B at t_3, and their being at rest will consist of their occupying A and B at t_3, followed by their occupying A and B at t_4. Since, in this case, a part of the cause is what happens at A_1 and B_1 at t_2 and a part of the effect, what happens at A and B at t_4, the two events are not contiguous (but here it must be remembered that we are speaking of these events as events rather than as objects, a move which is not really permissible in the context of Hume's theory), but a further and more important problem here is that their occupying positions A and B at t_4 is a separate and distinct

event or object from their occupying A and B at t_3, which is, in its turn, a separate and distinct event from their occupying A_1 and B_1 at t_2. There is, then, an intervening event between the events of t_2 and t_4 which constitutes part of the cause and part of the effect. In other words, the cause and effect are as follows: t_2t_3 (or A_1B_1 AB) and t_3t_4 (or AB AB). There is an event which is shared between them (their being at A and B at t_3) as an integral part of each, and the cause and the effect are not, therefore, separate and distinct, contiguous, or successive.

It might, of course, be objected that their stopping does not occur until t_4 or, in other words, that their stopping does not occur until an instant after their striking. But this is, I think, an obviously ad hoc objection made for no other reason than that it is necessary to bolster a particular theory, for we do ordinarily think that the balls stop when they strike and not an instant later. This objection, however, can be easily met. If we assume that the cause is their stopping or their being in contact at t_4, it is clear that there is no difference whatever between the cause and the effect except that some time has elapsed. The two objects (at t_3 and t_4) are identical; they are one and the same. But if this is denied the objection may be met in this way; their being in contact at t_3 is the cause of their being in contact (or, at least, of their not moving farther) at t_4, and their being in contact at t_4 is the cause of their being in con-

tact at t_5, and so on. Now, if we trace this chain of causes backwards, we find that every effect (as should be the case, given the rule of constant conjunction) has the same cause and every cause the same effect--except their being in contact at t_3, which is the effect of their being in positions A_1 and B_1 at t_2, and two very different causes will have precisely the same effect. Now, the only way to avoid the conclusion that cause and effect are, at least in this case, one and the same event, is to say that the cause of their stopping is not their being in positions A and B at t_3, but their being in positions A_1 and B_1 at t_2. In other words, it could be argued that the cause is not their striking at all but their moving toward one another (their being at A_1 and B_1 at t_2). On this view, it could be allowed that their striking and their stopping are one and the same event or two different ways of describing the event, and it would still be the case that their cause occurred an instant prior to them. Now, this objection appears to be compelling and in order to show that it is not it will be necessary to view the events in question in a somewhat different light. Let us say that, rather than the balls' striking being the cause of their stopping, ball A's striking ball B is the cause of B's stopping and that B's striking A is the cause of A's stopping. I believe it will be admitted, if somewhat reluctantly, that this is no more than just a different way of describing the event without changing it. Now, if the objection were accepted,

it would follow that the cause of ball A's stopping at t_3 is ball B's being in position B_1 at t_2. In this case, the cause is clearly antecedent to and temporally contiguous with its effect. It is not, however, spatially contiguous with it since the position B lies between the positions A and B_1.

THE NOTION OF CAUSE

The foregoing argument is directed solely against Hume. It is my belief, however, that essentially the same argument could be made against virtually all theories of causation. All or nearly all such theories, at least those that view causation as a relation holding between real events, presuppose the view that cause and effect are separate events, that the cause is an even which in some way produces or brings about some other event which is the effect. Of course, it is precisely this view of cause and effect, a view which is basic to our traditional (and Hume, however unorthodox his theories might have been in some other respects, lies, I think, directly in the mainstream of this tradition) and usual conception of causation, that creates and demands solution of the problem: How are two different events so connected that one can be said to be the cause of the other? The answer to that question, so puzzling to philosophers and physicists alike, I believe, is now clear. If we return to my billiard ball example, it is clear that, if the causal sequence is examined closely, no difference whatever can be discovered

between the event of the balls' striking and the event of their stopping. The proximate cause and the effect are one and the same event. Of course, my argument here rests upon a very carefully chosen example, and I cannot assert with certainty that the same analysis will apply to all examples. I can say, however, that after having devoted much thought to the subject I am unable to think of any examples to which such an analysis would not apply and which do not yield the same conclusion, that, in all cases, the event which is the cause and the event which is the effect are one and the same event. If my argument is conclusive, then it would follow that, as it has been traditionally understood, causation is impossible. If, then, causation is to have any meaning at all, the definitions of "cause" and "effect" must be made far more inclusive and our understanding of the relation between them radically altered.

For my own part, I think that any <u>complete</u> statement of a cause involves a proposition regarding the entire universe. For instance, to return to the billiard balls, it is obvious that their striking is not solely responsible for their stopping; to describe the event completely, it seems to me we must also specify all the conditions under which their striking occurs. Some are satisfied to say just that we must specify all the <u>relevant</u> conditions, but if we say as part of our description of any causal event that the conditions described are all the relevant conditions, we must, as a

necessary consequence, deny that any other events or conditions are relevant, which is to make a statement about the entire universe. But it is hardly sufficient just to deny that there are any other relevant conditions. If we are to know that our description of a cause is complete and accurate, we must know that there are no other conditions that are relevant. In other words, we must know that there are other conditions (if there are), precisely what those conditions are, and precisely what their relations, if any, to the event in question are. Failing this, we cannot say that the conditions described are all the relevant conditions, only that we know of no other conditions that are relevant, which, of course, would be tantamount to saying that we do not know whether the conditions described are all the relevant conditions or not. The cause, then, of a particular billiard ball's motion is not just its bing struck by another billiard ball: it is that, of course, but not that alone. The event of the billiard ball's being struck by another ball is an event which occurs in a certain context, a context from which it cannot, in fact, be dissassociated (although it may well be in fancy), and a context which, if it were substantially different, would occasion a different event. That context in which the event occurs, which, in a very real sense, constitutes the total event, is the entire universe, and a complete description of the event will include a complete description of the entire universe.

It follows that the complete statement of the cause of any event is impossible in practice, if not in principle. Again, this fact does not count as an argument either against causality or against causal entailment, but only as as argument against certain knowledge regarding the precise cause of any event. And if we accept the view that the complete description of a cause involves a complete description of the eintire universe (and the corollary that certain knowledge of a cause requires complete knowledge as to the state of the entire universe), yet deny causal entailment, we must also admit that the state of the entire universe at any instant is independent of the state of the entire universe at any other instant. It is for this reason that I would argue that the occurrence of any event which is uncaused is the beginning of an entire universe. If we accept the view that the complete description of a cause involves a complete description of the entire universe and deny causal entailment, it follows that the universe is at every instant a new, and not merely different universe without any connection with the universe which existed at the previous instant, which is a view that nobody accepts (although some may be blind enough to their own actions to think they do). It might be argued, of course, that there may be still <u>some</u> causal connection between the state of the universe at one instant and its state at the next; it could be argued, for instance, that the state of the universe at one instant is a necessary but not

sufficient condition for its state at the next. But this argument, I think, is simply absurd, for, if the present state of the universe is a necessary but not a sufficient condition for the next, it must ~~be~~ the case that something else is required to make it sufficient, but this something else, if it is not part of the universe at present, is nothing and, if it is nothing, the universe at the next instant is completely uncaused, for a cause, by anybody's definition, is not a partial cause (If X, Y, and Z are necessary and sufficient conditions for A, then each is a part of the cause of A. While X might be spoken of as a partial cause of A, in the sense that it is a part of the cause, it would not be a cause of A, and, if X and not Y and not Z and A, A would be uncaused).

The reader will recall that I suggested that, if causation were to be at all meaningful, we would have to make our definitions of cause and effect more inclusive and radically alter our conception of the relation between them. In the preceding two paragraphs I have been concerned with the former: in the remainder of this chapter I shall be concerned primarily with the latter and, as a result, with the implications of this change in our conception of cause and effect for determinism.

A TERMINOLOGICAL CHANGE

As a consequence of the principle of causal entailment, I would recommend a change in our usual terminology. Philosophers are accustomed to speak of antecedent and postcedent events. I would suggest "anterior" and "posterior" events. Antecedent and postcedent suggest that cause and effect are logically independent events separated by time; anterior and posterior, construed temporally, suggest that events are distinguished by time, that they are the same events distinguished, in a sense, by temporal position, as the anterior and posterior portions of a person's head are portions of the same thing distinguished by spatial position. Since the terms "cause" and "effect" are almost hopelessly bound up with our ordinary, philosophical conception of cause and effect as separate and distinct events, it would, no doubt, be best if these terms were, for the sake of clarity, dropped all together. These terms are, however, convenient and I shall have recourse to them occasionally.

THE CONCEPT OF DETERMINISM

What does the principle of causal entailment, as I have so far interpreted it, mean for the determinist thesis? Plainly, the determinist position will have to be restated as follows: Given the state of the entire universe at any instant B, there is another state of the universe at some

prior instant A, such that, given the state of the universe at A, its state at B will follow. The most significant change here I take to be the replacement of "must" by "will". "Must" suggests either that the connection between the two events is one of logical necessity, in which case it would be a timeless connection, or that the connection is, in some sense, one of requirement, but the universe can hardly be required to have some connection with the universe at some other time--no more than the Eiffel Tower can be required to be and to remain the Eiffel Tower (we can, of course, require that it continue to be called the "Eiffel Tower" or that it be painted and repaired occasionally to preserve it, but these requirements are not of the tower but of men). This usage of "must" would, I think, require the existence of three entities, the two events and some third entity in virtue of which the one event is required to follow the other. The usage of "must" as a logical connection arises, in part, from our thinking of cause and effect in terms of antecedent and postcedent, i.e., as independent events separated by time. No doubt, this conception arises in its turn from our traditional tendency to think of time as a succession of instants, but if we think of events as being related to one another as one instant is to its successor, I, too, would be reluctant to speak of necessary connections between events or of one event's entailing another. Such a connection would have to be a timeless connection; the state of the universe at any

instant would have to be contained in the state of the universe at the previous instant, as the conclusion to a valid syllogism is contained in the premisses, but this, plainly, is not the case. This difficulty can be overcome if we recognize the fact that time is not a succession of an infinite (or a finite) number of timeless instants.

SPACE, TIME, AND INFINITY

Any instant, however small, is the passage of some time. No two adjacent instants are independent; the one is simply the continuation of the other. Ordinarily we tend to reify "infinity"; I would suggest that "infinity" is simply an operational concept. It refers to the possibility of continuing certain mathematical operations, e.g., no matter with which number we stop counting, there is always a next higher or next lower number as the case may be. There is no such thing as an infinitely small number, but there is always the possibility of a smaller finite number. No more is there such a thing as an infinitely short portion of time, just the possibility of increasingly shorter finite periods of time. As an example, between any two points A and B there is a line. Now, the line AB, so long as it is not a chalk mark, a light ray, a corner of a room, or some such thing, is, in itself, nothing. It is, if you will, a conception or an idea; it is distance; it is the fact that, say, a ruler can be fit be-

tween the two points A and B: it is the fact that they are not the same point. In itself, it is nothing, and nothing may be divided for ever. If we divide it into two equal halves, we do not divide it into two equal parts. It does not have parts to begin with. If we put a ruler up to it, or, more accurately, up to the points A and B, and find that it measures twelve inches, we do not find that it has twelve or more parts. Nor do we find that the distance between the points A and B measures twelve inches: we find that it is twelve inches, or, more accurately, that A and B are twelve inches apart. Space, in itself, is nothing and, as nothing, it does not consist of parts. But time, too, is, in itself, nothing, and one instant is not, therefore, required or necessitated by the other. A person's head is one thing with anterior and posterior ends, yet an infinite number of lateral cross-sections may conceivably be made through it, which means just that no matter how thin the cross-sections are made there is always the possibility, in principle at least, of making them still thinner, and each cross section, when made, will occupy a finite portion of space with an anterior and posterior surface. Space, while it is in itself nothing, is defined by the existence of things. Time, too, while in itself nothing, is defined by the existence of things, but of things changing. It is not a succession; rather, it is a passing. To borrow a phrase from Aristotle, it is "a coming to be and a passing away." The events which we think of as cause and effect are

not separate events; that is not possible for reasons which I have given earlier. They are the coming to be and the passing away. They are the anterior and posterior ends of the same event.

TIME AND EVENTS

The concept of an event is typically restricted, in the domain of ordinary discourse, to the eventful, although it is not so restricted in the broad range of its actual meaning. No doubt, the fact that this restriction has been placed upon the concept in ordinary usage has contributed in a large part to the failure on the part of many to grasp the relation between cause and effect; we perceive some eventful occurrence and we assume that there must be some other, equally eventful occurrence which caused or produced it. The meaning of "event", however, cannot be so restricted that it is disassociated from the concept of change. Where there is change there is an event and, conversely, where there is an event, there is change. Aristotle, finding that time also could not be disassociated from change, concluded that it must be some aspect of change. Here he made an understandable mistake. Time is not an aspect of change, nor is it derived from change, nor, for that matter, is it something which underlies change. "Time" (note the quotation marks) is, in a sense, an aspect of change and is derived from change. But an event does not occur <u>in</u> time, nor is an event dependent in

any way upon time, for time and change do not exist separately. Nor, for that matter, are they distinct. They are one and the same. If, then, time is infinitely divisible, so, too, is an event. In itself, an event is nothing; it has no existence of its own. What exists is, of course, that which is changing.

An event, then, is a process of change, but, if a thing is changing, it must be the case that it is changing from something to something else. What it is changing from we may call the anterior event: what it is changing to, the posterior event. Now, I am not here suggesting any such thing as final causes. What I am trying to do is to elucidate the concept of "event", of "change." We can, I believe, say with certainty that, if a thing is now changing, it will be something else, that it will differ in some respect from what it is now. Of course, in saying that it will differ from what it is now, the implication is inevitably given that it is some thing now, that it is some discrete, determinate thing. But to assume that it is some definite thing is, of course, to deny that it is changing. That one can never step into the same river twice is appropriate to the view that time is a succession of indivisible instants and is compatible with change. It is not, however, compatible with changing, for on this view, it is true that one can never step into the same river once. The river is flowing; it is changing, and, as we step into it, it is continuallly different. An event is not different in some respect from what it was. It is a differing.

All this may well strike the reader as very inexact and vague, as, perhaps, a failure to achieve the rigorously disciplined mode of speech necessary for the precise communication demanded in philosophy, a failure which is due, no doubt, to laziness on the part of the writer or, perhaps, to some brand of mysticism. This, however, is not the case. the difficulty is that "now" is static. It is synonomous with "the present moment" or with "this instant." The concepts which we employ to discuss and to understand events with that degree of precision and specificity which we, as philosophers, all desire are, as they must be, static, but the events to which they refer are not. The events are changing. They are change, and it is not possible to speak about them in a way which reflects that change without speaking in participles (a mode of speech which is, admittedly, almost as obscure as speaking in tongues).[25]

---

25 These remarks may be better understood, perhaps, in terms of a principle of physics which I think they entail, the Heisenberg Principle of Uncertainty, which, I believe, states that it is not possible, at least for sub-atomic events, to determine both their position and velocity at the same time. Physicists generally account for this principle in terms of an instrument effect; that is, the effect of the instruments used to determine the positions and velocities of sub-atomic particles cease to be negligible where such small particles are concerned. But the uncertainty has nothing to do with instrument effects. If follows from the nature of an event and it is just as true of macroscopic as of microscopic events. If a thing has a definite, determinate position, it has, as a consequence, no velocity; it's position is not changing. If, on the other hand, its position is changing, if it has velocity, then it does not have a definite position. That the Heisenberg uncertainty

From the concept of an ongoing, continuous change, it is not possible to disassociate the concepts of becoming, of developing, of something which is, at the same time, "a coming to be and a passing away." A thing which is undergoing a currently ongoing, continuous change is always on the verge of becoming, it is never become, for once it has become some specific thing it ceases to change. So long as it is changing it is not a determinate thing: it is a that-tending-to-this but to the extent that it is a tending from the one to the other, it is neither. Consider, for example, the hour hand of a watch. Assume that the tip of the hour hand is exactly as wide as the hour markers; then, we may say that the hour hand is pointing to two o'clock when the edges of the hand are exactly congruous with the edges of the two o'clock marker and when the tip of the hour hand is as close as it can come to the inner edge of the marker. Now, let us ask: When does the hour hand point to the two o'clock marker? If the obvious answer, "When it is two o'clock (according to the watch)", is given, it need be pointed out only that it is two o'clock according to the watch only when the hour hand is pointing to the two o'clock marker. The obvious

---

principle seems to be confined to microscopic events has nothing to do with a greater instrument effect at that level: it is due to a necessary lack of precision in our instuments, a lack which is far more apparent at the microscopic than at the macroscopic level. Of course, I do not mean to deny that there is an instrument effect, only that the instrument effect is solely responsible for the uncertainty.

answer, therefore, does not answer the question. The only answer is: When the watch has stopped running at two o'clock. So long as the watch is running, the hour hand is not pointing to any definite time, and, if the watch is running, the correct answer is: Never. It ceases to point to the two o'clock marker as it begins to point to the two o'clock marker. It never points to the two o'clock marker; at two o'clock it passes the two o'clock marker. This may, perhaps, be better understood metaphorically: a bubble floating in air may be continuously changing in shape. Its state at any one instant will be different from its state at any anterior (antecedent, previous) instant, yet there is no instant during which we can say that it has exactly this shape and that it is not changing, although we may well imagine an instant so short that the change in its shape would be "negligible". This is, of course, only a metaphor, but I think it not at all unlikely that the universe is, in more ways than one, only a continuously changing bubble.

DETERMINISM AND A REFORMULATION OF THE PRINCIPLE OF CAUSAL ENTAILMENT

Determinism, as I pointed out earlier, is the view that every event has a sufficient cause. The principle of causal entailment, on the other hand, is the view that every event is a sufficient cause. When I first formulated the principle in this way, I noted that this formulation was

somewhat inaccurate. It should be obvious now what those inaccuracies are. In the first place, "cause" cannot be understood in the traditional sense as an event which somehow brings about or produces another, different event. But a further inaccuracy rests in the phrase "every event". Macroscopically, there is only one event, the universe. What we ordinarily think of as events, as causes and effects, are parts of that event, and, while they may be considered apart from their total context, they do not exist apart from it. (This will be taken up at greater length in the following chapter). From this it is clear that a more accurate statement of the principle of causal entailment would be: Given any event A, where A is the total event, the event A is a process of change, of becoming some other event B, where B is the total event, and where B is the posterior event of the anterior event A. This statement, admittedly, lacks somewhat of the beauty of precision; I can but hope that this lack is made up for by its contribution to the more sublime beauty of understanding.

The concept of determinism is logically indefensible and unjustifiable, whereas the principle of causal entailment, with the proper qualifications, is, I think, quite sound. Thus, given a complete description of any event during a time period t_1 (a description which, as should be obvious from the preceding discussion, it would not be possible in principle to obtain), it would be possible, in principle, to predict a

complete description of any event during any subsequent time period. The "effect", in other words, is discoverable in the "cause". It would not, however, be possible in principle to retrodict from such a description any event prior to t_1. That is, it is not possible on the basis of a description of any event over a "finite" period of time to infer that the event had, itself, a cause. There is always the possibility that the universe has only just begun.

Can we say, then, that the thesis of determinism is true? I do not know. Certainly, it is not entailed by the principle of causal entailment. But saying or assuming determinism is true is a very different thing from its being true, and there are certain difficulties which will arise if we assume it to be false. Some of these difficulties will be dealt with in the following chapter.

IV

DETERMINISM IN HUMAN AFFAIRS

Philosophers are not, for the most part, concerned to refute or even to challenge determinism or causal entailment with regard to the events of the world. The dispute, where there is any, centers around those events that comprise human activity. That is, the argument is that human actions are not, or need not be caused or, in other words, that a man's future actions are not a consequence, either directly or indirectly, of the current state of the universe. This view is denied by determinism: it is not denied, at least not directly, by the principle of causal entailment. However, we ought to ask what such a view of human activity would mean in light of the principle of causal entailment.

In the first place, it must be pointed out and re-emphasized that the complete description of a cause or of an event requires a complete description of the entire universe. Am I a part of the universe or am I not? If I am, then such a description of the current state of the universe would require a complete description of me as well, as a part of the universe, and we must then say that that future activity of mine which is uncaused is independent of my current state and my current activities. However, most persons are willing

to admit that at least some of my activities are caused, but this is a peculiar view, for, if it is accepted, it would seem to imply that I am a most schizophrenic sort of being, part of which is dependent upon the universe and itself for its activities and part of which is not. Perhaps, there is nothing logically inconsistent in such a view but it is certainly rather unbelievable. If we take the view that I am not a part of the universe, this becomes even more troublesome. On this view, the idea that some of my behavior is caused and some not would be virtually unintelligible. In the first place, we should have to say that I am not connected with the universe at all, that all of my actions and my very being are totally irrelevant to the events around me, that I should have no influence upon them and they none upon me, that, for me, they do not exist or that, if they exist, they exist as a part of me. In other words, it seems to me that this view implies solipsism in virtually its purest form. However, it would not be a solipsism asserting the non-existence of other entities; rather, it would be the case that other entities would be separate and distinct universes, that, in my universe, I, and I alone, exist. Such a view as this may be difficult to accept, certainly I do not accept it; however, such a view would enable us to understand how I might be a being which is the cause and the sole cause of its own behavior. That I would be such a being follows from the principle of causal entailment. Accepting the view that I am

a universe unto myself, then a complete description of any event would be a complete description of myself, of me. Taking the view that every event is a sufficient cause, it would follow that I, and I alone, am a sufficient cause for my own continued being and my own activity. But what it would mean in this case to say that some of my actions are caused and some not, I do not know. Clearly, those which are not caused would not be caused by me, since I would be the cause of all those of my actions which are caused.

But few people, I think, would wish to take this view. Those who do cannot be refuted; solipsism as a metaphysical doctrine admits of no refutation; the alledged refutation would after all, be just a part of me, an argument, as it were, to myself which would be just my own invention. Nevertheless, I assume that solipsism is false. I have no real grounds for this assumption (unless a simple inability to believe it can be admitted as grounds for calling it false); neither does anyone else. I assume, therefore, that there are other entities than myself and that I am a part of the universe (perhaps, I should say "of this universe"). In other words, I assume that a complete description of the universe would include a complete description of me as a part of that universe. From this it would appear that there could be no part of me which is not a part of the universe, since it would then not be a part of me. All of my future actions, then, are caused in the sense that they are causally entailed

by me as a part of the universe, since any action which were uncaused would not be my action.

Still, it might be possible for my arm, at some future time, simply to rise up in the air without cause. That is, it might be possible for such an event to occur without there having been any other event such that it would occur. However, I could not in such a case say that I had raised my arm, that is, my arm's rising could not be an event ascribable to me or to any other part of the universe. The principle of causal entailment does not exclude such a possibility as this. It does, however, require that it be qualified. We may not say just that there were no events such that my arm would rise up: we must also say that there were no events such that my arm would not rise up. In other words, we must say that what I am doing, the current state of my being, is irrelevant to any future state of my being, that the current contraction of my muscles which keeps my hand on this piece of paper, that gravity, which, in the absence of any counteracting forces, also serves to keep my hand in place, and so on all are suddenly and inexplicably irrelevant to the activities of my arm.

The state of my being at this moment is such that I am writing and that I intend to continue writing until this paragraph is finished: that is, my neurological activity, muscular activity, and surrounding events all are such that they entail my continued writing or, at least, my intention

to continue writing. If, despite this, I were suddenly to stop writing, if my arm were suddenly to rise up in the air, and if such an event were uncaused, that is, if there were no event such that it would happen, then would we not have to say that suddenly, inexplicably, all those events such that it would not happen had ceased and that what I intended I should do, that what the state of the universe entailed my doing was irrelevant to what I did, in fact, do (if one could speak of one's "doing" such an action). Furthermore, if such a thing could happen at some time when there is no cause for it, when there is no event such that it would happen, and when there are events such that it would not happen, then should we not have to say that it could happen at any time? If so, if such an event could happen at any time, does it not become simply a matter of chance such that, no matter what the state of the universe, there is nothing about that state such that such an event will or will not happen? If we say that it is purely a matter of chance, do we not have to say also that the state of the universe at any future time is independent of the state of the universe at the present time and that it is independent of the state of the universe at any previous time, just as the state of the universe at present is so independent, and if we say this do we not have to say also that the <u>apparent</u> uniformity of nature is not only rationally inexplicable but actually contrary to reason, since, if there are more things than one that could happen and

nothing constrains any of them to happen or to happen in a particular way or at any particular time, we should expect them all to happen in no particular order and without rhyme or reason?

I have finished that paragraph and my arm has not once risen inexplicably up into the air. My intentions and my expectations have been fulfilled. No better empirical proof may be given for the principle of causal entailment with regard to human affairs. My memory may be fallible: it may not have been caused: I may have come to be but an instant ago: indeed, my beginning may be at this moment, but, to admit this as grounds for denying any philosophical doctrine, even when the doctrine denied is that which asserts that a belief in such a possibility as fact is false, would be to deny the possibility of philosophy and, to that extent, self-refuting. I admit such a possibility as a logical possibility which, if recognized or thought to be true, would do away with the possibility of logic and all human reason, but I do not believe it, as no man believes it, because it cannot be believed.

I submit, then, to the view that determinism is true. This belief is not irrational: it simply cannot be proved. The two are not the same. There is no evidence for the truth of determinism which is not in itself in question. Neither is there evidence against determinism. The question is not really which is true; we cannot know that. The question is, given

what we do believe and the implications of either view for those beliefs, which can we accept? If we accept the view that determinism is true of human actions, then we must give up the belief that certain human actions are undetermined and, perhaps, some other beliefs attendant upon that. If, however, we accept the view that determinism is, with regard to human actions, false, then I submit that we should have to give up all or virtually all of our beliefs, since, on this view, nothing done by a human being would have any connection, except perhaps an accidental one, with anything else, and all beliefs about the universe and about ourselves would be irrational.

B

This section will be made as brief as possible. I intend here only to expand the major argument for determinism in human affairs, to point out briefly the way in which human behavior is determined, and to clarify one often misunderstood aspect of determinism.

It will be well, first, to set out the most basic facts at our disposal and then to determine the implications of those facts. In the first place, we may say that a man is a being of a particular sort. While there may be some disagreement about what sort of being he is, few would disagree with the view that he is a being of a particular sort and distinct from other beings such as bears, lions, and

elephants. From this fact alone, from the fact that man is a being of a particular sort, we may infer that, to the extent this is true, determinism is true. If determinism were entirely untrue, then we should expect man to be a being of no particular sort and to be indistinguishable from bears, lions, and elephants. But there are three and one-half billion men in the world and they are all beings of a particular sort: furthermore, all of their parents, for they all had parents, are beings of the same particular sort, and we may expect, with all the assurance human beings can muster, that, short of some such thing as a nuclear catastrophe, all of their children and their children's children will be beings of the same particular sort. Each of these men resembles the others to a greater or lesser degree and always to a greater degree than a bear, a lion, or an elephant. If nothing determined such things as these, then the fact that the physical appearances of men are so similar would be not only unexpected, but totally unintelligible.

Few philosophers, however, would wish to deny that such thihgs as phsical appearance are determined, but we need not stop here. We note also that nearly all men like the taste of sugar and dislike the taste of lemon (at least, of pure lemon), that nearly all men enjoy such things as sexual stimulation and dislike such things as placing their hands in fire. The list of examples could be extended at some length: these are not all of our basic likes and dislikes, but my

point should be clear. Given the statistical significance of such similarities, it is simply unreasonable to assume that they are not, in some way, determined. Granted that it may be all a matter of sheerest coincidence, the odds against it are astounding, and, given those odds, to assume that it is simply a matter of coincidence and undetermined would be unreasonable and irrational. We are forced to assume, I think, that there is something about men such that they like certain things and dislike certain others and that these likes and dislikes, or at least some of them, are simply given--in the sense that possessing them is part of what it is to be a man and that he possesses them for very much the same reason that he possesses any of the other properties that are a part of what it is to be a man (e.g., genetic endowment). This is not, of course, to say that a human neonate is born with a liking for sugar and a disliking for lemon. It is to say, rather, that a human neonate, because it is a human neonate, is such that, in the absence of any other relevant experience, it will like, or will exhibit signs of liking sugar when it first tastes it and of disliking lemon under the same circumstances.

This, I think, is the major argument for determinism. It is what I shall call the statistical argument. It does not prove that determinism is true: it proves neither that men's physical appearances are determined nor that their basic likes and dislikes are determined. What this argument attempts to

establish is simply that, given the statistically observed regularity in such things as physical appearance and basic likes and dislikes, it is simply unreasonable to assume that they are not determined. The odds are overwhelmingly against such a view. In the minds of some, such an argument as this might not be sufficient, but precisely the same sort of argument is used to show that non-human events are determined, and if such an argument is not sufficient to show that such things as physical appearance and basic likes and dislikes are determined, neither is it sufficient to show that the orbits of the planets, the flow of rivers, and so forth are determined. To assert, then, that the world of non-human events is determined and that such things as a person's appearance and basic likes and dislikes are not determined, when the same kind of evidence is offered for each, would be irrational.

Human beings, however, are not entirely similar and this may provide some ground for the belief that their likes and dislikes and, more importantly, their actions are not determined. I shall not attempt an argument against this view. It is sufficient for my purposes to show that such differences can be accounted for if determinism is true. We may begin by asking why it should be the case that men have, as a part of their natures, certain "inborn" likes and dislikes. The answer to this question is, I believe, fairly obvious. A man who possessed no natural likes and dislikes

would very quickly die. That certain kinds of food are nutritious and certain others poisonous and that, of those that are nutritious, men must eat a certain minimum amount and that, of those that are poisonous, men must eat no more than a certain maximum amount, and all this simply in order to live, are facts which no one, to my knowledge, would deny (at least, not for very long). But, without benefit of training, how would a man know what things are or are likely to be nutritious unless he had some unlearned liking for foods with a certain kind of taste, e.g., sweet? It is no good to argue here that men do have training in such matters: such an argument would simply beg the question. If my earliest training regarding edible foods came from my mother, why should I have followed her instructions--because she would give me approval if I did or spankings if I did not? Then why should I have cared about these? For approval, in whatever form, to reinforce some particular behavior, it must be the case that I care, in some respect, for that approval: i.e., it must be the case that approval is something I want or like. If approval were something about which I did not care one way or the other, then, to state the obvious, approval would be something about which I simply did not care, and whether I received it or not would make not a bit of difference to me, and if the same were true of disapproval, again, in whatever form, then how could I possibly have learned anything from my mother? Even assuming that I might

just possibly have learned under these conditions what foods were edible, how should I have learned which were desirable and why should I have acted in accordance with what I had learned? That these basic likes and dislikes, and others like them, for particular tastes, maternal affection, etc., determine the actions of animals and young children will, in all probability, be granted. The major objection arises with actions that do not seem to be determined by likes and dislikes or, on the other hand, to be contrary to our basic likes and dislikes. I have already argued in chapter two that this sort of objection is based upon a misunderstanding. What I now wish to do is give, in brief, the psychological explanation for the development of dispositions to act in ways which run contrary to our basic likes and dislikes.

It is possible to teach a chimpanzee to crush its own skull with its fist. It is possible to teach a child to wound itself. It is possible to teach a man to place himself in grave and imminent danger. To teach a chimpanzee to crush its own skull, simply give it a banana chip when it happens to touch its head. Wait to give it another until it touches its head with a little more force. Repeat this procedure, gradually increasing the requirement, until the chimpanzee knocks itself unconscious. To teach a child to wound itself or to engage in self-destructive behavior, simply ignore it until it hurts itself, then comfort it with affection. Repeat this procedure until the infliction of pain is deliberate. To

teach a man to place himself in danger, reward him with approval whenever he makes a remark suggesting that he values the lives of his fellows; reward him whenever he does something for his fellows, especially when it involves some element of self-sacrafice; punish or ignore him when he makes a remark or exhibits some behavior suggesting that he values his own life over the lives of his fellows. Praise him when he does anything dangerous to himself; belittle him when he refuses to. Repeat this procedure over the course of several years, then send him off to war. The process described in these three cases is that of operant conditioning. In the first case, the reward (reinforcer) is primary: it is unlearned and unconditioned. A chimpanzee could possibly be taught to dislike banana chips; it need not be taught to like them. Affection, in some of its forms (bodily contact, soft tones, etc.) is primary; in others, secondary (approving words). Praise is secondary; we learn to like it (assuming it is verbal).

The examples I have given here are not imaginary, although the third is a hypothetical reconstruction of the actual course of events. None of these procedures would work, however, if the reinforcers used were not reinforcing. If chimpanzees did not like banana chips, if infants did not like affection, and if men did not like praise, rewards consisting of these things would have no effect. Obviously, if this account is an accurate account of the development of

such behaviors, it is an adequate account of the development of behaviors which, in themselves, run contrary to our basic likes and dislikes. Of course, the same sort of account may be given to explain the dissimilarities in the behaviors of individuals. For instance, very few children engage in self-destructive behavior, but, then, few parents (whether consciously or unconsciously) differentially reinforce it.

An objection levelled against this sort of account, with regard to human behavior at least, is that human behavior is rule-governed. This objection, too, is based upon a misunderstanding. A rule is nothing more than a formulation of the contingencies of reinforcement; ordinarily, acting in accordance with such formulations is, in the long run, more reinforcing than not. Still we must learn to act in accordance with rules. Much of a child's education, at home and at school, is devoted to teaching him, not what the rules are, but to behave in accordance with them. "Always flush the toilet after use" is a rule which the child must be taught and which he must be taught to obey--by punishment when he does not, approval when he does. Later on the natural consequences of not following the rule will constitute sufficient reinforcement for maintaining this particular bit of rule-governed behavior.

Not all rules must be made explicit, however. Language learning, for instance, consists in a very large part of learning rules, yet a child seldom receives an

explicit formulation of those rules until long after he has begun to speak in accordance with them. At a very early age an infant utters sounds indiscriminately. Gradually he is taught (by differential reinforcement) to utter certain sounds (e.g., English phonemes) to the exclusion of others. He is taught by successive approximations to utter simple words (Dada, Mama) which are very easily and strongly reinforced when used ("mama" is reinforced by the appearance of mama or by her delight if it is uttered in her presence). Gradually other useful words are taught, followed by word combinations until, finally, the child "gets the idea" and finds learning new words and new combinations to be reinforcing in itself. He learns the rules implicitly. He learns, for instance, that words are of different types and, at the same time, that words of different types are used in different ways. He learns these rules not by being taught them directly, but by being differentially reinforced for following them (For instance, when "You give me cookie" is meant, "Me give you cookie" is likely to have precisely the opposite of the desired effect).

Rules can and do affect the outcomes of our rational choices. How this is so should by now be clear. Let us examine the following hypothetical reconstruction. A child, we may safely assume, finds the company of its mother to be reinforcing: that is, its behavior is reinforced by her presence and, possibly, punished by her absence or, at least, by the

withdrawal of her presence. Since the infant's bond to its mother is very strong, much of its early behavior can be shaped and developed simply be her carefully (but not necessarily consciously) reinforcing desired behaviors by her presence and by her punishing undesirable behaviors by her absence. Thus, a mother could extinguish much crying behavior by punishing it, by simply ignoring it, and by differentially reinforcing quiescent behavior, by keeping the baby company only when it is quiet or, at least, not crying. A baby is also reinforced and punished by certain of its mother's behaviors; possibly these likes and dislikes are "inborn", possibly they are learned by being associated with the mother's appearance and withdrawal. For instance, the mother's smile is reinforcing, her frown punishing. By the association of her smile with her behavior, the child will learn that a smile is a portent of good; a frown, of bad. As his experience and association with other people increases he will find that this is true of most people; when they smile they treat him well; when they frown, not so well. Of course, a child will like those people who treat him well and dislike those who do not. In all probability, he will come to like those people who smile and dislike those who do not. When he meets new people, he may well determine whether or not he likes them on the basis of nothing more than their smiles, and, at a later date, he may well formulate a rule to the effect that a good man smiles. His evaluation of a man, then, may well

depend upon whether or not he smiles. Now, this is the sense in which I believe it may be said that evaluations are causally entailed by facts. That a child should like men who smile and think them good and that he should dislike men who frown and think them bad is a consequence of his nature, of his being the sort of being he is, and of his individual experiences with smiling and frowning men. Of course, I cannot relate the exact causal conditions leading up to and including his evaluation in anything like detail. Furthermore, the example which I have given here is, admittedly, exceedingly simplistic in its lack of detail; however, it is not necessary for my purposes to give all the details. Those are matters for the psycholosgist to ascertain and will probably differ for each individual case. But the process, as I have described it, is sufficient, in its general outlines, to explain the sense in which values may be causally entailed by facts and to show how such entailment is possible. Other, more complicated examples will differ, I am convinced, only in the fact that they are more complicated. The principles behind them will be the same. As to whether or not I am correct in this belief, it is for the psychologist to determine.[1]

---

[1] To those who have been unfortunately guiled into skepticism (a mild word) about Skinnerian (operant) psychology by Noam Chomsky's review ("Verbal Behavior. By B.F. Skinner." Language, 35 (1959), pp. 26-58) I would suggest a reading of Kenneth MacCorquodale's reply ("On Chomsky's Review of Skinner's Verbal Behavior." Journal of the Experi-

It is an unfortunate fact that many philosophers who argue against determinism do so in the belief that determinism is the view that whatever I do is a matter of what happens to me, that I am caused to do things. This understanding of determinism has nowhere been better expressed than in the following passage from Thomas Reid:

> The law of nature respecting matter, is grounded upon this principle, that matter is an inert, inactive substance, which does not act, but is acted upon: and the law of necessity must be grounded upon the supposition, that an intelligent being is an inert, inactive substance, which does not act, but is acted upon.[2]

The picture of determinism presented in this and in similar passages, which may be found in almost any philosophical work written in favor of free will (where "free" is taken to entail "undetermined"), appears to be that of a man in chains being dragged about and forced to follow their pull. Of course, the chains may be attached far more subtly than this; they may be attached, not to our bodies, but to our wills, determining directly not what we do but what we wish to do. Such a view is not entirely inaccurate. Its chief fault lies in its creation of a gap or a distinction between me and the causes of my behavior (a distinction which is, probably, a consequence of the usual distinction between cause and effect).

---

mental Analysis of Behavior, 13 (1970), pp. 83-99.

2 "Some Arguments for Free Will." In Dworkin, Gerald, ed., Determinism, Free Will, and Moral Responsibility. Englewood Cliffs: Prentice-Hall, 1970, p. 66.

Analogously, we could say that these writers view the relation between my behavior and the cause of my behavior like that between the motion of a caboose and that of the engine. But a better analogy and a more accurate view would be the relation between the motion of the engine and its cause. The cause of the engine's motion cannot be separated from the engine. But, it will be replied, it takes a man to turn the key. Yes, it does, but what does this prove? We could, afterall, easily imagine a train run by a computer. Yes, it will be replied, but it takes a man to program the computer. Well, does it? And if it does, then we may ask: Where is the man who programmed the acorn to become an oak, who programmed the zygote to become a fish, a monkey, or a man? The motion of the engine is self-caused to the extent that the cause of the engine's motion cannot be separated from the engine itself, but, while the cause of the engine's motion cannot be separated from the engine, neither can it be identified with the engine. We cannot say, for instance, just that the pull of the engine is the cause of the engine's and, hence, of the caboose's motion; we must also include the push of the tracks (or their resistance). The tracks may, to be sure, be distinguished from the locomotive's wheels, and a distinction may, perhaps, be made between the push of the wheels and the push of the tracks, but they cannot be separated; that is, they cannot be made to act independently of one another. The two conditions are integrally

related as a part of the cause of the train's motion or, more accurately, as integrally a part of the train's motion. If we wish to know why the train is moving rather than standing still, that the wheels of the locomotive are turning and pushing against the tracks may be sufficient to answer our question or it may not. Whether or not it is a sufficient answer will depend on our interests and our purpose in asking. A satisfactory answer may have to include also the fact that the engine is running; it may have to include the fact that the engine is connected, via a drive shaft and transmission, to the wheels. Depending on one's interest it may have to include the fact that explosions are occuring at regular intervals in the engine's cylinders, that oxygen is entering by way of valves and what not into those cylinders together with fuel and so on. It may have to include the fact that the brakes are not on, or, if they are on, that the torque of the engine is sufficient to overcome their resistance. This process could go on until the entire universe had been described and no one of the conditions mentioned would be the cause of the train's motion. The train's motion is constituted by all of them and all of them together.

Precisely the same thing is true, if deter-inism is true, of a man's actions. There is no one event or set of conditions which act upon a man to produce some action. Events no more act upon me than I act upon them. To say that my leg jerked because it was struck by a hammer may be sufficient

for the layman: it may not be for the physiologist, and the physiologist's account may not be sufficient for the biochemist, nor his for the physicist. In the end the complete description of my action will include a complete description of me as a part of the universe and, of the conditions described, no one and no single group of them will constitute the cause of my action. My action will be constituted by all of them and all of them together. I am not acted upon by events; I am integrally a part of them.

Again, I do not make this point in order to deny the truth of this view of determinism, but to point out that it often rests upon a misconception of the relation between me and the causes of my behavior. This view is accurate, however, to the extent that it shows my behavior to be the outcome of circumstances none of which are, in the final analysis, completely under my control. But need it follow from this that my behavior is, therefore, under the control of something other than me? I think it plain that it need not follow. My behavior is, rather, under the control of both. At the moment, I control my pen, but so does the paper and, furthermore, so does my pen control me. To say that I control my pen is not, however, inaccurate. It would be inaccurate only if it were taken to mean that I am, in some sense, in complete control of my pen, that I and my movements are the sole and sufficient condition for everything done or accomplished by my pen. The same would apply to the statement that my pen

controls me. This statement, as it stands, is not inaccurate, but it is clearly inaccurate if it is taken to mean that my pen is, in some sense, in complete control of me. The statement, then, that what I do is, if determinism is true, not under my control is in one sense true but in another sense false. It is false if it is taken to mean that I, as a distinct and individual being am under the one way direction or control of things other than myself. In other words, if A and B are distinct things which, taken together, constitute C, it would be true in one sense that the existence of C is under the control of A since, if B and not A, then not C, but in another sense to say that C is under the control of A is false since, if A, it does not follow that C.[3] If I am a being such that I like the taste of sugar, then it might be possible to control some of my behavior, under certain conditions, by

---

3 This argument will also serve as a reply to the standard objection to soft-determinism. Soft-determinism may be taken to be, roughly, the view that determinism is true and that I am the cause of (some of) my actions. The objection is, roughly, that what I do is caused, if determinism is true, by events which are, eventually, antecedent to me or to my doing anything (such as "willing") which results in my performing some particular action. Since what I do, then, is not under my control, I am not the cause of my actions. Now, this objection is, as it stands, patently absurd, for, if an event is not a cause simply in virtue of the fact that it was itself caused, then it would seem to follow from determinism, the doctrine that everything is caused, that nothing is caused. But if the objection could be taken in such a way that it would not have this implication, then the argument given in the text would apply: I am the cause of my actions, but not the sole or the complete cause, in the sense that I am integrally a part of the conditions leading up to and including my acting as I do.

reinforcing it with sugar. In such a case, it would be an accurate description to say that my behavior is under the control of some particular set of contingencies of reinforcement, but such a description would be inaccurate if it were continued with "but not at all under my control." The conditions which constitute me and the conditions which constitute the contingencies of reinforcement together constitute the event or part of the event which constitutes my behavior. If, for some reason, we must say that one or the other controls, which we choose will depend, as much as upon anything else, upon our point of view.

V

RATIONAL CHOICE, DELIBERATION, AND DETERMINISM

At the outset of this thesis, I stated that I would be concerned to examine two arguments, those offered by Carl Ginet and Richard Taylor, to the effect that the concepts of choice and determinism are inconsistent and that, therefore, determinism must be assumed to be false. It should be pointed out now that neither of these two men makes the claim that determinism <u>is</u> false nor that they have proved it to be false. An alternative conclusion to their arguments, as each admits, would be that men do not make choices.[1] On the other hand, each of these two men is convinced, and I am inclined to agree, that the belief that men do make choices is far more basic to our usual conception of ourselves and our place in the world than any philosophical theory such as determinism, and, while the fact that "choice" is such a fundamental concept does not prove that men make choices and certainly does not prove that determinism is false, it would be simply irrational, in the face of the relative importance which we attach to these two concepts to abandon the more fundamental in favor of the less.

By this time the reader should be aware that I am not

---

1 Actually, Taylor's argument is not concerned so much with choice as with deliberation, but this will be taken up later.

in agreement with the position taken by Ginet and Taylor. In the first place, I am not convinced that their arguments are sound, and I shall offer my own counter-arguments to show that they are not. But, in the second place, I am not convinced that, even if their arguments are granted, we should be forced to choose between the beliefs that men do make choices and that determinism is true. To know, with all the certainty possible in human knowledge, that men do make choices is not necessarily to know precisely what it is to make a choice, and a proper analysis of the concept of choice or, at least, of rational choice, may show that all that need be abandoned, if determinism is true and if it is inconsistent with the belief that men make choices, is a particular, erroneous conception of what it is to make a choice. At least this is what I believe and what I shall attempt to show. I shall begin with Ginet's argument and I shall attempt to show (1) that, if the first part of his argument is correct, the premisses necessary to establish the truth of the second part cannot both be true, (2) that, on the basis of that part of Ginet's argument which may be accepted, determinism cannot be taken to imply that, given complete knowledge of the present state of the universe, it would be possible in principle to predict everything that will occur in the future, and (3) that Ginet has an erroneous conception of "choice". With regard to Taylor's argument, I shall attempt to show (1) that it is liable to the same objections as those

levelled against Ginet and (2) that it rests partly upon an equivocation in his use of "cause" or "sufficient condition" and in his use of "it is up to me."

GINET'S ARGUMENT

In his paper "Can the Will Be Caused?"[2] Carl Ginet argues that it is "conceptually impossible" for a choice, a decision, or a volition (each of which he takes, for the purposes of his argument, to be substitutable for "will") to be caused. That this is so, he says, follows from the two following propositions:

> (A) It is conceptually impossible for a person to know what a decision of his is going to be before he makes it.
>
> (B) If it were conceptually possible for a decision to be caused, then it would be conceptually possible for a person to know what a decision of his was going to be before making it.[3]

Ginet's argument is devided into two parts; in the first part he argues for the truth of (A) and, in the second, for (B). About proposition (A) I shall say nothing except that I accept both the proposition and Ginet's argument for it. Ginet attempts to argue for the truth of (B) in roughly the following way: He argues that, if it were possible for a choice or a decision to be caused, then it must be logically

---

2 In Dworkin, Gerald, ed., Determinism, Free Will, and Moral Responsibility. Englewood Cliffs: Prentice-Hall, Inc. 1970, pp. 119-126

3 Ibid. p. 120.

possible for a person to know in advance what a decision of his would be and this would be possible if the two following conditions were met:

> (1) The decider knew prior to his decision the causal law that circumstances of the kind that were going to cause it are always accompanied by a decision of that kind.
>
> (2) The decider knew prior to his decision that circumstances of the required kind existed or would exist.[4]

For the sake of brevity and clarity, I shall rephrase these two conditions as follows:

> (1a) The decider knew the causal law that some set of conditions C is a sufficient condition for a choice X.
>
> (2a) The decider knew at a time t that C would obtain at a time t_1.

Given that both (1a) and (2a) are true, then, according to Ginet, the decider would know at a time t that at a time t_1 he would make the choice X, which would be inconsistent with (A). But the question is can both (1a) and (2a) be true?

The problem with Ginet's argument, at least as I understand it, is simply this: If, according to proposition (A), it is conceptually impossible for me to know in advance of my choice what I will choose and if I am to choose between some set of alternatives X and Y, then, assuming that I am going to choose X, I must not know prior to such time as my choice is actually made that I will choose X. In other words, proposition (A) states a necessary condition for the event of

---

4 Ibid. p. 123.

my choosing. Accordingly, if there is a set of conditions C sufficient for my choosing X, then, among the set of conditions constituting C, there must be some condition such that I do not know that I shall choose X and, if C, then either I do not know prior to the time that I have made my choice, the causal law that the set of conditions C is followed by the choice X (condition 1a) or I do not know that C (condition 2a). Given that (A) is true, then, it follows that either (1a) or (2a) must be false.

While Ginet argues that we cannot deny the possibility of conditions (1a) and (2a), he makes the mistake of arguing for them separately. He argues that we cannot deny the possibility of (1a) and he argues that we cannot deny the possibilty of (2a), but he does not argue that we cannot deny the possibility of (1a) and (2a): that is, he does not argue, as he must if his case is to be established, that it is possible for both (1a) and (2a) to be true at the same time. That a decider could know that a given set of conditions, C, obtains is not, by itself, inconsitent with his making the choice X, nor is his knowing the causal law that C is followed by the choice X inconsistent with his choosing X. What would be so inconsistent is his knowing both that C obtains or will obtain and the causal law connecting C with the choice X. Condition (1a), since it states a causal law (whatever that may be), Ginet takes to be a universal proposition, and he argues that "If a completely universal proposition can be

known by anyone, then it can be known by everyone."[5] Whether or not this is true, I shall not attempt to say; however, if it is assumed to be true, and if it is assumed that I know it to be true, then I know, too, that I cannot know that C obtains, if and when it obtains, for at least one person, namely, myself. I could know, perhaps, that condition C obtained for someone else (which would be a particular proposition and not, therfore, entailing that he could also know), but I could not know that condition C obtained or would obtain for me. Ginet's argument against this implication of (1a), that if I know (1a), then I know that I cannot know that (2a), is decidedly inadequate. He argues,

> How could the possibility of the second condition...be excluded? One can, of course, describe a set of circumstances that it would be logically impossible for the decider to know in advance of his decision. (One need only include in the set the circumstance that the decider remains ignorant of certain other circumstances in the set at least until the time of the decision....). And a set of circumstances would not be a less plausible candidate for the cause of a decision merely because it had this feature. But neither could a set of circumstances be ruled out as a candidate for the cause merely because it lacked this feature.[6]

But, as I have tried to show, for any set of circumstances to be sufficient for some decision or choice of mine, that I should be ignorant of some aspect of the set (either that they entail my making a particular choice, condition (1), or that they obtain, condition (2), is not just a plausible

---

5 Ibid. pp. 123-4.

6 Ibid. pp. 124.

feature of such a set of circumstances, but a necessary one. What follows from this is that, if it is a fact that my future is such that I shall make at least one choice and if it is the case that determinism is true, I cannot, in principle, know my entire future, or, alternatively, if I do know my entire future, I know as a part of that knowledge or, at least, can infer from it, that I shall make no choices.

Let us suppose for a moment, however, that Ginet's argument, or at least that portion of his argument that I have outlined above, could not be refuted. Need it follow that determinism is false, or is there some other alternative? Ginet, himself, admits that we could deny that men make choices, but that men make choices is, I think, too fundamental a datum to deny in favor of a theory (if, indeed, it would even be possible to do such a thing without making a choice). Is there, then, yet another alternative? If we assume that determinism is true, must we deny that men make choices or could we say, instead, that that conception of what it is to make a choice which is inconsistent with determinism is wrong and that what needs to be denied is not that men make choices but a particular conception of choice? Ginet interprets thinking of a decision or choice as caused as thinking of a decision as a "specific event which, like a flash or a bang, can be identified independently of inquiry

into its causes."[7] Perhaps, this is so, but is this what it is, necessarily, to think of a choice or a decision as determined? It would serve little purpose to repeat, at this point, the first two chapters of this thesis, but if I am correct in the analysis that I have given there of rational choice, then, certainly, a rational choice is not a specific event which can be identified independently of its causes and certainly it is not an event like a flash or a bang. Of the two senses of choice that I distinguished in chapter one, the second is, in the context of a discussion of determinism, the more fundamental. But in the second sense, to make a rational choice is to go through a certain process, a process which we may call the process of choosing. Since it is a process, however, as opposed to a specific, discriminable, and distinct event, it cannot be thought of as being caused in the sense described by Ginet, nor can its occurrence be ascertained without inquiry into its "causes," for each step, stage, or part of the process is, given the total situation, a sufficient condition for the next and is, in that sense, the "cause" of the next, but no one part of the process is identifiable with the whole, and it is the whole process of considering the alternatives and their properties in light of one's criteria that, in this sense at least, constitutes the choice or the choosing. But on Ginet's view, if determinism

---

7 Ibid. p. 124.

were true, the event of my choosing would be something that happens to me and something which I could know to be happening to me without knowing how or why. But if rational choice is understood to be a process, as something which I do, then it is clear that I know the event of choosing is happening to me because I know that I am making a choice. In other words, given the views of rational choice and causation developed earlier in this thesis, the activity of my choosing is integrally a part of the events leading up to and including (is a "cause of") my choice, and my inquiry into the causes of my making a particular choice that I am in process of making is my choosing. Ginet believes that if determinism were true and decisions were caused, it would be possible for a decider to sit back and "watch a series of causally connected events and circumstances produce a decision of his."[8] If my views of rational choice and determinism are accurate, it would appear that the only way to make sense of this statement would be to assume that there are two selves, one which watches and one which chooses. The metaphysical problems raised by such a view are, of course, outrageous, but even if determinism asserted such a view of the self (as, of course, it does not) it would not follow from the fact that the self which watches could know in advance what the choice would be that the self which chooses

---

8 Ibid. p. 123.

could know.

Now, it is only when "choice" is being used in this sense, in a sense that involves choosing, that it is impossible for a person to know in advance what his choice will be, but since choosing is not an event of the kind described by Ginet but, rather, a process, there need be no reason to assume that, if determinism were true, it would be possible for a person to discover what a future choice of his will be. In other words, since a choice is not, in this sense, an event the occurrence of which is ascertainable independently of inquiry into its causes, in order to know that I will make a particular choice at some future date, it is necessary that I know that there will exist conditions sufficient for my making that choice, but since one of the conditions necessary for my making a choice in this sense is that I do not know what I am going to do or what I am going to choose to do, then, either I could not know what choice those conditions are sufficient for or that those conditions exist. Of course, if a choice were an event of the kind described by Ginet, then it would be possible, at least in principle, for me to know that I would make a particular choice (if, for instance, I could see into the future) but this possibility would not follow, I think, from determinism, since, if the basis for the possibility of my knowing what I will choose is the fact that determinism is true, then I would know what I will choose only by knowing its causes and that those causes would exist.

TAYLOR'S ARGUMENT

In his book, Action and Purpose,[9] Richard Taylor offers an argument which is, in many respects, very similar to Ginet's. Taylor, however, takes a somewhat different approach to the problem. Rather than attempting to show that the concept of choice, itself, is inconsistent with determinism, Taylor attempts to show that deliberation is inconsistent with determinism. I do not believe that this is so, and I shall attempt to show that it is not by a step-by-step analysis of Taylor's argument.

Like all sound philosophical arguments, Taylor's begins with a definition, "Deliberation," he says,

> is a process of active, purposeful thought, having as its aim or goal a decision to act, under circumstances in which more than one action is, or at least is believed to be, possible for him who deliberates.[10]

Now, this definition is one with which, with some qualification, I can agree. Deliberation, I take to be a process of rational choice, as I have described that process, where the alternatives from which one (or more) is to be chosen is a

---

9 Englewood Cliffs: Prentice-Hall, Inc., 1966, chapter twelve. Taylor offers essentially the same argument in his earlier book, Metaphysics (Englewood Cliffs: Prentice-Hall, Inc., 1963, chapter four) and in his paper "Deliberation and Foreknowledge" (American Philosophical Quarterly, I, no. 1 (Jan., 1964), 73-80). While the argument is essentially unchanged, it is presented in its fullest, most sophisticated form in Action and Purpose. Accordingly, I shall be concerned with Taylor's argument only as it occurs in that work. All references and quotations are from chapter twelve, unless otherwise noted.

10 Ibid. p. 168.

set of alternatives consisting of different possible actions or courses of action. Thus, deliberation differs from the process of rational choice, if it differs at all, only in that it is directed to a narrower range of objects. There is another difference, not very important in itself. The concept of deliberation differs from the concept of rational choice in that its meanings are not so broad. "Rational choice", as I pointed out in the first chapter, has two senses. In the first sense, the sense of "reasonable" choice, it refers to the object of choice (which, at least in the case of deliberation, is always an action or a course of action); in the second sense, it refers to the process of choice, the "process of active, purposeful thought," whereby the choice is made. Except for the fact that, as I have pointed out above, its range of application is narrower, it would have been possible in my earlier discussion of rational choice to have employed the term "deliberation" in place of the phrase "rational choice". This alternative was foregone, however, for methodological reasons. Since the process of choosing rationally, or of deliberating, may always be interrupted before a choice is actually made, it is understandable that the concepts of deliberating and choosing should be sometimes thought to be more different than they are. I believe, however, that they do not differ in any significant respects (other than those mentioned above) and that they refer to the same thing. When we choose, i.e., when we go through the process

of choosing, we do so for the purpose of selecting one alternative or one set of alternatives from among the set of alternatives which we take to be available; in other words, choosing has "as its aim or goal, a decision to act, under circumstances in which more than one action is, or at least is believed to be, possible." And, of course, the same is true of deliberation. It should be noted, however, that the aim of deliberation is not just a decision to act, if that is understood quite literally; it is, rather, a decision whether to act or not to act or whether to perform this act or that act, that is, a decision to perform one from among a set of apparent alternatives.

It should be noted, too, that Taylor takes the fact that deliberation is a process which is both <u>active</u> and <u>purposeful</u> to be of "primary significance" in his definition. For the moment, it is sufficient to note that I do not attach the same implications to these concepts that Taylor does, since he believes, or at least seems to believe, that these concepts are also inconsistent with determinism. I shall not undertake in this thesis to offer a direct refutation of Taylor's arguments concerning these concepts; however, it will be necessary during the course of the following discussion to make a few remarks, in passing, about them.

From his definition and the fact that deliberation is concerned only with the future, Taylor believes that the four following "restrictions" follow:

(1) that the subject of one's deliberation is something conceived as future, never present or past;

(2) that the subject of deliberation is an action, and not some event unconnected with action;

(3) that this is always one's own action and never that of another agent; and

(4) that it is an action which is conceived as merely possible, and never as one that is either unavoidable or impossible.[11]

The first three of these conditions may stand as they are given here. The fourth, however, will require some clarification and, perhaps, some qualification as well.

Taylor goes on to argue that one cannot deliberate about an action if one believe it to be "already inevitable or unavoidable." This, he says, follows from the fact that one can deliberate only about what is within one's "power both to do and forego" (not to be taken in a sense which is self-contradictory).[12] Now, at this point Taylor runs a very grave risk of presenting an argument which simply begs the question. If "it is in my power both to do and to forego" is taken to mean that I am free with respect to the act in question (in a sense of "free" which is incompatible with determinism) then to say that it is within my power to do and to forego is to say that I am, with respect to the act in question, undetermined and that determinism is, to that

---

11 Ibid., p. 171. In Metaphysics he refers to these conditions as "presuppositions."

12 Ibid., pp. 173-4.

extent, false. From this, of course, it would follow that deliberation is, by definition, inconsistent with determinism, but the real question is "Does deliberation presuppose that whether or not I perform the act is 'up to me' in the sense that it is within my power both to do and to forego and, if so, does it presuppose this in a sense which is inconsistent with determinism's being true?" If deliberation does make this presupposition in this sense, then there is no need for argument: it is inconsistent with determinism, but in the context of philosophical discussion, at least, that deliberation does make such a presupposition and that it makes it in this sense is something to be demonstrated, not just stated. This difficulty, however, may be passed over for the moment.

Taylor continues his argument by stating that to call something inevitable or unavoidable is to deny that it is within my power both to do and to forego. But this, I think, is also a matter for argument, since it depends upon the sense in which "inevitable" is being used or, at least, what it is being used to refer to, as well as upon what is meant by something's being "up to me."

The next point Taylor makes is one which has already been granted in the discussion of Ginet's argument. This is the point that one cannot deliberate about what to do if one already knows what one is going to do. About this point there

need be no discussion.[13] From this Taylor rightly believes it follows, that a man cannot deliberate about whether or not to perform an action if he believes that conditions exist or will exist which are, in themselves, sufficient for his performing that particular action and if he knows what those conditions are. Now, this point is, again, one which may be granted, but it needs some clarification. For example, suppose that during the course of a long automobile trip I must cross a river. I may cross it by a bridge or by a ferry. I deliberate about which to take. Now, suppose that, during the course of my deliberation, I find that the bridge has been washed out. In discovering this, I discover conditions which are sufficient for rendering a certain course of action (taking the bridge) impossible and for rendering another course of action (taking the ferry) inevitable. It is no longer up to me whether or not I shall take the bridge because there is, in effect, no bridge to take, and, of course, if there is no bridge to be taken, then it is not within my power to take it. To say, however, that it is within my power to forego taking the bridge is rather meaningless. It is always within my power to do what I must do and cannot avoid doing. The

---

13 It is not necessary that I argue for this point, since, if it is false, it presents no difficulty for the view that I wish to defend, but, for those interested in counter-examples, see John Canfield's paper "Knowing about Future Decisions" (Analysis, 22.6 (June, 1962), pp. 127-129.) and the reply by Peter Swiggart ("Doing and Deciding to Do" Analysis, 23.1 (Oct., 1962), pp. 17-19.).

point of this example is just that the conditions rendering a particular action of mine inevitable, at least in this case, have nothing to do with my deliberation. In light of these conditions, any further deliberation on my part would be manifestly, to use a phrase of Taylor's, "otiose and pointless." The alternative actions involved here are, in effect, not possible actions. The action of my crossing the bridge is an impossible one, for there is no bridge to cross.

From the preceding argument, Taylor believes it follows that "one's deliberate acts cannot be caused, in the usual sense, or, if they are, then one cannot know that those causes exist at the time he deliberates."[14] Now, there is some question here about what it is for an action to be caused "in the usual sense," but I think, from the context, that Taylor takes it to mean something like "caused in such a way that whether or not I perform the action in question will not be the outcome of my deliberation." To show that this is the correct interpretation, I may point to the example Taylor immediately provides and from which he concludes "that in such cases [where one's action is caused "in the usual sense"] one's deliberation is otiose and pointless, since what one then does is not the result of his deliberation at all."[15]

---

14 Ibid., pp. 176-7

15 Ibid., p. 177.

From this point, Taylor goes on to argue that "if one does not know what he is going to do, but nevertheless knows that conditions already exist which are causally sufficient for his doing whatever it is that he is going to do, then he cannot deliberate about what to do, even though he may not know what those conditions are."[16] That this is so, he again attributes to the fact that one may deliberate only in the belief that what one does is within one's power "to do and forego." From the examples which Taylor gives to illustrate this point, it would appear that he is using "conditions...which are causally sufficient" in the same way that he earlier used "caused in the usual sense." To take one of his examples, Adam has been invited to spend the night at the house of a friend. He deliberates about whether or not to stay until Brown announces that he knows what Adam is going to do and that he knows this on the basis of conditions which are causally sufficient for rendering what Adam will do to be inevitable without telling Adam what those conditions are or what he is going to do. As an example of what those conditions might be, Taylor suggests that Brown might know that the last train had left and that there was no other way for Adam to get home. In this case, it is clear that the conditions are such that Adam's deciding what to do will be pointless. He has no choice. There are no alternatives to

---

16 Ibid.

choose from and what he does will not be the result of his decision. Of course, if Adam believes Brown he can no longer deliberate for the obvious reason that, if he believed Brown, he would believe that he had nothing to deliberate about, and whether or not he wishes to go home (or to stay) he must do what existing conditions allow. Again, there is the problem here of what Taylor means by its being within one's power "both to do and to forego," but this is a phrase which he takes to be more or less synonomous with "it is up to me" and, from his example, it would seem to be clear that what he means here is that what I do is "up to me" if it is the result of my decision, deliberation, or choice.

Taylor goes on to suggest that this point may be "generalized, such that, if a man believes that there are, or ever will be, conditions, not themselves within his control yet sufficient for his doing whatever it is that he is going to do, then he cannot deliberate about what to do."[17] Now, this may be a fair generalization or it may not, and which it is depends upon whether or not Taylor has changed what he means by "sufficient conditions." I confess that I am not quite sure what he means here; however, it would appear from the examples which he provides that he is using this phrase in the same sense as has been given previously. For instance, in one of his examples (that of a man one of

---

17 Ibid., pp. 178-9.

whose future actions will be determined by a roulette wheel), he concludes that "having got this far into the game it is no longer up to him...there is nothing left for him to decide."[18]

At this point Taylor makes rather a giant leap to the proposition that "no one could in principle ever know what another man is going to do as a result of forthcoming deliberation."[19] Taylor's argument here is interesting; however, it is also fallacious. While the faults in his argument are real, they are also subtle and it is not easy to make them clear. To elucidate these faults, then, will require some close attention to detail.

Taylor argues for this point in the following manner: If a person knew what another was going to do as the result of deliberation, he would have to know on "the basis of some kind of evidence; that is, on the basis of...certain conditions that were causally sufficient for the agent's doing the thing in question."[20] This much we may grant; if it is true that A can predict what B will do, A's ability to make such a prediction must (assuming he does not have "second sight") be based on his knowledge of conditions which will produce B's action. Taylor goes on, however, to make a claim which should now be familiar, "if there were such conditions

---

18 Ibid., pp. 179-80

19 Ibid., p. 180

20 Ibid.

then they could also be known by, or made known to, the agent himself, such that he too could infer what he was going to do. This, however, is impossible, so long as the agent has not yet decided what to do."[21] Now, this is obviously just Ginet's argument in a somewhat different form,[22] and precisely the same objections that applied there will apply here. If A knows that B will do X and knows this because he knows B will do X as a result of his deliberation or decision to do X, then A knows that among the conditions sufficient for B's doing X will be the condition that B has decided to do X. Since A's knowledge must be based on knowledge of conditions sufficient to produce B's action and since these conditions have not yet been realized--B has not yet decided--A must be aware of conditions which are or will be sufficient to produce B's decision. Now, among the conditions sufficient for any person's making a decision will always be found the condition that the decider does not know, for some reason, what he will decide. A then knows that B does not know and will not know until he has decided what he will decide. From this A must know or at least be able to infer that even if he attempts to make B aware of those conditions of which he is aware, he will not succeed in making B aware that he will decide to do X before he actually decides. In

---

21 Ibid.

22 As Taylor himself acknowledges in his earlier paper "Deliberation and Foreknowledge," although not in his book.

other words, if A knows that B will decide to do X, then A knows that all of the conditions necessary for B's making such a decision are or will be met: since one of those conditions is that B not know what he is going to decide to do, it follows from the fact that A knows that B does not know. Thus, if a person were to know "what another man is going to do as a result of forthcoming deliberation" it would not follow, as Taylor suggests, that the other man could know also (if the other man did know, he could not decide--which would refute what the person making the prediction was supposed to know) but, rather, that he could not know. Taylor's argument, therefore, fails at this point for precisely the same reasons as does Ginet's.

This is the severest problem in Taylor's argument, but there are others as well. Taylor has attempted to show earlier that, if there are conditions which render what a person does inevitable, the person cannot deliberate about what to do if he knows such conditions exist. The present argument would seem, then, to be just an extension of the earlier one, and it would seem that if we accept the earlier argument, we must also accept this argument. However, this is not the case, for this reason: the sufficient conditions referred to there are not the same as those referred to here. The sufficient conditions which Taylor spoke of earlier were, as I have shown, those which affect the alternatives (e.g., "the last train has left," "the bridge is washed out," etc.),

and they determine what the person will do by limiting the alternatives open to him, by eliminating the choice situation. What these conditions render inevitable is what the person does and they do this by removing the alternatives open to him. In the argument at issue, however, what is rendered inevitable is not directly what the person does, but what the person decides to do, and the sufficient conditions referred to in this argument must be those that affect the person's decision. That there is this difference between the two arguments Taylor does not make clear, but it is not at all obvious from the fact that conditons irrelevant to the person's decision which render a certain action inevitable make his deciding "otiose and pointless" (and impossible if he is aware that such conditions exist) that decisions themselves cannot be caused. The conditions referred to earlier do not cause the decision, they remove the necessity for a decision, and, if the agent is aware of them, the possibility of his making a decision. Certainly it does not follow from this that decisions cannot be caused. There is a neat trick which may be employed here, however, to make it follow that decisions cannot be caused: we may change the definition of "it is up to me." Where conditions sufficient for a certain action render a decision impossible, what I do is not then "up to me." This follows from the fact that if what I do is "up to me", what I do must be the result of my decision to do it. Where I cannot decide, therefore, what I

do cannot be up to me. Now, if we understand "it is up to me" to mean that my decision is not caused, then, if there are sufficient conditions for my making a decision, it follows that what I do is not up to me, and if its being up to me is taken as a necessary condition of its being the result of my decision, it follows that my decisions cannot be caused. It also begs the question, and it is essentially the argument Taylor offers here. He says, "Indeed, the agent cannot even believe that any such conditions...exist, and at the same time believe that it is within his power both to do, and to forego doing the thing in question."[23] This is true, however, only if "it is within my power both to do, and to forego" or "it is up to me" is taken to mean that my decision is not caused, and this begs the question. In his previous arguments, Taylor has rightly argued that "it is up to me" is incompatible with my action's being caused in such a way that which action I perform is not the result of my decision. He has not argued, however, at any point, that "it is up to me" is incompatible with my decision's being caused, and that it is incompatible with my decision's being caused obviously does not follow from his earlier definition of "it is up to me." This being so, it is clear that, since the prediction taken by Taylor to be impossible in this argument is what another man is going to do as a result of his deliberation,

---

23 Ibid.

there would seem to be nothing about the act or the prediction of it which is inconsistent with its being up to him.

In order to show that such a prediction could, in principle at least, be made and that it would not be inconsistent with the agent's deliberating or with the act's being "up to him," I shall give the following example: Dick is a man who owes a very large debt which he cannot pay and which will, in all probability, cause him to declare bankruptcy, with subsequent loss of nearly all that he owns and, quite possibly, the loss of his wife and family as well as the respect of his colleagues. Dick has been entrusted, however, with a very large sum of money by a friend and business associate, Tom. After having entrusted this sum of money to Dick, Tom becomes aware, for the first time, of Dick's financial circumstances; he also learns that, since receiving the money, Dick has been making discrete inquiries about travel arrangements to Buenos Aires and extradition agreements between his own and the Argentine government. Tom, needless to say, becomes worried about the safety of his money; however, he doesn't wish to make a direct accusation on the basis of such evidence. He consults Harry for advice. Harry is a bosom buddy of Dick's: they grew up together, went to school together, had a double wedding (marrying twin sisters), are business partners, and get together several times a week on a friendly basis. In short, Dick and Harry know each other "like a book." Harry advises Tom in this manner:

"Don't worry. I know dick. He is probably considering absconding with the money, but he won't do it. He is just too moral. Pretty soon he will be calling me to ask for advice." Next day, Dick, who has been all the while deliberating about whether or not to abscond with the money, calls his friend Harry for advice. Harry advises him in this way: "Don't worry, Dick; you will do the right thing. I know you will." Dick believes him, but, perhaps, does not know what he is going to do, what is the right thing. In this case, it would perhaps be true that Dick could no longer deliberate. But surpose the conversation had gone in this way: "Dick," Harry says, "I know you like a book." "I know you do, Harry," Dick replies. "And I could tell you what you are going to do," Harry continues. "I believe you could, Harry," says Dick, a bit despondently (Why are his conversations with Harry always so uninformative, he wonders). "But," says Harry, "I won't." "I knew you would say that," Dick replies, a bit preoccupied. "Because," Harry goes on doggedly, "I believe you should make up your own mind." "Well, you know, Harry, I knew you would feel this way. Still, it would save me a lot of personal torment if you would just tell me what I would otherwise have to choose so I can simply do it and have done with it," but Dick knows his friend well enough to know that this entreaty will prove futile. "Well, Dick," says Harry, "you know I always have your best interests at heart, and I think the struggle will be good for your soul. You really ought to

understand yourself a little better, you know." "Perhaps," says Dick, "but if you will excuse me I have a lot of thinking to do." Next day, Dick calls Harry and tells him that, after much thought, he has decided, finally, to return the money. "I knew you would do that," says Harry.

Now, this example may be a bit far-fetched, but there is nothing about it that makes it, in principle, impossible. Certainly, more familiar, everyday examples of the same sort could be found. And, obviously, there is no good reason to say that what Dick does, in this case, is not up to him. Harry knows what Dick will do because he knows what Dick will decide to do, and he knows this, not because he is aware of some conditions which will make what Dick does inevitable (such as Tom's having him arrested), but because he knows Dick.

Taylor offers an example of his own to show that this sort of thing cannot happen. Suppose a man B when he is confronted with a certain choice always makes the same choice after deliberating about the matter. In this case, it would be possible to predict what he would choose. But, says Taylor, in such a case, our knowledge of what he will do is not knowledge of what he will do as a result of his deliberation but of his habits. Now, this is a very strange habit, but, if the habit is just that of always deliberating in such a situation, then, certainly, this is not inconsitent with his always choosing the same thing as the result of his

habitual deliberations. Assuming that the situation and he were the same, why should his decision differ? If his choice is the result of his deliberation, whether or not he has the habit of deliberating, what he does is, at least as Taylor has been using the phrase, "up to him." Admittedly, it might be strange for a man to be a habitual deliberator; nevertheless, there would seem to be nothing in the idea such that it would be impossible. On the other hand, if the habit is making that particular choice, then Taylor's analysis of the example would be correct to the extent that the act was not the result of his deliberation, but it need not follow that it is not up to him or that it is not his choice, assuming, of course, that a choice can become a habit. At any rate, the mere fact that a man usually or always does the same thing in the same situation does not entail that what he does is habitual or that it is done because of habit and not from choice, unless, of course, "habit" is defined simply as "regularity."

Taylor goes on to give another example which is, I think, even more odd. Suppose I know a man is deliberating about whether or not to leave a room and suppose I know, too, that the house is on fire and that the man will soon become aware of this fact. I know, then, according to Taylor, that "what he is going to do will <u>not</u> be the result of his deliberation, but of his knowledge of the circumstances."[24] Perhaps,

---

24 <u>Ibid</u>., p. 181.

given the circumstances, this is so, but certainly it does seem odd to oppose one's deliberating about what to do in a given set of circumstances so emphatically to one's knowledge of those circumstances.

The conclusion which Taylor draws from his argument is, of course, that "deliberation is inconsistent with determinism."[25] He has not, I think, established his case for this view. But he goes on to say that it will not do to say that deliberation may be one of the causes of human actions. This is so, he says, because "It is part of the very concept of deliberation that it applies to situations in which there are, or are at least believed to be, alternative possible courses of action, and that as a result of one's deliberation, he might do either one."[26] Now, this is very intriguing and, if I may be permitted the liberty of saying so, the oddest statement Taylor has, to my knowledge, yet made. For, literally interpreted, it means that even after having deliberated and decided, I may or I may not do what I have decided to do. Understood in this way, it is clear that deliberation would, if determinism were not true, be even more "otiose and pointless" than if determinism were true. Of course, it is doubtful that Taylor means it quite literally. What he no doubt means is the quite trivial fact that one deliberates only when there are alternative possible courses of action,

---

25 Ibid. 26 Ibid., p. 183.

and, depending on the course taken by one's deliberation, one might take one alternative or the other. In the same sense, we may say that, depending on how the eight-ball is struck, it may roll into the corner pocket or it may not. So long as we do not know how the ball will be struck, we do not know where it will roll and we can only say that as a result of its being struck, it might do either one. Where there are alternatives open to the ball, it might do either one; where there are alternatives open to me, I might do either one. Taylor makes the remark quoted above to explain the sense in which saying "that some agent does something as a result of deliberation is very far from saying that he is caused to do it by his deliberation."[27] Indeed, this is so, but from the fact that saying the one is the result of the other is different from saying it is the (causal) effect of the other, it most emphatically does not follow that it is to deny that it is the effect of the other. There would be nothing wrong with saying "The eight-ball rolled into the corner pocket as a result of its being struck," and, while this is very far from saying that it was caused to do what it did by its being struck, it is obviously not to deny it. There is no reason to assume, then, that deliberation may not be one of the causes of human actions.

---

27 Ibid., pp. 182-3.

THE FUNDAMENTAL NATURE OF "CHOICE"

While both Ginet and Taylor have attempted to show that the concepts of determinism and choice or deliberation are inconsistent, neither has attempted a very careful analysis of both these concepts. Taylor has attempted an analysis of "cause", but his analysis of deliberation as an activity is far from complete. Ginet, while he offers an interpretation of what it would mean to think of a decision as caused, has not attempted to defend this interpretation as the only possible one in view of determinism and does not even pretend to have offered an analysis of the concept.

On the other hand, I believe that the concept of rational choice is so fundamental that we cannot even understand what it means to say that human activity is determined without an understanding of what it is to make a rational choice. I have defined a rational choice as one made on the basis of or in accordance with reasons (alternatively, one could say "as a result of a consideration of those reasons") where a reason is some fact about the object seen in light of a standard or criterion of choice and where standards and criteria are reducible to our basic likes and dislikes. Given this conception of the process of rational choice (or of deliberation) and my explication of the concept of determinism in light of the principle of causal entailment, it is

fairly easy to see how these two concepts could be fit together. The real question, I think, whether or not determinism is true, is "Why do I choose what I do when I choose rationally?" The answer, I believe, is consistent both with the view that men do make rational choices and that determinism is true; namely, I so choose because I am a being such as I am and the situation is such as it is. If my choice were not the result of my being the person that I am, having the likes and dislikes, purposes, and beliefs that I have, it would be difficult to see how it could be described as my choice at all, and were my choice not at all dependent upon the situation and the alternatives, or, at the very least, upon my knowledge and beliefs about the situation and the alternatives, then, indeed, deliberation would be pointless. This, of course, is not to say that deliberation is not pointless.

REASONS AS CAUSES

There is another topic which should be taken up briefly before this thesis is brought to a close. This is the distinction made so much of by Richard Taylor and other contemporary philosophers between reasons and causes.[28] It is not a distinction, however sound the arguments for it, which concerns me or which should, for that matter, concern anyone

---

28 Ibid., chapter 10.

else. "Causes" as they are ordinarily understood do not exist. To speak of the "cause" of an event is always to describe, in an abbreviated and always artificial fashion, the actual course of events or, to use a much more accurate phrase of John Dewey's, the "state of affairs." I need not specify the "cause" of any event or of any action, nor need any determinist. What needs to be described, in as accurate a manner as possible, admitting that complete accuracy in this case is always impossible, is the series of events or conditions leading up to and including the event to be accounted for (although to think of the state of affairs as a "series" of events is already to have falsified it). I do not claim that my reasons for making a particular choice, however weighty, "cause" that choice. I do not need to. What I do claim is that, if determinism is true, my having those reasons is an integral part of the event leading up to and including my making that choice and that any account of my making that particular choice as opposed to some other, if my choice is rationally made, must, if it is to be at all complete, include an account of those reasons as determining factors.

RATIONAL EXPLICABILITY

I have argued, too, that if determinism were false, human behavior and events in general would be rationally inexplicable. Taylor takes issue with this point, arguing

that such things as reasons and purposes, while they do not provide causal explanations, do explain human actions and that the claim that human actions would be rationally inexplicable if determinsim were false is, therefore, groundless.[29] This argument is, in a sense, valid, but it misses the point. On Taylor's view, to explain an act in terms of reasons, purposes, goals, etc., is not to deny it is caused. It is to give a different kind of explanation and says nothing about the existence or non-existence of conditions sufficient for just that action. It is, rather, to describe some, but not all of the conditions surrounding the act, that it was done with reason or with purpose, and, thus, to make the act "intelligible." Such an "explanation," of course, does not explain the act, at least not all by itself and in any usual sense of "explain." At best, it describes the act in terms of other things; it puts it in a certain kind of framework. Sometimes such an "explanation" of how a given act "fits in" with surrounding circumstances is all I want, but more often I wish to know why it should fit in. I want, in such cases, to know not "What did you do that for?" nor "What did you hope to accomplish?" (both senses in which "Why did you do that?" might be taken), but "Why did you do that?" in the sense of "What led you to do that?" To take an example, the psychiatrist asks his patient why he did something and the

---

29 Ibid., pp. 140-147.

patient gives him a reply of the first type. But what the psychiatrist wants to get at is an explanation of the second type. It is the fact that explanations of the first type can be given that makes actions rationalizable. It is the fact that explanations of the second type can be given (assuming they can) that makes them rational, and it is in this sense that actions become rationally inexplicable if determinism is false. To take an example used by Taylor, if I ask a guest at a party to leave, I could, as an explanation for my act, say that he was becoming rude and I did not want him to spoil the whole party. As Taylor correctly points out, this would be an explanation in terms of my purpose which would, quite possibly, make my action intelligible to others. But such an "explanation," if it were not the case that I had asked the man to leave because, under those circumstances, I wanted him to leave or because I wanted him not to spoil the party, would clearly, at least as the words are ordinarily used, be a case of rationalization, and my action, if no explantion of the other kind could be given for it, if it could not be shown or, at least, assumed to be, in some manner, connected either with the situation or with my alleged purpose, would be, at best, non-rational and, in this sense, rationally inexplicable, and that I should have performed this action rather than some one of the other, innumerable possible actions and that I should have had this purpose rather than some one of the other innumerable possible

purposes would be totally unintelligible. It is in this sense that I believe human actions would be rationally inexplicable if determinism were completely false, but I do not claim, nor have I claimed that human actions ever are, in fact, rationally explicable.

BIBLIOGRAPHY

Canfield, John. "Knowing about Future Decisions," Analysis, 22.6 (June, 1962), 127-129.

Chomsky, Noam. "Verbal Behavior. By B.F. Skinner," Language, 35 (1959), 26-58.

Davidson, Donald. "Actions, Reasons, and Causes," Journal of Philosophy, LX (Nov. 7, 1963), 685-700.

Dewey, John. Experience and Nature. New York: Dover, 1958.

Epicurus. "Letter to Menocceus," Geer, Russell M., trans., Epicurus: Letters, Principle Doctrines, and Vatican Sayings. Indianapolis: Bobbs-Merrill, 1965.

Gauthier, David. Practical Reasoning. London: Oxford, 1963.

Ginet, Carl. "Can the Will Be Caused?" Philosophical Review, LXXI (1962), 49-55. Reprinted in Dworkin, Gerald, ed., Determinism, Free Will, and Moral Responsibility. Englewood Cliffs: Prentice-Hall, 1970.

Hampshire, Stuart. Thought and Action. London: Chatto and Windus, 1959.

Hartnack, Justus. "Some Remarks on Causality," Journal of Philosophy, L (1953), 466-471.

Heisenberg, Werner. Physics and Philosophy: The Revolution in Modern Science. New York: Harper and Row, 1958.

Hobbes, Thomas. Leviathan, Parts I and II, ed., Schneider, Herbert W. Indianapolis: Bobbs-Merrill, 1958.

Hume, David. Hume's Treatise of Human Nature. ed., L.A. Selby-Bigge. London: Oxford, 1888.

--------. An Inquiry Concerning Human Understanding, ed., Hendel, Charles W. Indianapolis: Bobbs-Merrill, 1955.

Lorenz, Konrad. Evolution and Modification of Behavior. Chicago: University of Chicago, 1965.

MacCorquodale, Kenneth. "On Chomsky's Review of Skinner's Verbal Behavior," Journal of the Experimental Analysis of Behavior, 13 (Jan., 1970), 83-99.

Nowell-Smith, P.H. Ethics. Middlesex: Penguin Books, 1954.

Oppenheim, Felix. "Rational Choice," Journal of Philosophy, L (June 4, 1953), 341-350.

Rawls, John. A Theory of Justice. Cambridge: Harvard University, 1971.

Reid, Thomas. "Some Arguments for Free Will," in Dworkin, Gerald, ed., Determinism, Free Will, and Moral Responsibility. Englewood Cliffs: Prentice-Hall, 1970.

Russell, Bertrand. An Inquiry into Meaning and Truth. Middlesex: Penguin Books, 1962.

Skinner, B.F. Science and Human Behavior. New York: MacMillan, 1953.

--------. Verbal Behavior. New York: Appleton-Century-Crofts, 1957.

--------. Contingencies of Reinforcement. New York: Appleton-Century-Crofts, 1969.

Swiggart, Peter. "Doing and Deciding to Do," Analysis, 23.1 (Oct., 1962), 17-19.

Taylor, Richard. Metaphysics. Englewood Cliffs: Prentice-Hall, 1963.

--------. "Deliberation and Foreknowledge," American Philosophical Quarterly, I (Jan., 1964), 73-80. Reprinted in Berofsky, Bernard. Free Will and Determinism. New York:

--------. Action and Purpose. Englewood Cliffs: Prentice-Hall, 1966.

Warnock, G.J. "Hume on Causation" in Pears, D.F., ed., David Hume: A Symposium. London: McMillan, 1963.

Winch, Peter. The Idea of a Social Science and Its Relation to Philosophy. London: Routledge & Kegan Paul, 1958.