Freedom and Miracles

JOHN MARTIN FISCHER
YALE UNIVERSITY

There is a "new incompatibilism" in the air.[1] One argument in the family of "new" arguments for incompatibilism might be called the "modal" argument for incompatibilism. I have presented a version of the modal argument (following Carl Ginet), and Peter Van Inwagen has recently presented basically the same sort of argument.[2] (I shall also refer to the modal argument for incompatibilism as the "Ginet/Van Inwagen argument".) Elsewhere, I argued that the Ginet/ Van Inwagen argument is importantly incomplete; there is no *single* interpretation of the premises of the argument on which it is uncontroversially sound.[3] Both Ginet and Van Inwagen have defended the modal argument against the sort of criticism which I developed.[4] In this paper, I propose to look again at the modal argument for incompatibilism. I shall present a simplified version of the argument, and I intend to show that the compatibilist can explain the examples adduced by Ginet and Van Inwagen in defense of their argument by *narrower* principles than those used in the argument, principles which *don't* support incompatibilism. Thus, I shall defend a version of my original claim that the incompatibilist's argument is *incomplete*. I do not here purport to defend compatibilism; rather, my project is to make clear where the *gaps* are in the modal argument for incompatibilism. Further, I shall point out that, whereas the incompatibilist's argument is not entirely convincing, it has an important virtue — it exposes a commitment of certain compatibilists which appears to stand in need of some explanation and justification. In the course of this discussion, I hope to shed light on the relationship between human freedom and the laws of nature. I shall argue that the compatibilist has a distinctive and not obviously false view about this relationship.

NOÛS 22 (1988) 235-252

I. THREE INTERPRETATIONS

The modal argument makes use of a propositional-operator-form, "$N_{St}p$", which abbreviates

> "p and it was not in agent S's power at time t to cause it not to be the case that p."[5]

Inserting universal quantification on the agent-variable, "S", we create another propositional-operator-form, "N_tp", which is

> "p and, for any S, it was not in S's power at t to cause it not to be the case that p."

Inserting universal quantification on both variables, we get, "Np":

> "p and, for any S and any t, it was not in S's power at t to cause it not to be the case that p."

The argument assumes the validity of the following principles:

(a) It is logically necessary that p
$\therefore$ Np

(b) *N* (if p then q) $\quad$ N_t (if p then q)
Np $\quad$ N_tp
$\therefore$ *Nq* $\quad$ $\therefore$ N_tq

(c) The laws of nature entail that p.
$\therefore$ Np

(d) Everything needed to make it the case that p happened before t.
$\therefore$ N_tp

Given these assumptions, the argument shows that if causal determinism is true, then: for any truth p as to what happens in the world, Np. That is, causal determinism is incompatible with freedom to do otherwise.

Peter Van Inwagen's version of the modal argument for incompatibilism uses the following construal of the propositional-operator, "Np":

> "p and no one has, or ever had, any choice about whether p".[6]

I am not certain what it is to "have a choice" about whether p, but I take it that one "has a choice" about whether p only if one can (in some suitable sense) cause it not to be the case that p. That is, I shall interpret Van Inwagen's construal of the relevant modal operator as equivalent to Ginet's construal.

I shall now set forth one version of the Ginet/Van Inwagen argument. The argument uses the following defined symbols:

Determinism (D) is the thesis that, for any given time, a complete description of the state of the world at that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time.

"L" expresses the conjunction of all the (true) laws of nature.

"B" expresses a proposition that details the total state of the world at some remote past time before any moral agent existed.

"A" expresses any proposition you like that attributes to you the performance of one of your recent actions that seemed to you to be in your power not to perform (up to the time you performed it).

Here is the argument:[7]

1. Determinism obtains.	hypothesis
2. It is logically necessary that $(L \& B \supset A)$	1, def. (D)
3. $N(L \& B \supset A)$	2, a
4. $N(L \supset [B \supset A])$	3
5. $N(L)$	c
6. $N(B \supset A)$	4, 5, b
7. $N(B)$	d
8. $N(A)$	6, 7, b

The argument shows (if sound) that if determinism is true, no person is ever free to do other than what he actually does.

In evaluating the Ginet/Van Inwagen argument, it is critical to disambiguate the key phrase, "It was in S's power at t to cause it to be the case that p".[8] I shall distinguish *three* different interpretations of the phrase; these different interpretations will allow us to articulate and distinguish various compatibilist responses to the argument.

Let us stipulate that sometimes a proposition "mentions" or "contains" an event e, and that we can identify this event. A proposition of the form, "e occurs (doesn't occur) at t" is atomic. Here "e" is an event broadly construed, where states of affairs, acts, and omissions are all considered events. (Of course, a state of affairs obtains (or fails to obtain) — it does not "occur" (or fail to occur). For simplicity's sake, I shall speak broadly of events (including states of affairs) occurring or failing to occur.) Also, I shall assume here that actions are events, and a proposition which says that an agent performs an act will be assimilated to a proposition saying that the event occurs, for the purpose of distinguishing between atomic and non-atomic propositions. When an agent performs an act, I shall say that he causes an appropriate event to

occur. Thus, when A scratches his back, he causes the event, "A's back being scratched", to occur. Nothing of substance in *this* discussion will hang on these simplifying assumptions.

Clearly there are propositions which are non-atomic; particularly pertinent to our evaluation of the incompatibilist argument are propositions which are *conjunctions* of atomic propositions. I believe that it is important to distinguish different conceptions of the ability to cause *conjunctive* propositions to obtain, in evaluating the Ginet/Van Inwagen argument; thus far, commentators on the argument have been insufficiently attentive to the distinction between atomic and non-atomic propositions and their role in the argument. What exactly is it to cause a conjunction of atomic propositions to obtain? Must one cause *each* of the conjuncts to obtain? If Jack scratches his back on Monday, has he caused the proposition, "John F. Kennedy was assassinated and Jack scratches his back on Monday" to obtain? I claim that there are (at least) the following three different interpretations of "S had it in his power at t to cause it to be the case that p" which are relevant to the Ginet/Van Inwagen argument:

(i) (a) p is equivalent to the proposition, "e_1 occurs (doesn't occur) at t_1" and S had it in his power at t to cause e_1's occurrence (non-occurrence) at t_1, or
(b) p is equivalent to the proposition, "e_1 occurs (doesn't occur) at t_1 and e_2 occurs (doesn't occur) at t_2", and
(1) e_1 occurs (doesn't occur) at t_1,
(2) S had it in his power at t to cause e_2's occurrence (non-occurrence) at t_2, and
(3) If S were to cause e_2 to occur (not occur) at t_2, then e_1 would occur (not occur) at t_1.

(ii) p is of any form, and S had it in his power at t to cause some event e such that if e were to occur, p would obtain.

(iii) is exactly the same as (i), except for the third clause of (b):
(b) (3) If S were to cause e_2 to occur at t_2, he would not thereby cause e_1 to not occur (occur) at t_1.[9]

Note that, on all of the accounts, an agent can cause it to be the case that p just in case he can cause it to be the case that a proposition logically equivalent to p obtains. Also, let us assume that an agent can make it not the case that p (can make it the case that p is false) if and only if he can make it the case that not-p.

It might be useful to point out that (ii) is to be so understood that it holds if p is *any truth* and e is any event that S *did in fact* cause to occur. So I had it in my power yesterday to cause it to be the case that the United States is in North America. Thus (ii)

yields a rather broad sense of the technical phrase, "S can cause it to be the case that p". It is quite clear how interpretation (ii) differs from both (i) and (iii): in the case of a proposition which says that e_1 occurs at t_1, all (ii) requires is that S be able to cause *some* event such that if it were to occur, e_1 would occur at t_1, but S need not be able to cause e_1's occurring at t_1. But how do (i) and (iii) differ?

In order to see this, let us consider the following case.[10] Suppose that determinism is true, and that some condition C obtained at t_1 which causally necessitated A's failing to scratch his back at t_2. Imagine also that the case is one in which a compatibilist wants to say that the conditions did not *prevent* A from scratching his back at t_2 — he could, intuitively, have done so. Further, suppose that the compatibilist believes (as he might, but need not) that had A scratched his back at t_2, then C wouldn't have obtained at t_1. That is, the kind of compatibilist we are now considering holds the "backtracking" counterfactual, "If A had scratched his back at t_2, then C wouldn't have obtained at t_1" to be true (rather than the "local-miracle" counterfactual, "If A had scratched his back at t_2, then some natural law which actually obtained wouldn't have obtained").

Consider whether A can at t_2 cause it not to be the case that condition C obtained at t_1. On *both* interpretations (i) and (iii), the answer is *no*: there is no backwards causation. But now consider whether A can at t_2 cause it not to be the case that if condition C obtained at t_1, then A would refrain from scratching his back at t_2. On interpretation (i), the compatibilist (who accepts the truth of the "backtracker") will again say *no*. In order to cause it not to be the case that if condition C obtained at t_1, then A would refrain from scratching his back at t_2, S must be able to cause it to be the case that not — (If condiction C obtained at t_1, then A would refrain from scraching his back at t_2), i.e., he must be able to cause it to be the case that (condition C obtained at t_1 and A scratches his back at t_2). But, whereas A can at t_2 cause it to be the case that he scratches his back at t_2, if he were to do so (assuming the backtracking conditional), it would have been the case that condition C *didn't* obtain at t_1. Thus, via condition (ib3), A can't at t_2 cause it to be the case that condition C obtained at t_1 and A scratches his back at t_2. The intuition behind interpretation (i) is that, in determining which conjunctive propositions an agent can cause to obtain, one demands that the agent be able to cause at least *one* "mentioned" event (contrary to interpretation (ii)), and that if he does so, the the other mentioned event would *also* occur (contrary to interpretation (iii)).

In contrast, on interpretation (iii), A *can* at t_2 cause it to be the case that condition C obtained at t_1 and A scratches his back at t_2. This is because A can at t_2 scratch his back at t_2, condition C actually obtained at t_1, and A's scratching his back at t_2 wouldn't *cause* condition C's not having obtained at t_1. The intuition behind interpretation (iii) is that, if some proposition r actually obtains, and one can cause s (a proposition compatible with r) to be true, then it seems that one can cause the proposition "r and s" to be true (and cause "if r then not-s" to be false).

Having distinguished these three interpretations of the key phrase, "S had it in his power at t to cause it to be the case that p", we can now map out various compatibilist responses to the basic argument for incompatibilism (which will be developed further below). The compatibilist begins by pointing out that common usage does not single out one interpretation of the key phrase. If (i) is employed, then whereas principles (a) and (d) are valid, either (b) or (c) is unacceptable.[11] If (ii) is employed, then (a) and (b) are valid, but not (both) (c) and (d). And if (iii) is employed, then (a), (b), and (d) are valid, but *not* (c). Thus, there is no single interpretation on which *all* of the principles are uncontroversially valid, and the basic argument for incompatibilism is not uncontroversially sound.

The compatibilist, as I have presented him, need not deny that there are *some* interpretations on which principle (b) is valid. Specifically, he is quite willing to grant that it is valid on (ii) (as well as (iii)); but this concession is not troublesome, since he wishes to deny that, on (ii), both (c) and (d) are valid (and that, on (iii), (c) is valid). The first claim I wish to discuss is the compatibilist's rejection of (bi); later, I shall discuss the rejection of (ciii).

II. THE DENIAL OF PRINCIPLE (bi)

We have already briefly discussed the sort of case on the basis of which a compatibilist might deny the validity of (bi). Let's spell it out more explicitly here. Suppose that causal determinism is true and that the obtaining of condition C at t_1 causally necessitates A's failing to scratch his back at t_2. Suppose also that we grant that the "backtracking" counterfactual, "If A had scratched his back at t_2, then condition C wouldn't have obtained at t_1" is true. (Of course, whether this counterfactual is true is controversial; I shall discuss approaches which deny that it is true below.[12]) A compatibilist might describe the case as follows.

(1) N_{At_2} (Condition C obtained at t_1)

is true, on interpretation (i), since no agent can initiate a causal chain extending backwards in time. Further,

(2) N_{At_2} (If condition C obtained at t_1, then A failed to scratch his back at t_2.)

is also true, since if A were to scratch his back at t_2, then condition C *wouldn't* have obtained at t_1. Thus, A can't at t_2 cause it to be the case (on interpretation (i)) that (condition C obtained at t_1 and A scratches his back at t_2). However, there is no reason to deny that A can scratch his back at t_2; indeed, if the causal conditions are of the appropriate sort, the compatibilist will insist (and no reason had been adduced to deny) that he *can* do so. Hence, although (1) and (2) are true in the example,

(3) N_{At_2} (A fails to scratch his back of t_2)

is *false*, according to the compatibilist. Thus, (bi) is held to be *invalid*. What is important to the structure of the case, as described by the compatibilist, is that A *can* at t_2 cause *some* event (his scratching his back) such that if it were to occur, then condition C *wouldn't* have occurred. Thus, though

(1) N_{At_2} (Condition C obtained at t_1)

is true on interpretation (i), it is *false* on interpretation (ii).

That is, the case is of the following form. On interpretation (i), both

(1) $N_{St}p$

and

(2) N_{St} (If p, then q)

are held to be true. But the compatibilist wishes to deny

(3) N_{St} (q).

In other words, the compatibilist wishes to deny (bi). Now, since the compatibilist *accepts* (bii), and it is a fact that for any proposition p, if Np is true on interpretation (ii), Np is true on interpretation (i), the compatibilist must deny either (1) or (2) on interpretation (ii), if he is to deny (3), on interpretation (i). But given our assumptions, it is clear that (1) is false on interpretation (ii), and thus the compatibilist can deny (3), on interpretation (i), and thus deny the validity of (bi). In this proposed counterexample to (bi), both of the premises are true, on interpretation (i), but it is crucial that one be *false* on interpretation (ii); otherwise, (bii) would apply.

So the compatibilist will agree to the truth of the premises (1) and (2), on interpretation (i), but deny (bi) and thus deny (3i); and he will accept (bii) but deny (1), on interpretation (ii), and thus deny (3ii). Whereas Ginet admits that he has no argument against the latter sort of compatibilist, he would claim, I believe, that (bi) is valid. And it must be admitted that the case which I have sketched above is essentially controversial; it has the *structure* required for a counterexample to (bi), but to assert that it is a convincing counterexample would be question-begging — it assumes

straightforwardly that incompatibilism is false. Further, both Ginet and Van Inwagen have produced a number of examples which appear to have the form of (bi) and which *don't* presuppose the *truth* of incompatibilism. They argue that, since there are a number of plausible examples which seem to have the form of (bi) and which don't presuppose the truth of incompatibilism, and the only counterexamples (so far produced) to (bi) presuppose the *falsity* of incompatibilism, this is some reason, in the absence of other considerations against (bi), to adopt (bi).[13] In order to assess this sort of argument, we must consider the Ginet/Van Inwagen examples:

> If it rained hard in Ithaca this morning then there was water flowing in Cascadilla gorge this afternoon, and it was never in anyone's power to cause it to be the case that it rained hard in Ithaca this morning and there was not water flowing in Cascadilla gorge this afternoon.
> It rained hard in Ithaca this morning, and it was never in anyone's power to cause it to be not the case that it rained hard in Ithaca this morning.
> ∴ There was water flowing in Cascadilla gorge this afternoon and it was never in anyone's power to cause it to be the case that there was not water flowing in Cascadilla gorge this afternoon.
>
> If Alice has asthma then she sometimes has difficulty breathing, and it was never in anyone's power to cause it to be the case that she has asthma and does not sometimes have difficulty breathing.
> Alice has asthma and it was never in anyone's power to cause it to be the case that Alice does not have asthma.
> ∴ Alice sometimes has difficulty breathing, and it was never in anyone's power to cause it to be the case that she does not sometimes have difficulty breathing.
>
> If the sun explodes in 3000 A.D. then all life on earth will be destroyed in 3000 A.D., and it was never in anyone's power to cause it to be the case that the sun will explode in 3000 A.D. but not all life on earth will be destroyed then.
> The sun will explode in 3000 A.D., and it was never in anyone's power to cause it to be the case that the sun will not explode in 3000 A.D.
> ∴ All life on earth will be destroyed in 3000 A.D., and it was never in anyone's power to cause it to be the case that not all life on earth will be destroyed then.[14]

What exactly do these examples establish? Is it true, as Ginet and Van Inwagen assert, that in the absence of a counterexample to (bi) that doesn't presuppose the falsity of incompatibilism, such examples make it reasonable to accept (bi)? I think not. The problem is that *all* of the Ginet/Van Inwagen examples can be explained by (bii) as easily as by (bi); that is, none of the examples supports

(bi) *rather than* (bii). And given that the compatibilist can accept (bii), the incompatibilist does not *yet* have a convincing argument, simply in virtue of adducing the examples. If a group of examples supports principle P2 as strongly as P1, but only P1 renders a given argument valid, then simply in virtue of adducing the examples, one hasn't yet supported the argument.

In order to assess the examples, let's focus on the third example. Of course, it's the incompatibilist's assertion that if the premises of the argument are both true, then the conclusion must be true. Consider the second premise:

> The sun will explode in 3000 A.D., and it was never in anyone's power to cause it to be the case that the sun will not explode in 3000 A.D.

Now it seems that if the premise were true, on interpretation (i), it would *also* be true, on interpretation (ii). That is, because of the content of the premise, if it is true that

> The sun will explode in 3000 A.D., and it was never in anyone's power to initiate a causal sequence issuing in the sun's not exploding in 3000 A.D.,

then it would *also* be true that

> The sun will explode in 3000 A.D., and it was never in anyone's power to cause some event e, such that if e were to occur, then the sun would not explode in 3000 A.D.

The same sort of point holds for the first premise. That is, because of the *content* of the premises in the example, the example can be seen as having the true premises

(1) Np

and

(2) N (If p, then q)

on *both* interpretations (i) and (ii). Of course, if this is so, then the compatibilist *must* accept

(3) Nq

on interpretation (ii), since he accepts (bii), and thus also accept (3), on interpretation (i). (Remember that, for any proposition p, if Np is true on (ii), it is true on (i).)

But clearly we do *not* have here an example which supports (bi) *over* (bii). In order to support (bi) over (bii), an example must have the following sort of structure. It must be a case where both premises

(1) Np

and

(2) N (if p, then q)

are true, on interpretation (i), but one of the premises, say (1), is *false* on interpretation (ii), but where we would *still* say that

(3) Nq

is *true*, on interpretation (i). If the premises are both true on (ii), then the fact that the conclusion is true on (i) might be explained by the fact that (bii) is valid and that, for any proposition, if Np is true on (ii), it is true on (i). Remember that, in the compatibilist's example discussed above,

(1) N_{At_2} (Condition C obtained)

was claimed to be true on interpreation (i) but *false* on interpretation (ii). Thus, (bii) does not apply, and the compatibilist can deny

(3) N_{At_2} (S failed to scratch his back at t_2)

on interpretation (i).

It is striking that *none* of the Ginet/Van Inwagen examples has the appropriate structure to support (bi) *over* (bii). In all of the examples, the content of the premises is such that, if they are true on (i), they are also true on (ii). Thus, for all Ginet and Van Inwagen have said, (bii) and not (bi) is valid. So, they do not *yet* have a convincing argument for incompatibilism. They demand that the compatibilist provide a non-question-begging counterexample to (bi), in the absence of which we accept (bi). But this is unfair. It is appropriate to respond that, in the absence of any examples which support (bi) rather than (bii), we needn't accept (bi), and thus we needn't accept incompatibilism.

Not all compatibilists will wish to adopt the approach developed above. This is because not all compatibilists want to accept that, in the examples such as that of A's failing to scratch his back, the backtracking counterfactual, "If A had scratched his back at t_2, then condition C wouldn't have obtained at t_1" is true. Rather, they argue that if A had scratched his back at t_2, then C still would have obtained at t_1, but some natural law which actually obtained *wouldn't* have obtained. That is, had A scratched his back at t_2, a "miracle" (violation of *actual* natural law) would have occurred (after t_1 but prior to t_2). Such a compatibilist must accept (ai), (bi), and (di), but he will reject (ci). This approach can be motivated by the same sort of consideration as motivates the compatibilist's rejection of (ciii). I now turn to this sort of view, discussing the rejection of (ciii).

III. THE DENIAL OF PRINCIPLE (ciii)

Suppose now that the incompatibilist employs interpretation (iii) of the key phrase, "S had it in his power at t to make it the case that p". It is evident that the compatibilist will not wish to deny (biii). To see this, consider again the example of A's failing to scratch

his back. Given that the compatibilist wishes to say that A *can* at t_2 scratch his back, i.e., he wants to deny

(3) N_{At_2} (A fails to scratch his back at t_2),

he must *also* deny, on interpretation (iii),

(2) N_{At_2} (If condition C obtained at t_1, then A fails to scratch his back at t_2).

This is because, in the example, condition C actually obtained at t_1, S can (according to the compatibilist) at t_2 scratch his back at t_2, and his so doing wouldn't *cause* condition C's not obtaining at t_1. Thus, the compatibilist is not in a position to accept (1) and (2) but deny (3) — a denial of (3) must also bring a denial of (2). But, whereas, on interpretation (iii), the compatibilist must accept (b), he will deny (c).

In order to assess (ciii), we need to look rather carefully at the relationship between natural laws and human freedom. My claim will be, as above, that the examples adduced by Ginet and Van Inwagen can be explained by a *narrower* principle than (ciii), a principle which doesn't lead to incompatibilism. Let us consider the sorts of examples adduced by Van Inwagen and endorsed by Ginet as supporting (ciii). Van Inwagen claims that from the proposition that

(A) Jones, a physicist, can construct a particle accelerator that would cause protons to travel at twice the speed of light,

it would be correct to deduce that

(B) It is not a law of nature that nothing ever travels faster than the speed of light.[15]

Also, Van Inwagen says:

> Suppose a bureaucrat of the future orders an engineer to build a spaceship capable of travelling faster than light. The engineer tells the bureaucrat that it's a law of nature that nothing travels faster than light. The bureaucrat concedes this difficulty, but counsels perseverence: 'I'm sure', he says, 'that if you work hard and are very clever, you'll find some way to go faster than light, even though it's a law of nature that nothing does.' Clearly his demand is simply incoherent.[16]

Similarly, Van Inwagen asks us to suppose that, as a matter of fact, if any human being is deprived of vitamin C, he develops scurvy. Now Van Inwagen says:

> . . . suppose also that there is a certain group of biologists and bureaucrats who want to institute a program of selective breeding that is intended to produce a population of human beings who are able to get along without vitamin C. Let us further suppose that wiser counsel prevails, and these people are disuaded from this idiotic

> and immoral undertaking: but suppose that *if* they had been allowed to have their way, they (or their descendants) *would have* succeeded: eventually there *would have been* human beings who did not [need vitamin C, in order to avoid scurvy]. In that case, it seems to me, we should hardly want to say that [it is a law of nature that if a human being is deprived of vitamin C, he gets scurvy].[17]

The conclusion which Ginet (and I believe also Van Inwagen) draws from these examples is that (ciii) is true. But note that all of the Ginet/Van Inwagen examples are cases in which it is intuitively plausible to say that a person can't perform an act *which itself would be or cause a law-breaking event.*[18] My claim, following Lewis, is that the compatibilist can agree that no human agent can perform an act that would *be* or *cause* a law-breaking event, but nevertheless deny (ciii). That is, the compatibilist can deny that no agent can cause to be false (in the relevant sense) a proposition entailed by the laws of nature: this is because one might be able to cause to be false a proposition entailed by the laws of nature *without* performing some act which is or would cause a law-breaking event.

There are two sorts of approaches to denying (ciii), depending on one's views about counterfactuals. Let us again consider the case of A's failing to scratch his back, and we shall first assume that the "backtracking" counterfactual is true in the example. The compatibilist claims here that A can at t_2 scratch his back at t_2. Thus, he must deny, on interpretation (iii),

(2) N_{At_2} (If condition C obtained at t_1, then A fails to scratch his back at t_2)

But since "If condition C obtained at t_1, then A failed to scratch his back at t_2" is, by hypothesis, entailed by the laws of nature, the compatibilist must reject (ciii). But he will point out that if A had scratched his back at t_2, condition C wouldn't have obtained at t_1 (but all the actually obtaining natural laws *would have obtained*). Thus, if A had scratched his back at t_2, his scratching his back wouldn't itself have been or caused an event which breaks an actually obtaining law; indeed, if A had scratched his back at t_2, there would not *ever have been* any violation of an actually obtaining law — rather, the past would have been different from what it actually was. Thus, from the claim that one can't perform an act which itself would be or cause a violation of natural law, it doesn't follow that A couldn't have at t_2 scratched his back at t_2. Here is a case, then, in which it might be true that an agent can't perform an act which is or would cause a violation of natural law, but can "cause to be false" a proposition entailed by the natural laws. None of the Ginet/Van Inwagen examples *establishes* that this sort of compatibilist position is untenable.

Note that, whereas it is in general true that if Np is true on interpretation (ii), it is also true on (i), it is *not* in general true that if Np is true on (ii), it is also true on (iii). It might be the case, as the example shows, that an agent can't so act that "if r, then s" would be false, but that he can cause it to be false (in sense (iii)) that if r, then s.

Now, let us consider the example, supposing that the backtracking counterfactual is false. It is here imagined that if A had scratched his back at t_2, then condition C would still have obtained at t_1, but some law which actually obtained *wouldn't* have obtained at some time after t_1 but prior to t_2. Again, it is open to the compatibilist to claim that A can at t_2 scratch his back at t_2, and thus to deny (ciii). The reason is that, whereas if A had scratched his back at t_2, some actually obtaining law would have had to have been (at some point) violated, A's scratching his back wouldn't itself be or cause a law-breaking event. Of course, A couldn't have scratched his back faster than the speed of light! That is, if in scratching his back, A would have been moving his fingers faster than the speed of light, then his scratching his back at t_2 would *itself* have been a law-breaking event, and thus he couldn't have scratched his back at t_2.[19] And of course, A couldn't have built a machine which would have scratched people's backs at faster than the speed of light. If in scratching backs, the machine moved faster than the speed of light, then A would have *caused* some event which would be a law-breaking event. But what is envisaged is the normal sort of back-scratching, which wouldn't itself be or cause a law-breaking event. Again, the Ginet/Van Inwagen examples do not provide any sort of argument against this compatibilist position; these examples can be explained by the principle that one can't perform an act which itself would be or cause a violation of natural law, a principle which is narrower than (ciii) and which doesn't lead to incompatibilism.

Let us look at the example a bit more carefully. We suppose that A can at t_2 scratch his back at t_2, and in doing so, A would cause the event, "A's back is scratched at t_2", to occur. Since A can at t_2 cause this event to occur, then (on interpretation (iii)) A can at t_2 cause it to be the case that condition C obtained at t_1 and A's back is scratched at t_2. So A can cause to be the case a proposition incompatible with the laws of nature. But does it follow that A can cause the "compound event", "C obtained at t_1 and A's back is scratched at t_2"? If so, then A can cause a law-breaking event, and the distinction between performing an act which would be or cause a law-breaking event and causing to be false a proposition entailed by the laws of nature would collapse.

I claim that it does *not* follow from A's being able at t_2 to cause

it to be the case that Condition C obtained at t_1 and A's back is scratched at t_2 that A can cause the compound event to occur. I take it that causation is a relation between events which are such that no part of the effect *precedes* the cause. Since the event, "Condition C obtained at t_1" is a part of the conjunctive event, "Condition C obtained at t_1 and A's back is scratched at t_2", and since it precedes t_2, A's activity at t_2 *cannot* cause the conjunctive event to occur, even though it does cause to be the case the proposition stating that the conjunctive event occurs. Similarly, whereas one doesn't cause the sun to shine, and therefore A doesn't cause the compound event, "The sun shines at t_1 and A's back is scratched at t_2", A does cause it to be the case that the sun shines at t_1 and A's back is scratched at t_2 (on interpretation (iii)).

I have said that a compatibilist can explain Van Inwagen's example by reference to the principle that no one can perform an act which would *be* or *cause* a law-breaking event. But what exactly is a "law-breaking event"? I mean by "law-breaking event" an event whose occurrence is logically incompatible with the laws of nature. Thus, the occurrence of the event of flying faster than the speed of light is logically incompatible with the law of nature that nothing travels faster than the speed of light. But whereas the compatibilist's principle seems adequate to explain Van Inwagen's examples, it is not so apparent that it can also explain an example due to Ginet.[20] Suppose that it is true (as it seems to be) that if I were to cause this typewriter to be on Jupiter one minute from now, then I would have caused protons to travel faster than the speed of light. Now, it is quite obvious that I *can't* now cause this typewriter to be on Jupiter one minute from now. But the occurrence of "the typewriter's being on Jupiter one minute from now" is *not* logically incompatible with the laws of nature (but only with the laws of nature *and* the state of the universe now). So, how can the compatibilist explain why it is that I can't now cause my typewriter to be on Jupiter one minute from now?

In order to explain Ginet's example, I need to provide a more explicit interpretation of the compatibilist's principle. The principle is to be interpreted as follows: No agent can at t_1 perform an act a at t_2 such that (i) a would be identical to some event b's occurring at t_3 or a would cause some event b's occurring at t_3, and (ii) b's occurring at t_3 is inconsistent with the laws of nature. ($t_1 \leq t_2 \leq t_3$) so I can't now do anything to cause my typewriter to be on Jupiter one minute from now, since my doing that now would cause protons to travel faster than light in the next minute, which is inconsistent with the laws of nature. And this interpretation of the principle is consistent with compatibilism: if determinism is true

and I don't scratch my back at t_2, nevertheless, I may be able at t_2 to scratch it at t_2, since my scratching it at t_2 needn't be identical to or cause anything inconsistent with the laws of nature. Rather, it might be the case that if I were to scratch my back at t_2, some law would have been violated prior (perhaps immediately prior) to t_2.

There may however be a residual uneasiness with the compatibilist line I have been developing. Perhaps the incompatibilist will *grant* the distinction between performing an act which would be or cause a law-breaking event and causing to be false a proposition entailed by the laws of nature. And perhaps he will even grant that, strictly speaking, all his examples conclusively establish is that one can't perform an act which would be or cause a law-breaking event; they do not establish the stronger claim that one can't cause to be false a proposition entailed by the laws of nature. But his crucial claim is that it is plausible to accept the stronger claim if one accepts the weaker claim. That is, why exactly should we discriminate between causing a law-breaking event and causing it to be false that a proposition entailed by a law of nature obtains?[21]

This is, I think, a legitimate and important question. Whereas it is true that there is a gap in the incompatibilist's argument, the incompatibilist has at least shown that the compatibilist (of a certain sort) is committed to a distinction which might appear to be rather "fine". There is an incompleteness in the incompatibilist's argument, and thus it is *open* to a person to reject it, but in rejecting it, he may be committed to a distinction between claims about our abilities which is hard to explain and justify. I believe that it is a virtue of the Ginet/Van Inwagen style of argument that it exposes this commitment of the compatibilist (who denies (ciii) and accepts the "local-miracle counterfactual").

David Lewis says that, whereas compatibilism seems to be committed to an incredible consequence, we must distinguish what's incredible from what's the consequence. That is, we must distinguish (in the framework I have been discussing) the claim that one can perform an act which would be or cause a law-breaking event from the denial of (ciii) — the claim that one can, in the relevant sense, cause a proposition entailed by the laws of nature to be false. Van Inwagen claims, on a number of occasions, that "it is a feature of laws of nature that they impose limits on our abilities".[22] This may be so, but I have argued that the *form* of the limits may be different from what is envisaged by the incompatibilist.[23]

IV. CONCLUSION

I have attempted to defend the thesis that there is no single interpretation on which all of the incompatibilist's principles are uncon-

troversially valid. On interpretation (i), neither (b) nor (c) is clearly valid; on (ii), neither (c) nor (d) is obviously valid; and on (iii), (c) is not uncontroversially valid. I have not claimed that these critical principles are *invalid* — to do so, on the basis of the examples discussed above, would be unfair. But it would be equally unfair to claim that, in the absence of non-question-begging counterexamples, we should *accept* the principles. This is because all of the examples so far adduced fail to support the incompatibilist principles *over* narrower principles acceptable to a compatibilist. We are looking at an argument which purports to *establish* incompatibilism; if all of the data which support the argument can be explained by principles weaker than those employed by the argument, then we don't yet have an entirely convincing argument for incompatibilism. But even if there is a gap in the argument, I have pointed out that a certain kind of compatibilist is committed to a distinction in our abilities which may seem implausible. It is a virtue of the argument that it brings out this commitment.

For the compatibilist, what do miracles — violations of natural laws — have to do with freedom? For a libertarian, a free act must be *ungoverned* by laws — freedom to do otherwise requires the absense of deterministic laws. Not so, for the compatibilist. And some might have thought that a compatibilist requires that agents be free to cause law-breaking events — to fly faster than a speeding bullet and bend steel with their bare hands. Again, not so. What's required, on the compatibilist picture, is the ability, in the relevant sense, to cause to be false a proposition entailed by the laws of nature. Whether this is incredible, or merely a consequence of compatibilism, is a difficult question which I can't resolve here.[24]

NOTES

[1]SEE, FOR EXAMPLE: CARL GINET, "MIGHT WE HAVE NO CHOICE?", IN K. LEHRER (ED.), *Freedom and Determinism* (Random House, New York), pp. 87-104; and "The Conditional Analysis of Freedom", in Peter Van Inwagen (ed.), *Time and Cause* (Reidel, Dordrecht), 1980, pp. 171-186; David Wiggins, "Towards a Reasonable Libertarianism", in T. Honderich (ed.), *Essays on Freedom of Action* (Routledge and Kegan Paul, Boston 1973), pp. 31-62; Peter Van Inwagen, "The Incompatibility of Free Will and Determinism", *Philosophical Studies* 27, 1975, pp. 185-199, and *An Essay on Free Will* (Clarendon Press, Oxford, 1983), pp. 55-105. Michael Slote discusses the "new" arguments for incompatibilism in: Michael Slote, "Selective Necessity and the Freewill Problem", *Journal of Philosophy*, 79 pp. 5-24. Also, for a recent discussion of Van Inwagen's argument, see: Terence Horgan, "Compatibilism and the Consequence Argument", *Philosophical Studies* 47, 1985, pp. 339-356.

[2]John Martin Fischer, "Incompatibilism", *Philosophical Studies* 43, 1983, pp. 127-137; and Van Inwagen, 1983, pp. 93-105.

[3]Fischer, 1983.

[4]Carl Ginet explicitly defends his argument against my criticism in: Carl Ginet, "In Defense of Incompatibilism", *Philosophical Studies* 44, 1983, pp. 391-400; Peter Van Inwagen defends the argument against a similar criticism in: Van Inwagen, 1983, pp. 101-104.

[5]In presenting the assumptions of the argument, I closely follow Ginet, 1983. However, in Ginet, 1983, the proposition-operator-form is taken to abbreviate:

"p and it was not in agent S's power at time t to make it not the case that p."

In "Incompatibilism", I argued that the phrase, "It was in S's power at t to make it the case that p", is crucially ambiguous. Ginet attributes to me the claim that there are the following two interpretations of the phrase (Ginet, 1983):

(iG) S had it in his power at t to cause it to be the case that p.

(iiG) S had it in his power at t to perform some act e^* such that if S had performed e^*, then it would have been the case that p.

I did indeed present what I called a "causal" and "non-causal" interpretation of the crucial phrase, but Ginet's two alternatives are not exactly what I suggested. Actually, I only specified the interpretations for propositions of a certain sort, "atomic" propositions which say that some event occurred at a time. Let us assume that the proposition p is the proposition that event e occurs at t_0. My two interpretations of "it was in S's power at t to make it the case that p", as presented in Fischer, 1983, are:

(iF) S had it in his power at t to cause e's occurrence at t_0 (i.e., to initiate at t a chain of events issuing in e's occurrence at t_0).

(iiF) S had it in his power at t to perform some act e^* such that if e^* were to occur, then e would have occurred at t_0.

Let us take as "given" that sometimes a proposition "mentions" or "contains" an event e, and that we can identify this event. If the proposition p "mentions" the event e (p, let us suppose, says that e occurs), then one might be able to cause some event e^* which is *sufficient* for p's obtaining (and thus be able to cause it to be the case that p) *without* being able to cause the occurrence of the *mentioned* event e. Thus, Ginet's (iG) does not capture the sense of (iF). Further, since it is plausible to suppose that whenever an agent performs an act e^* he causes some event e^{**}, then we could rewrite (iiG) as:

(iiG) S had it in his power at t to cause some event e^{**} such that if e^{**} had occurred, then it would have been the case that p.

Now it is unclear, for all that Ginet has said, that (iiG) differs at all from (iG); why isn't (iiG) simply an *explication* of (iG)? It is evident, I believe, that Ginet's (iG) is no less ambiguous than the phrase it purports to analyze; when one makes it the case that p, one causes it to be the case that p, but there are various different ways in which one might cause it to be the case that p. Thus, it will be most useful to interpret the phrase, "S made it the case that p", as "S caused it to be the case that p", and then distinguish different ways in which an agent can cause a proposition to obtain.

[6]Van Inwagen, 1983; p. 93. In Van Inwagen, 1983, Van Inwagen presents a family of three arguments for incompatibilism. In presenting the "non-modal argument (the "first formal argument"), he gives an account of the locution, "S can render p false" (p. 68). It is unclear exactly what the relationship is between this locution and the one used in the modal argument, "S has no choice about p". In any case, I have discussed Van Inwagen's account of "S can render p false" and the non-modal argument in: "Van Inwagen on Free Will", forthcoming, *The Philosophical Quarterly*. Perhaps Van Inwagen would not accept *any* of the accounts of "S can at t cause it to be the case that p" proposed below in the text of this paper; still, Van Inwagen accepts the argument-structure as I have presented it, and he would owe us an explicit account of "S can render p false" (or the related locution, "S can at t cause it to be the case that p").

[7]This argument is a simplified version of the argument I attributed to Ginet in Fischer, 1983. It closely follows Van Inwagen's presentation in Van Inwagen, 1984, pp. 93-94.

[8]Unhappily, the two interpretations I offered in "Incompatibilism", (iF) and (iiF), apply only when p is of the "atomic" form, "e occurs at t". I believed that these two sorts of interpretation could be straightforwardly extended to cases where p is non-atomic, but consideration of the basic argument for incompatibilism and Ginet's defense of it have made me realize that it is important explicitly to consider cases in which the proposition is non-atomic.

[9]It is clear that accounts (i) and (iii) of "S had it in his power at t to cause it to be the case that p" are only *partial*; they are only specified for atomic propositions and conjunctions with two conjuncts. It would be an illuminating project to produce a perfectly general account, but it is not necessary for my purposes here to do so. In this paper I shall

simply assume that accounts (i) and (iii) can be suitably generalized; note that neither Van Inwagen nor Ginet specifies the interpretation of the pertinent operator as applied to non-atomic propositions, but they simply rely on an intuitive understanding of the operators. I need only specify the differences between (i) and (iii) with respect to those kinds of propositions on which (i) and (iii) differ which are relevant to the debate about the incompatibilist's argument, while assuming that the generalized versions of (i) and (iii) will apply suitably to more complex propositions (including those non-atomic propositions which occur in the incompatibilist's argument but which aren't the focus of the current debate). An example of a non-atomic proposition which occurs in the Ginet/Van Inwagen argument which is not a conjunction of two conjuncts is:

$$(4) \quad N(L \supset (B \supset A))$$

Of course, if there is no clear intuitive content to the operator as applied to such propositions, then the Ginet/Van Inwagen argument doesn't even get off the ground.

[10]I discussed this case in: Fischer, 1983, pp. 131-132.

[11]Which principle one rejects will depend on one's view about counterfactuals; this point is discussed in the following section.

[12]For a defense of the view that the "backtracker" is true, see Jonathan Bennett, "Counterfactuals and Temporal Direction", *Philosophical Review*, 73, 1984, pp. 57-91. In contrast, see David Lewis, "Counterfactual Dependence and Time's Arrow", *Noûs*, 13, 1979, pp. 455-476.

[13]Ginet explicitly endorses this kind of argument at Ginet, 1983, p. 395. Here, of course, I am assuming that Ginet's (b-i) is the same as my (b-i); if not, I'm not sure what the interpretation of his (b-i) is. Van Inwagen endorses a structurally similar argument on behalf of his "principle β" at Van Inwagen, 1983, pp. 101-104.

[14]Ginet, 1984; for similar examples, see Van Inwagen, 1983, p. 98.

[15]Van Inwagen, 1975, p. 193.

[16]Van Inwagen, 1983, p. 62.

[17]Peter Van Inwagen, "Laws and Counterfactuals", *Noûs*, 13, 1979, p. 446.

[18]I am indebted for this point to David Lewis, "Are We Free to Break the Laws?" *Theoria*, 47, 1981, pp. 113-121. Lewis here defends compatibilism against Van Inwagen's non-modal version of the basic argument for incompatibilism.

[19]Again, I am indebted to Lewis, 1981.

[20]See Ginet, 1983, p. 398. I am also indebted to extremely useful personal correspondence with Ginet.

[21]I pointed out above that, whereas A can't at t_2 cause the compound event, "Condition C obtains at t_1 and A's back is scratched at t_2", A can at t_2 cause it to be false that Condition C obtains at t_1 and A refrains from scratching his back at t_2. So, on the compatibilist's picture, A *can't* perform an action which would be or cause a law-breaking event, but A *can* at t_2 cause an event, "A's back is scratched at t_2", which would be a *part of* a compound event, "Condition C obtained at t_1 and A's back is scratched at t_2", which *would be* a law-breaking event. On this picture, one can't cause law-breaking events, but one can cause parts of law-breaking events. Whereas the incompatibilist's examples do not, strictly speaking, establish his point, the incompatibilist's challenge to the compatibilist is to *justify* the distinction in our abilities to which he is committed. This project is not undertaken by Lewis.

[22]Van Inwagen, 1979, p. 446; and Van Inwagen, 1983, p. 62.

[23]In my original paper, "Incompatibilism", I suggested that the compatibilist would accept (c) on interpretations such as (i) and (iii), but deny it on (ii). Ginet's criticism of this position in Ginet, 1983, is convincing, and my argument in this section has obviously been different.

[24]I have benefitted from conversations with Phillip Bricker and Anthony Brueckner. The paper has improved as a result of comments by the reader for *Noûs*, and I am very grateful for the generous and useful comments by Carl Ginet on a previous version of this paper. My work on this paper has been supported by a Fellowship for Independent Study and Research from the National Endowment for the Humanities.