DETERMINISM AND MORAL RESPONSIBILITY

H. B. DALRYMPLE

The purpose of this paper is to examine the adequacy of some efforts to reconcile determinism and moral responsibility. But perhaps I should begin by stating as clearly as I can what might have led some unsophisticated people to believe there is an incompatibility.

Although determinism has been said to stand for "nothing clear," a very simple definition will be good enough for my purposes. The doctrine of determinism is the doctrine that for every event A there is a set of preceding events B such that B having occurred A is inevitable. The doctrine of determinism as applied to Jones' actions is the doctrine that for every action of Jones A there is a set of preceding events B such that B having occurred A is inevitable. Some unsophisticated people have argued that if Jones' actions are determined in the sense defined that Jones is not to be held responsible for anything that he does. Other, more sophisticated, people have argued that these simple-minded folks are suffering from an illusion.

This definition of determinism, like all definitions of determinism, is subject to the complaint that it cannot be disconfirmed; that is, even if science fails to find such a set of events B, it can always be claimed that there is such a set of events that has not been found. It was some such consideration as this that probably led Kant to regard the sentence, "Every event has a cause," as being synthetic *a priori*. And it is the same sort of consideration that seems to be responsible for some analysts saying that sentences like the above are meaningless; that is, since the sentences cannot be falsified, they are without any empirical meaning according to the verifiability criterion of meaning. But perhaps at least it can be shown that there are situations in which people would employ explanations that would approach one of the determinist paradigms.

Let us put the matter in this way. Suppose Smith believes that at least some of Jones' actions are indeterminate, what might induce him to change his mind and concede that all of Jones' actions are determined? Just the demonstrated ability of some person (say Frazier) to predict successfully how Jones behaves. Suppose, furthermore, that Frazier gradually acquires his ability to predict Jones' behavior and that Smith witnesses the gradual improvement in Frazier's ability. At what point will Smith change his mind and concede that what Jones will do is strictly determined? Well, probably long before Frazier's predictions reach 100 per cent accuracy. Which is to say that Smith will at a certain point be willing to admit that Frazier's failures in prediction are the result not of nature's ways but of Frazier's mistakes or ignorance.

The problem of whether moral responsibility can be reconciled with determinism can now be put in a more precise way. If Smith as a result of Frazier's impressive performance in predicting Jones' behavior comes to genuinely believe that every one of Jones' actions is fully determined in the sense defined, will he then hold Jones morally responsible for any of his actions whatsoever? Incompatibilists whether of the libertarian or determinist schools argue that Jones would not be morally responsible in these circumstances while compatibilists argue that these circumstances are consistent with Jones being morally responsible.

I shall now consider a major effort to show that moral responsibility is compatible with determinism--the attempt to show compatibility by means of the analysis of the expression "could have done otherwise." The other major effort is an attempt to show that freedom should be contrasted with compulsion, not with determinism. The argument from "could have done otherwise," is a rebuttal of an argument in favor of incompatibilism. The argument in favor of incompatibilism runs as follows: People generally (philosophers and nonphilosophers) are agreed that moral praise and blame are deserved only when a person could have done otherwise. If this is so, it would seem to follow that if determinism is correct that nobody is ever deserving of praise or blame. For if determinism is correct there are no occasions in which a person could have done otherwise than he did in fact do. Compatibilists deny the implication, however, holding that determinism is compatible with "could have done otherwise" and hence with moral responsibility. "X could have done otherwise," is held to mean something like "X would have done otherwise had X chosen to do otherwise, " and the latter expression is compatible with determinism. I hold that compatibilists have confused the meaning of expressions containing "could have done other than x" with expressions containing "could have done x and meaning of expressions containing "could not have done other than x" with expressions containing "could not have done x." For the sake of simplicity, I concentrate upon cases where the question at issue is whether blame should be attributed to an agent. The two kinds of expressions appear in the following forms:

1.1. X is a wrong action and Jones couldn't have done other than x. (Jones is not blamed for doing x.)

1.2. X is a wrong action and Jones could have done other than x. (Jones is blamed for doing x.)

2.1. X is a right action. Jones didn't do x, but he could have done x. (Jones is blamed for not doing x.)

2.2. X is a right action. Jones didn't do x, but he couldn't have done x. (Jones is not blamed for not doing x.)

The compatibilist's analysis of 2.1 and 2.2 is approximately correct. Consider the following examples of 2.1. Jones is blamed for not

saving Smith because we say that Jones could have saved Smith; i.e., if Jones had tried, he would have saved Smith. Or, alternatively, Jones is blamed for not saving Smith, because Jones knows how to swim, he was at the scene of the drowning and he was in good physical condition. The youth is blamed for doing his sum incorrectly, because we say that he could have done his sum correctly; i.e., he would have done his sum correctly if he had tried (or tried hard enough). In these two instances, and in many others, in which a person is thought to be deserving of blame, it makes sense to elucidate expressions of the form "A could have done x," in some such fashion as the above.

Consider the following examples of 2.2. Jones is not blamed for not saving Smith because we say that Jones couldn't have saved Smith; i.e., Jones wouldn't have saved Smith if Jones had tried. Or, alternatively, Jones is not blamed for not saving Smith, because Jones doesn't know how to swim and there was no other means available to Jones by which he could have saved Smith. The youth is not blamed for doing his sum incorrectly, because we say that he wouldn't have done his sum correctly if he had tried (or tried as hard as we had any right to expect). In both of these instances, and in many others in which a person is not blamed (or excused from responsibility) it makes sense to elucidate expressions of the form, "A could not have done x," in some such fashion as the above. The point that the compatibilist is making is that moral responsibility implies ability and opportunity.

What I want now to show is that the behavior of "A could have done other than x," and "A could not have done other than x" is different from the above. A proper analysis will lead straight into the Hume-Mill-Schlick theory. I begin with "A could not have done other than x," since its meaning seems to be relatively clear and unambiguous. "A could not have done other than x" seems to have about the same meaning as a number of other expressions: "A had to x," "A had no choice," "If A had had a choice, he wouldn't have x-ed," "A had no choice but to x," and "A had no other choice than to x." The following paradigms show that these expressions are at least roughly equivalent.

Jimmy, bitterly protesting, is enrolled in the first grade. Jimmy's situation is described as "Jimmy had to go to school," "Jimmy had no choice," "If Jimmy had had a choice, he wouldn't have gone to school," "Jimmy had no choice but to go to school," and "Jimmy had no other choice than to go to school.

At gunpoint, a man jumps out of a window, landing in a flower garden. To the enraged gardener the man explains "I'm sorry, I had to do it; they made mo do it." He might have said "I could not have done other than jump out the window," "I had no choice about jumping out the

window; I had to jump," "If I'd had a choice, I wouldn't have jumped in your flower garden," or "I had no other choice than to jump in your flower garden."

At this point, a confusion about the proper use of expressions of this sort needs to be cleared up. In our second paradigm, a man is forced at gunpoint to jump out the window. The man excuses himself by saying that he had to jump. If the gardener does not accept this as a valid use of "I had to jump," the question of what he would accept as a valid use becomes a problem. The man being pushed out the window, perhaps? But then the man would not say that he had to jump, but rather that he hadn't jumped, or perhaps he would say that his body landing in the flower garden was none of his doing. In short, the expressions in this group apply only to human actions.

The expression that needs to be scrutinized carefully is the expression "A had no choice." By itself, the statement seems to indicate in the clearest possible terms that there was no choice. The trouble is that the sentence makes no sense to people who are unaware of the relevant situation; that is, the sentence will need to be completed in some way. Here are two ways in which the sentence can be completed with sacrificing the flavor of the original: "A had no choice about x-ing; he had to x," and "If A had had a choice, he wouldn't have x-ed." Also, notice that "A had no choice," seems to be the denial of "A had a choice," (another incomplete expression), and "A had a choice," is best completed by "A had a choice of x-ing or not x-ing," so that "A had no choice," would become "A had no choice of x-ing or not x-ing." "A had no choice of x-ing or not x-ing," was not mentioned in my list of equivalent expressions, but it can be made without much strain to fit contexts in which the other expressions are also appropriate. For example, consider the case of a "shotgun wedding." Jones, who doesn't know that the wedding is a "shotgun wedding," says "Jack wasn't such a bad guy after all. He had a choice of marrying her or leaving her, but he did the right thing." Smith who knows the kind of wedding the wedding was, replies "You're wrong, Jones. Jack didn't have a choice of marrying her or not marrying her. Old MacDonald made him do it."

Of course, "A had no choice," may be completed by expressions like "A had no other choice than to x," and these expressions seem to imply that A chose to x and that choosing to x was the only choice that A had. But I believe instead that the expressions are just misleading ways of designating the situations in which choosing does not occur. Notice that it makes perfectly good sense to "A had no other choices than to x, to y or to z," or to say "A had no other choices than to x or to y." But now, by an illicit transfer, this becomes "A had no other choice than to x," along with suggestions that nobody ever intended.

The analysis of "A could have done other than x" follows a similar line. What needs to be shown is that expressions of the form "A could have done other than x" are equivalent to expressions with forms like these: "A didn't have to x" and "A had a choice of x-ing or not x-ing." For example, Jack, though sorely tempted to leave the girl, finally decides to do the right thing. Jones says, "Jack isn't as bad a guy as we had thought, because after all he didn't have to marry the girl." Instead he might have said "Jack could have left the girl," or "Jack had a choice of marrying her or not marrying her, but he married her." The point that is made is this: A is not excused for doing x or A can claim credit for doing x if it can be shown that A could have done other than x, but to show that A could have done other than x is simply to show that A had a choice of x-ing or not x-ing. Or, what amounts to the same thing, to blame A for doing x or to praise A for doing x because A could have done other than x is simply to notice not only that A did x but that x was something that A chose to do.

What has been argued above can be restated in terms of the Hume-Mill-Schlick theory. Let A's chosen actions be defined as the actions that he does of his own free will, and let A's actions that are not chosen by him be defined as the actions that he is compelled to do. Then he is excused, or not given credit for, the actions that he is compelled to do, and he is praised or blamed, rewarded or punished, for the actions that he does of his own free will.

A recapitulation is in order. A being morally responsible for x implies that (1) A can do x (A has the ability and the opportunity to do x) and (2) A could have done other than x (x was something that A chose to do). (1) is clearly compatible with determinism, since abilities and opportunities are determined in large part at least by education, inheritance of wealth and position, genes, etc. What about (2)? Chosen, no less than compelled, actions are after all bodily movements, and bodily movements are natural events. So I can see no good reason why chosen actions should not be subject to causal explanations, and, if they are subject to causal explanations, then the thesis that chosen actions are completely determined could be true. But the possibility remains that moral responsibility implies something more than (1) and (2). Perhaps in addition to (1) and (2) it also implies (3) that A's action is not an inevitable action.

An inductive consideration in favor of (3) comes from looking at (1) and (2) in a certain way. Let x be some objectively right action. Then A can be excused for not doing x if (1) A does not have the ability or the opportunity to do x or (2) A is compelled not to do x. But if A does not have the ability or the opportunity to do x, then (3) it is inevitable that he will not do x. And if A is compelled not to do x, it is inevitable that he will not do x.

will not do x. (These may be analytic truths, but it does not matter.) Thus, (3) being a common feature of (1) and (2), the possibility exists that it is (3)—the inevitability of not doing x—that is the feature in (1) and (2) that constitutes the univocal meaning of "excusing circumstances."

My principal argument in opposition to the compatibilist position takes the form of an analysis of how the word "responsible" behaves. "Responsible" is defined in a variety of ways, but I believe the following is the core meaning of "responsible": "A is responsible for x if A is the author or cause of x." Here are two examples. The car doesn't start, and I explain that the batteries are responsible. The garbage pail is turned over, and I explain that the neighbor's dog is responsible. If asked the further question of how I know that the dog is responsible, I explain that I saw him do it. These examples are remarkable for their simplicity; more complicated examples result when the immediate cause is seen as the inevitable result of something else. The example of the dead battery can be used to show the complication. I demand a new battery or my money back, because I brought the battery just a month ago and it has already gone dead twice. After looking inside the engine, the man at the garage refuses to reimburse on the grounds that the battery was not responsible for my difficulties. Instead, the motor was responsible; something in the motor was running the battery down and would run any battery down, regardless of its quality. My final example is an imaginary one: a machine has more parts than any person or group of persons could examine in a lifetime or in several lifetimes. At first, A seems to be responsible for the untoward occurrence x, but then it is found out that the defect in A is the inevitable consequence of a defeat in B, and the defect in B is an inevitable result of a defect in C, and so on. In such a situation, none of the machine's parts would be regarded as being responsible.

There is no reason to believe that the behavior of the word, "responsible," is any different in moral than in nonmoral context. Suppose, in Frazier's utopian experiment in human engineering, that Jones has the job of deciding which infants will be the victims of infanticides. He conditions Smith to carry out the infanticides. Smith is not compelled because he enjoys his work. Who would somebody who was not a member of the community blame? Jones, perhaps? But it turns out that Jones (who also enjoys his work) is a product of Frazier's foolproof methods of human engineering. So who is to blame. Frazier?

University of Texas at Arlington