

Journal of Philosophy, Inc.

Responsibility and Control

Author(s): John Martin Fischer

Source: *The Journal of Philosophy*, Vol. 79, No. 1 (Jan., 1982), pp. 24-40

Published by: [Journal of Philosophy, Inc.](#)

Stable URL: <http://www.jstor.org/stable/2026344>

Accessed: 24/02/2011 13:02

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=jphil>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Journal of Philosophy, Inc. is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Philosophy*.

<http://www.jstor.org>

modal logics lacking agglomerativity and our main principle,²⁵ but there is an obvious need for technical investigations of the logic and semantics of the selective necessities we have mentioned and various others that are undoubtedly waiting to be discovered. I believe such projects may be worth pursuing, quite independently of any relation that exists between selectivity and the specific problem of free will, for the insights they are likely to give us into modal thinking in general.

MICHAEL SLOTE

Trinity College, Dublin

RESPONSIBILITY AND CONTROL*

MOST philosophers have held that a person is morally responsible for what he has done only if he could have done otherwise. If responsibility requires freedom to do otherwise and this freedom is incompatible with causal determinism, then responsibility is incompatible with determinism.

Harry Frankfurt has contended recently that freedom to do otherwise is *not* necessary for responsibility¹; Frankfurt thus attempts to show that the question of the consistency of determinism with responsibility can be separated from that of the consistency of determinism with freedom to do otherwise. I shall discuss two ways of responding to Frankfurt's examples. The first response claims that Frankfurt has *not* successfully dissociated responsibility from "control" (that is, freedom to do otherwise), in which case the compatibility of determinism with responsibility continues to depend on the compatibility of determinism with freedom to do otherwise. The second response concedes that Frankfurt *has* successfully pried

²⁵ See, e.g., Krister Segerberg, *An Essay in Classical Modal Logic* (Uppsala: University Press, 1971), and Brian Chellas, *Modal Logic: An Introduction* (New York: Cambridge, 1980). However, what also needs to be explored (though there is no space to explore it here) is the possibility that there are indexical aspects to (some of) the necessity operators of ordinary language, that (various) apparent failures of agglomerativity and the like involve subtle shiftings of context rather than any flouting of "normal" modal principles, and that recent arguments for incompatibilism commit some sort of fallacy of equivocation.

* I am indebted to Carl Ginet, T. H. Irwin, and especially John G. Bennett for useful comments. I benefited from reading this paper at colloquia at Cornell, Stanford, and Dartmouth.

¹ "Alternate Possibilities and Moral Responsibility," this JOURNAL, XLVI, 23 (Dec. 4, 1969): 829-839.

apart responsibility and control, but it claims that he has not thereby established the consistency of responsibility with determinism. Though it is usually thought that determinism threatens responsibility because it erodes control, I show how an incompatibilist about determinism and responsibility can agree with Frankfurt that responsibility need not require control.

I. FRANKFURT'S EXAMPLES

Let me start by reviewing some of the larger issues into which the topic of this paper fits. An incompatibilist about responsibility and determinism might make an argument against compatibilism by appealing to what Frankfurt calls the "principle of alternate possibilities." This principle says that

A person is morally responsible for what he has done only if he could have done otherwise.

The incompatibilist argues that if universal causal determinism is true, then no one could ever have done other than he actually did. Consequently, by the principle of alternate possibilities, no one is ever morally responsible for what he has done, if causal determinism is true.

In understanding this argument it is important to see that something like the principle of alternate possibilities is usually accepted by *both* compatibilists and incompatibilists. This is why the incompatibilist's argument is powerful. Usually both camps associate responsibility with control, but, whereas the incompatibilist holds that determinism rules out control, the compatibilist argues that one might have been able to do otherwise, in the relevant sense, even though one's action was in fact causally determined by antecedent events over which one had no control.

Harry Frankfurt has suggested a new attack on the incompatibilist's argument. Instead of disputing the incompatibility of determinism with control, he attacks the principle of alternate possibilities itself—that is, the claim that responsibility requires control.

Frankfurt's examples have the following general form²: A person is in the process of deciding which of n alternative acts $A_1, \dots, A_k, \dots, A_n$ to perform. He believes (correctly) that he cannot avoid performing some one of the acts, although there is no one act of which he believes that he can't avoid performing it. He decides to perform and, acting on this decision, does perform A_k . But, unknown to him, there were various factors that *would have* prevented

²In summarizing the form of Frankfurt's examples, I closely follow Peter van Inwagen, "Ability and Responsibility," *Philosophical Review*, LXXXVII, 2 (April 1978): 201-224, esp. p. 202.

him from performing (and also from deciding to perform) any of A_1, \dots, A_n except A_k . These factors would have come into play if he had shown any tendency toward performing (or deciding to perform) any of the alternatives except A_k . But, since he in fact showed no such tendency, these factors played no role whatever in his actually deciding to perform and in his performing A_k .

I want to describe a case of the Frankfurt sort in some detail; this is a somewhat extreme case, but it poses all the problems in an especially pressing way. Black is a nefarious neurosurgeon. In performing an operation on Jones to remove a brain tumor, Black inserts a mechanism into Jones's brain which enables Black to monitor and control Jones's activities. Jones, meanwhile, knows nothing of this. Black exercises this control through a computer which he has programmed so that, among other things, it monitors Jones's voting behavior. If Jones shows an inclination to decide to vote for Carter, then the computer, through the mechanism in Jones's brain, intervenes to assure that he actually decides to vote for Reagan and does so vote. But if Jones decides on his own to vote for Reagan, the computer does nothing but continue to monitor—without affecting—the goings-on in Jones's head.

Suppose Jones decides to vote for Reagan on his own, just as he would have if Black had *not* inserted the mechanism into his head. Then Frankfurt claims that Jones is responsible for voting for Reagan, regardless of the fact that he could not have done otherwise. His voting for Reagan is something we can charge to his credit or discredit; it expresses something about Jones's character. But the principle of alternate possibilities does not seem to be satisfied; there is no way that Jones could have avoided voting for Reagan—if he had been on the verge of doing so, the computer would have intervened and ensured that he vote for Reagan. If the computer *had* intervened, then I think it is obvious that Jones would not have been responsible. But since the computer did *not* intervene, it is plausible to think, along with Frankfurt, that Jones is responsible.

Frankfurt endorses a revision of the principle of alternate possibilities, one which, he claims, does not conflict with the view that moral responsibility is compatible with determinism:

A person is not morally responsible for what he has done if he did it *only* because he could not have done otherwise (*op. cit.*, 838/9).

On this revision, if a person "really wanted" to do what he did and did it because he really wanted to do so, then he can be morally responsible for his act, even if his act was causally determined. In a

later article, Frankfurt sets up the apparatus of higher-order desires to explain what it is for an agent to act because he really wants to.³

In order better to understand Frankfurt's account of responsibility, we should consider one of Don Locke's criticisms of Frankfurt.⁴ Locke asks us to compare a willing and an unwilling drug addict; both, we suppose, act on irresistible desires to take a drug, but whereas the unwilling addict struggles against this desire, the willing addict embraces it. Locke claims, essentially, that a contented slave is still a slave. Both addicts, insofar as they are addicts and thus act on irresistible desires, act unfreely. He claims that the willing and unwilling addicts are both not responsible for taking the drug, even though it seems that Frankfurt must say that the willing addict is responsible, but the unwilling addict is not.

Although Frankfurt himself is not explicit about this, it is useful, on Frankfurt's theory, to distinguish between two "willing" addicts. The first is actually motivated to take the drug by the thought that, were he to try to abstain, he would experience an irresistible desire for the drug; the irresistibility of the desire (or his belief in its irresistibility) is his sole reason for action. Yet, were he to consider the issue of whether it is good that he has an irresistible desire for the drug, he would be pleased; he doesn't mind being that sort of person. The second willing addict is actually motivated by his desire for the drug; the irresistibility of this desire plays no role in explaining his behavior—he may be quite unaware of its irresistibility. Frankfurt should say of the second but not of the first willing addict that he (like the nonaddict) is responsible for taking the drug.

This follows from two components of Frankfurt's theory:

- (1) What explains the second willing addict's taking the drug may be exactly the same as what explains the taking of the drug by a nonaddict who takes it simply because he likes taking it,⁵

and thus

- (2) What the second willing addict's action reveals about him is the same as what is revealed by the nonaddict. It is not the same as what the action of the unwilling addict reveals (*loc. cit.*).

³"Freedom of the Will and the Concept of a Person," this JOURNAL, LXVIII, 1 (Jan. 14, 1971): 5-20.

⁴"Three Concepts of Free Action: I," *Proceedings of the Aristotelian Society*, suppl. vol. II (1975): 95-112, pp. 99/100.

⁵"Three Concepts of Free Action: II," *ibid.*: 113-125, p. 118.

If one associates responsibility ascriptions with acts that reveal features of an agent's character, then Frankfurt can be defended against Locke's criticism.⁶ On Frankfurt's account of responsibility, if the fact that a desire is irresistible plays a certain role in an agent's deliberation, the agent is not responsible. That is, if an agent believes that a desire is irresistible and if this belief is a part of his reason for acting on the desire, then the agent is not responsible for so acting. But, if the fact that the agent couldn't have done otherwise plays no role in the agent's deliberation, then the agent may be responsible for what he did; this explains why the second sort of willing addict, but not the unwilling addict, may be responsible for taking the drug, and it explains why responsibility is compatible with determinism. Responsibility is compatible with determinism, on this theory, since an agent may not know that determinism is true (if it is) or, even if he does know the truth of determinism, he need not know *which* act (among his alternatives) is determined; the lack of an open alternative need not play a role in the agent's deliberations even if determinism is true (and he knows this) (*ibid.*, 412/3).

Frankfurt's theory of responsibility poses a formidable challenge to the incompatibilist about determinism and responsibility. I shall now discuss two approaches to meeting this challenge. The first strategy claims that Frankfurt has *not* dissociated responsibility from control.

II. THE FIRST CRITICISM OF FRANKFURT

There is considerable disagreement about what sorts of things we hold persons responsible for. Some philosophers believe that we hold agents responsible only for their choices, and not for their acts (if these are different from their choices); some believe that we hold agents responsible for their acts, but not for the consequences of these acts, whereas others insist that we hold agents responsible for the consequences of their acts (as well as the choices and acts themselves). Among philosophers who believe that we hold agents responsible for events of certain sorts, there is disagreement as to the ontological status of these events: are they to be construed as particulars or universals? For the purposes of this discussion, I shall attempt to steer clear of such disagreements, as much as possible. I shall assume that we might hold persons responsible for events construed as particulars or universals and that these events might be acts, the consequences of these acts, or events related to these acts

⁶ This association is emphasized in Robert Cummins, "Could Have Done Otherwise," *The Personalist*, LX, 4 (October 1979): 411-414.

by relations such as Alvin Goldman's "level-generation" (or some similar relation). I want to be very liberal about the metaphysics of responsibility.

Consider again Jones and the "counterfactual intervener," Black. Black wants Jones to vote for Ronald Reagan and would bring it about (through the computer) that Jones decides to vote for Reagan and then does so, should Jones show an inclination to decide to vote for Carter. But Jones decides on his own to vote for Reagan and does so. Is this a counterexample to the principle of alternate possibilities? Is Jones responsible for voting for Reagan, even though he couldn't have done otherwise?

A version of the first approach to denying that Frankfurt's examples undermine the principle of alternate possibilities is developed by Peter van Inwagen (*op. cit.*). I call this the *associationist strategy*—the strategy that insists on the *association* of responsibility with control. The associationist divides the principle of alternate possibilities into one principle applicable to particulars and one applicable to universals:

- (PAP1) A person is morally responsible for a particular event only if he could have prevented it.
- (PAP2) A person is morally responsible for the obtaining of a state of affairs only if he could have prevented the obtaining of that state of affairs.

Consider first the principle applicable to event particulars. Is Jones responsible for the particular event of his voting for Reagan? The associationist adopts the following principle of event individuation:

- (E) x is the same particular event as y if and only if x and y have the same causes.

I shall call this the *essentialist principle*, since it asserts the essentiality of the causal genesis of an event. This principle is supposed to license certain counterfactual inferences; it is supposed to provide a way of identifying events across possible worlds. That is, event a in world α is the same particular event as event b in world β just in case a and b have the same causal antecedents. Donald Davidson argues that events x and y which occur in the actual world are numerically the same just in case they have the same causes and effects; the criterion accepted by van Inwagen applies the Davidsonian intuition to counterfactual talk about events; it provides a way of answering the question, "Under certain circum-

stances which didn't actually obtain, would y have been the same event as the actual event, x ?"

If we accept the essentialist principle, then we can see that the story of Jones and Black might not undermine the principle of alternate possibilities. If we accept this approach we can say that Jones is responsible for the particular act of voting for Reagan because *it* was avoidable; if he had shown signs of deciding otherwise he would have voted for Reagan surely enough, but that would have been a *different* act of voting for Reagan from the actual act, because it would have had different causes. So the particular act of voting for Reagan need not have occurred; it was avoidable.

Now the associationist considers whether Jones is responsible for the fact that Jones votes for Reagan. Is Jones responsible for the obtaining of a certain state of affairs which *could* have been instantiated in various different ways? The associationist (again, following van Inwagen) claims that Jones is *not* responsible for its being the case that he votes for Reagan, and he argues as follows. Jones is *not* responsible for its being the case that either he votes for Reagan on his own or he is caused by Black's computer to vote for Reagan. If he were responsible for this state of affairs, then why not also for something for which he is obviously *not* responsible: its being the case that either he votes for Reagan on his own or $2 + 2 = 4$? But if he is not responsible for its being the case that either he votes for Reagan on his own or he is caused by Black's computer to vote for Reagan, Jones is not responsible for its being the case that he votes for Reagan.

The associationist then responds to Frankfurt as follows: Jones is responsible for the particular voting for Reagan, but *this* was avoidable; Jones is *not* responsible for the fact that he votes for Reagan, and he couldn't have avoided that. Thus the association of responsibility with control is preserved. Further, there is a universal for the obtaining of which Jones is responsible: this universal is "Jones voting for Reagan on his own" (or perhaps, "Jones voting for Reagan in the normal way"). But Jones *could* have prevented the obtaining of this state of affairs; again the association of responsibility with control is maintained.

I find the associationist's approach unconvincing. First, this approach to blocking Frankfurt's claim depends on the essentialist principle for event individuation, but this principle might be challenged.⁷ Consider, for instance, a complex event such as Japan's at-

⁷ W. R. Carter, "On Transworld Event Identity," *Philosophical Review*, LXXXVIII, 3 (July 1979): 443-452.

tack on Pearl Harbor in 1941. Suppose the Emperor of Japan actually initiated the attack by saying, "Commence the attack." If instead he had initiated the attack by saying, "Destroy Pearl Harbor now", it may seem implausible to conclude that that would have been a different attack on Pearl Harbor from the actual attack. It is hard to see how to resolve the question of the adequacy of the essentialist principle of event individuation, but I contend that one need *not* accept this principle in order to deny that Frankfurt's examples undermine incompatibilism about determinism and responsibility; though van Inwagen's criticism of Frankfurt depends on the essentialist principle, I shall present a criticism of Frankfurt that does not depend on any such controversial principle of event individuation.

But there is a second problem with the associationist approach. Consider again whether Jones is responsible for the particular event that is his voting for Reagan. Even if we accept the essentialist principle of event individuation, in what sense could Jones have done otherwise? When we associate responsibility with control, we normally mean that a person is responsible for a particular event only if there is some alternate sequence open to the agent in which he performs a different act (or brings about a different event by performing some act that issues in the event). But if Black's computer were to intervene, it is not clear that Jones would be acting (in the relevant sense) at all. Were Black's computer to intervene and directly manipulate Jones's brain state, we might say that Jones's bodily movements would not in the appropriate sense be *his* actions (or actions at all).

But, even if one insists that there is a sense in which Jones acts in the alternate sequence, we can see that there is a problem for the associationist. When we demand that an agent have control, we mean that there must exist an alternate sequence in which the agent chooses and acts as a result of his character or practical reasoning; but, in Jones's case, in the alternate sequence Jones is caused by an external agent to choose and to act, and the external agent produces a choice and an act *unrelated* to the intention Jones begins to manifest.

For the agent to have control, in the relevant sense, there must be an alternate sequence in which the agent does otherwise as a result of *an appropriate sort* of chain of events. In Jones's case, if his practical reasoning inclines him to vote for Carter, he'll nevertheless be made to choose to vote for Reagan and to vote for Reagan, and his practical reasoning will be superseded. For the agent to have *deliberate control* (in the sense required for responsibility),

there must be an alternate sequence in which there is an action rationalized by his practical reasoning; but in Jones's case, in the alternate sequence his choice and act do not flow from his practical reasoning in the appropriate way.

If we grant to van Inwagen the essentialist principle, then he may have established that a different event would have occurred in the alternate sequence, but he hasn't shown that Jones was capable of exercising deliberate control; van Inwagen thus confuses the ability deliberately to do otherwise with the possibility of something different occurring. Van Inwagen's mistake is to assume that the *only* way in which Frankfurt's example could threaten the principle of alternate possibilities would be by presenting an alternate sequence in which the same event particular (as the actual event) occurs; I have argued that this is false.⁸ The rather attenuated alternate possibility preserved by the associationist doesn't look to me very much like the possibility of deliberately doing otherwise—it doesn't preserve Jones's *control* over the outcome.

One who agrees with the position I have been developing will insist on the following revision of the principle of alternate possibilities applicable to particulars:

(PAP1*) *S* is responsible for event *e* only if there exists some property *F* such that *F*(*e*) and an alternate sequence open to *S* in which *S* brings about $\sim F$ (*e'*) [*e* ≠ *e'*] as a result of an intention to do so.

On (PAP1*), Jones is not responsible for voting for Reagan, since he is incapable of exercising deliberate control. Let '*F*' name the property, "not being caused to vote for Reagan by the computer." If *e* is Jones's actual act of voting for Reagan and *e'* is the act in the alternate sequence, *e* has *F* and *e'* lacks *F*, but in the alternate sequence *S* does *not* bring it about that *e'* lacks *F* as a result of an intention that *e'* lack *F*. Adoption of (PAP1*) shows how van Inwagen's criticisms can be attacked *without* rejecting the essentialist principle of event individuation. An associationist should adopt (PAP1*), but Frankfurt's examples undermine (PAP1*).

In light of these worries about the associationist strategy, it will be prudent for the incompatibilist to develop another response to Frankfurt's examples. This second strategy, which dissociates responsibility from control, is a radical departure from the conventional incompatibilist approach; but I shall argue that it is an appealing position.

⁸ Van Inwagen's assumption plays a crucial role at van Inwagen, *op. cit.*, p. 210.

IV. THE SECOND CRITICISM OF FRANKFURT

The first sort of critic of Frankfurt insists that Frankfurt has not separated responsibility from control, and thus Frankfurt hasn't shown responsibility to be compatible with determinism. Underlying this sort of criticism is the traditional view that, if determinism threatens responsibility, it does so in virtue of undermining control. But I shall argue that an incompatibilist about determinism and responsibility can agree with Frankfurt that an agent might be morally responsible for doing something that he is *not* free to avoid doing. For the incompatibilist, the reason why determinism threatens responsibility *need not* be that it undermines control.

To see this, we should consider again the example of Jones and the "counterfactual intervener," Black. Suppose that the world actually proceeds via a sequence that is *not* causally deterministic; that is, though there are some causal laws, not all events are causally determined. Suppose further that the world proceeds in just the sort of way in which a libertarian says it must, if agents are to be morally responsible for what they do. Although an agent's desires and purposes explain his choices and acts, they do not causally necessitate them; the agent freely "identifies" with some of his desires, where this identification is not causally necessitated. Perhaps the identification is explained in terms of agent causality, although this notion need not play a role. In this sort of world, one in which human choices and actions are not causally necessitated, the libertarian can certainly say that Jones is morally responsible for voting for Reagan, even if Black *would have* brought it about that Jones vote for Reagan, if Jones had shown signs of deciding to vote for Carter. That is, nothing about Frankfurt's example *requires* the actual sequence issuing in the decision and action to proceed in a deterministic way; if it proceeds in a nondeterministic way that satisfies the libertarian, then Jones can be held responsible, even though he could not have done otherwise.

The kernel of truth in Frankfurt's example is that responsibility attributions are based on what happens in the actual sequence; an incompatibilist about responsibility and determinism can agree with this and thus admit that, if determinism is false, an agent who couldn't have done otherwise might be responsible for his action. But of course this doesn't show that *determinism* is compatible with moral responsibility; determinism is a doctrine about what happens in the actual sequence. The point could be put as follows: there are two ways in which it might be true that one couldn't have done otherwise. In the first way, the actual sequence compels the

agent to do what he does, so he couldn't have initiated an alternate sequence; in the second way, there is no actual-sequence compulsion, but the alternate sequence would prevent the agent from doing other than he actually does. Frankfurt's examples involve alternate-sequence compulsion; the incompatibilist about determinism and responsibility can agree with Frankfurt that in such cases an agent can be responsible even while lacking control, but he will insist that, since determinism involves *actual-sequence* compulsion, Frankfurt's examples do not establish that responsibility is compatible with determinism.

So whereas van Inwagen's criticism of Frankfurt contends that he has not dissociated responsibility from control, the approach I am developing concedes this dissociation, but argues that the transition from this dissociation to the compatibility of determinism with responsibility is a spurious transition. This is because the reason why determinism threatens responsibility is not *that* it undermines control, but because of the way in which it undermines control; determinism involves actual-sequence compulsion, and such compulsion might be incompatible with moral responsibility. Thus it is open to an incompatibilist about responsibility and determinism to accept the kernel of truth in Frankfurt's examples without swallowing the compatibility contention, and to defend his position without relying on a controversial principle of event individuation.

The account of responsibility which I have suggested that the incompatibilist might adopt is, like Frankfurt's account, an "actual-sequence" approach; such an approach might take the following general form:

An act (or decision) is unfree (compelled) if and only if (a) it is causally determined, or (b) it is not in an appropriate sense the agent's act (or decision), or (c) it issues from a desire of an intensity *i*, and (i) the desire's having intensity *i* explains why the act occurs, and (ii) it is unreasonable to expect the agent to resist desires of intensity *i*.

This is a plausible libertarian actual-sequence approach to responsibility, which will be elaborated on below; it entails that Jones is morally responsible for voting for Reagan (in the world in which determinism is false), even though Jones couldn't have done otherwise. Of course, if one had information about Black's disposition, then it might be unreasonable to expect Jones to do otherwise, but this sort of expectation would not be unreasonable because of the *intensity* of Jones's desire. Although some philosophers have argued that the causal determination of an event need not entail the possibility, in principle, of predicting its occurrence, this ex-

ample shows that the possibility in principle of predicting the occurrence of an event need not entail its causal determination.

Of course, what is predictable in advance is that Jones will vote for Reagan, though we can't predict which sequence will actually take place. When we predict that an event will occur, we predict that some event of a certain kind will occur. The example shows that the possibility, in principle, of predicting that Jones will vote for Reagan doesn't entail that Jones's particular act of voting for Reagan is causally determined. Although one might worry about determinism because one worries about predictability, a wedge can be driven between determination and predictability; and when the two pull apart, the incompatibilist's fundamental concern is with determination.

V. A POSSIBLE OBJECTION

I wish now to consider a possible objection to the approach I have developed. Someone might say that even in the libertarian world I sketched above in which Jones votes for Reagan on his own and Black does not play any causal role—nevertheless Jones's voting for Reagan is causally determined. This is because, given the setup of the actual sequence including the dispositional properties of agents such as Black, it is true that antecedent states of the world plus the causal laws entail that Jones will vote for Reagan.

Call the actual sequence "*A*," and the counterfactual sequence, in which Black intervenes, "*C*." Say that the obtaining of a state of affairs *P* (construed broadly to include the occurrence of an event or sequence of events) at a time *T* (a point or interval) is nomologically inevitable (at *T*) just in case some state of affairs which obtains prior to *T*, together with the laws of nature, entails that *P* obtains at *T*.

In the example discussed above, neither *A* (the actual sequence) nor, of course, *C* (the counterfactual sequence) is nomologically inevitable (at the relevant time), although Jones's voting for Reagan is nomologically inevitable. Hence it might be claimed that Jones's voting for Reagan is (in the relevant sense) causally determined, and I have not presented an example of an uncompelled (i.e., non-causally-determined) act to which there was no alternate possibility.

But this objection is not cogent. If the doctrine of causal determinism is true, then each state of affairs that actually obtains at each time is nomologically inevitable. Hence, a causally determined act (in the sense that follows from the truth of causal determinism) is and must be a part of a sequence of states of affairs all of which are nomologically inevitable. But the actual sequence *A* (in the example discussed above) is *not* such a sequence; for in-

stance, the state of affairs, "Black's not intervening in Jones's decision" is a non-nomologically-inevitable component of the actual sequence (as is the state of affairs, "Jones's deciding on his own to vote for Reagan"). So even if one thinks that the nomological inevitability of Jones's voting for Reagan is compatible with his moral responsibility for it, it doesn't follow that Jones's responsibility for voting for Reagan is compatible with its being causally determined.

The objection might rest on a confusion between events construed as particulars and as universals. Inevitability is most naturally applied to event universals. (I'm not sure that it is coherent to talk of the inevitability of an event particular.) If it is inevitable that Jones will vote for Reagan, then perhaps it is causally determined that Jones will perform *some* act that is a voting for Reagan; but this leaves it open whether Jones's particular act issues from a sequence that is not causally deterministic. Inevitability, which is a property of event universals, need not entail the causal determination of the particular events that instantiate the universals.⁹

The approach I suggest to responsibility is strongly *path-dependent*. That is, where there are various alternate paths to an event of a certain sort, one focuses on the path that actually leads to the event. There may be no deterministic process on this path, though the event is inevitable. The control model of responsibility is path-independent in the sense that the fact of lack of control is consistent with various paths with radically different properties leading to the event.

The example discussed above also illustrates clause (b) of the account of responsibility. A libertarian will want to say that, since sequence *A* issues in Jones's voting for Reagan, Jones is responsible for so acting, but, if sequence *C* had occurred, then he would *not* have been responsible. But in *C* Jones's voting for Reagan does not issue from a causally deterministic sequence (in the sense sketched above); there are states of affairs (such as "Jones's showing signs of voting for Carter") which are non-nomologically-necessitated components of *C*. But, because of the direct intervention of Black, there is a clear sense in which the decision to vote for Reagan is not (in an appropriate sense) *Jones's* own decision.

⁹The account I have given above of nomological inevitability is a standard account of causal determination. The example shows that this account does *not* adequately capture the notion of causal determination as applied to event particulars. Elsewhere I discuss an example which points to the inadequacy of the "standard" sort of account of causal determinism (the doctrine that all events are causally determined)—Lehrer's "duplicate universe" example; see my "Lehrer's New Move: 'Can' in Theory and Practice", *Theoria*, XLV, 2 (1979): 49-62.

Consider the frequently discussed case of the demonic neurologist who directly manipulates a person's brain to induce all his desires, beliefs, and decisions. Now contrast this case with a counterfactually intervening demonic neurologist (similar to Black), a neurologist who does not actually intervene, but would do so if the person showed any sign of deciding otherwise than he actually decides. Neither the person subject to the actual intervener nor the person subject to the counterfactual intervener *controls* his decisions and actions; but, although both lack control, the second person's acts are *his* in a way in which the first person's are not. An act can be yours without its being up to you; you can be in charge without being in control. Although both lack control, the first person (subject to the actual intervener) is a marionette manipulated by someone else, but the second is not (though he *would* be under other circumstances). A theory of responsibility should reflect this difference; a path-dependent theory (but not a control model) can capture this intuition.

VI. ACTUAL-SEQUENCE THEORIES OF RESPONSIBILITY

On some theories, responsibility is associated with control (freedom to do otherwise); but we have seen that, on these theories, persons might not be held responsible for all the acts for which we believe they are culpable. On other theories, responsibility is associated with willing action (or acting freely); this is Frankfurt's theory, a modified version of which is accepted by Gary Watson.¹⁰ I have suggested a different theory—an actual-sequence approach—which associates responsibility neither with freedom to do otherwise nor with acting freely, but with free, or uncompelled action.

Patricia Greenspan has argued that acts can be unfree even though the agent was free to do otherwise¹¹; I have argued that acts can be free though the agent was *not* free to do otherwise. Also, an agent can act freely (in the Frankfurt-Watson sense) though the act is unfree (in my sense); that is, one can "identify with" an act that is compelled. (I thus preserve Locke's intuition about the happy slave.) Whereas some philosophers (including G. E. M. Anscombe, and most recently, Richard Sorabji¹²) argue that responsibility is associated with acts that are *caused* but not rendered inevitable by those causes, I have suggested that one might allow that responsible acts are sometimes inevitable but uncaused (or not causally determined) events.

¹⁰ "Free Agency," this JOURNAL, LXXII, 8 (April 24, 1975): 205-220.

¹¹ "Behavior Control and Freedom of Action," *Philosophical Review*, LXXXVII, 2 (April 1978): 231-233.

¹² *Necessity, Cause and Blame* (London: Duckworth, 1980).

The traditional account of responsibility claims that what is wrong with determinism is that it erodes control; the account I offer claims that what is wrong with lack of control is that it usually (but not always) indicates actual-sequence compulsion. Thus the reason why lack of control normally rules out responsibility is that it normally points to actual-sequence compulsion. But when lack of control is not accompanied by actual-sequence compulsion, we need not rule out responsibility.

The actual-sequence approach to responsibility which I have sketched solves a problem faced by Frankfurt's theory. Consider Frankfurt's statement in "Alternate Possibilities and Moral Responsibility" (*op. cit.*):

Now if someone had no alternative to performing a certain action [and did indeed perform it] but did not perform it because he could not have done otherwise, then he would have performed exactly the same action even if he *could* have done otherwise . . . Whatever it was that actually led the person to do what he did, or that made him do it, would have led him to do it or made him do it even if it had been possible for him to do something else instead.

Thus it would have made no difference, so far as concerns his action or how he came to perform it, if the circumstances that made it impossible for him to avoid performing it had not prevailed . . . When a fact is in this way irrelevant to the problem of accounting for a person's action it seems quite gratuitous to assign it any weight in the assessment of his moral responsibility (837).

But, if Frankfurt is right, then it seems as though an agent will be responsible for too many things. This criticism, including the following example, was presented to me by Carl Ginet. Suppose someone had no alternative to failing to stop the rain but it is not the case that he failed to stop it *because* he could not have done otherwise; he thought (incorrectly, of course) that he could stop it any time he wanted to by uttering a certain prayer. It follows that he would have failed to stop the rain even if he could have done otherwise. Whatever it was that actually led him to fail to stop the rain would have led him to do so even if it had been possible for him to stop it. Thus it would have made no difference, so far as concerns his failing to stop the rain or how he came to do so, if the circumstances that made it impossible for him to avoid failing to stop the rain had not prevailed. If Frankfurt is right, then when a fact is in this way irrelevant to the problem of accounting for someone's failing to stop the rain, it seems quite gratuitous to assign it any weight in the assessment of his moral responsibility.

So it seems that Frankfurt will need to say that this agent is mor-

ally responsible for failing to stop the rain; other persons might be morally responsible for failing to stop the Earth's rotation, etc. But a theory of moral responsibility which has this sort of result is inconsistent with our intuitions about moral responsibility.

On the actual-sequence account of responsibility I sketched above, we can explain why the agent is not morally responsible for failing to stop the rain. Even in a libertarian world (in which human choices or actions are not causally necessitated), there are certain causal laws governing natural phenomena which (given present technology) entail that no agent can stop the rain by uttering a prayer. Thus the actual sequence of events proceeds in such a way that the agent's not stopping the rain is causally necessitated. Similarly, the physical laws that obtain (even in a libertarian world) are such that (given present technology) it is causally necessitated that no person can stop the Earth's rotation. If we accept the claim that actual-sequence causal necessitation is incompatible with responsibility, we can explain why no agent is morally responsible for failing to stop the rain. An actual-sequence model, but *not* of the Frankfurt-Watson sort, can explain why agents are not responsible for (say) failing to stop the rain.

VII. CONCLUSION

It might be objected against the approach I have sketched that causal determination needn't constitute *compulsion*; after all, it is consistent with Frankfurt's presentation of his example that Jones's voting for Reagan is actually causally determined. Hence it might seem that responsibility is obviously compatible with determinism. Though this may be so, it is crucial to see that the battleground has now shifted; the question now is whether causal necessitation in the actual sequence constitutes compulsion, that is, whether causally deterministic actual sequences lack components required for moral responsibility. The compatibilist about responsibility and determinism will say no, but the incompatibilist will say yes; they will both point to features of causally deterministic sequences in order to establish their positions, but Frankfurt-style examples concerning what would happen in alternate sequences will not support either position. Frankfurt-style examples will be irrelevant to resolving *this* dispute.

In summary, Frankfurt's examples seem to separate responsibility from control and hence to show responsibility compatible with determinism even if determinism is inconsistent with control. One strategy of response is to challenge the divorce of responsibility from control. I discussed two problems with this strategy: the tenuousness of the essentialist principle of event individuation and the

fact that alternate-sequence interventions of the Frankfurt sort do not establish that the agent has the deliberate control required for responsibility. I developed a second strategy of response: this strategy concedes the dissociation of responsibility from control but argues that this dissociation needn't entail the consistency of determinism with responsibility.

I have not argued *for* incompatibilism about determinism and responsibility; I have had the more modest project of showing how the incompatibilist is not forced into inconsistency by Frankfurt's examples. Both the compatibilist and the incompatibilist alike can unite in conceding that enough information is encoded in the actual sequence to ground our responsibility attributions; as philosophers we need to decode this information and see whether it is consistent with deterministic causation.

JOHN MARTIN FISCHER

Yale University

COMMENTS AND CRITICISM

A PROBLEM FOR DEFEASIBILITY THEORIES*

WHAT all defeasibility theories of inferential knowledge have in common is the belief that Gettier and Gettier-inspired counterexamples to the traditional definition of knowledge involve a defeater which prevents such cases from being true cases of knowledge. In order to say that *S* has inferential (or nonbasic) knowledge of some proposition, it is not enough that *S* have a justified true belief: *S*'s justification must also not be defeated.

Just what constitutes defeat has been the subject of much controversy, but underlying the various attempts that have been made at giving such a characterization, one intuition stands out—that a defeater is such that it is true and if a justified true believer were aware of it, then he would no longer be justified in believing whatever it is he believes (the reason being that the defeater undermines his original evidence for his belief).

* I am indebted to Walter Horn, Ernest Sosa, James Van Cleve, Peter Klein, Carl Ginet, and the members of Ginet's 1980 Theory of Knowledge seminar for helpful comments on this topic.