

## A New Incompatibilism

N. M. L. NATHAN

### I

In an early and famous essay on freedom A. J. Ayer attributed our fear of determinism partly to the half-conscious survival of an animistic conception of causality, in which an unhappy effect tries 'vainly to escape from the clutches of an overmastering cause'.<sup>1</sup> The general argument was this. (i) No determinism is incompatible with freedom unless it is a thesis not just about regularities in the world but also about regularities which in some sense must obtain. But (ii) the 'must' here will be unintelligible or obviously inappropriate unless it is definable in terms of a limitation on human abilities. And (iii) even if some version of determinism is reasonable to believe, there is no reason to believe in a determinism defined in terms of such a limitation. Subsequent speculations about natural necessity have not I think done anything to undermine this argument. But a version of 'incompatibilism' has recently emerged in which the necessity of determinism is for once quite openly defined in terms of a limitation on human abilities. Instead of claiming that unfreedom is the indirect and unobvious consequence of an independently intelligible natural or causal necessitation, the new doctrine appeals quite openly to the underived 'conceptual truth' that if anyone has it in his power to render some proposition false, then that proposition is not a law of nature, or at least not a law of physics. In this paper I try to apply Ayer's deflationary strategy to the new voluntaristic incompatibilism, and consider whether the doctrine is trivialised by the absence of any reason to believe in the particular brand of determinism whose freedom-excluding consequences it so successfully brings out. In section II I expound the new doctrine (VI from now onwards) and compare it with a more traditional incompatibilism. In the remaining sections I criticise the only seeming reasons I can think of for believing in determinism as VI defines it.

<sup>1</sup> A. J. Ayer, 'Freedom and Necessity', *Philosophical Essays* (London, 1959), p. 283.

## II

Here then to begin with is an argument for VI. Suppose you do  $x$  at time  $t_1$  and this action is determined in the sense that the proposition that you do it then is entailed by the conjunction of (1) a true proposition describing the whole state of the universe at some previous time  $t_{-1}$ , and (2) a set of true generalisations none of which is true only because somebody does not exercise his ability to do what would make it false. It then follows that your action lacks the freedom it would have if, at a time  $t_0$ , after  $t_{-1}$  but before  $t_1$ , you were able either to do  $x$  at  $t_1$  or not to do  $x$  at  $t_1$ . For the proposition that you do  $x$  at  $t_1$  is entailed by the conjunction of (1) and (2). So you are able at  $t_0$  not to do  $x$  at  $t_1$  only if either (a) you are able at  $t_0$  to render false the proposition mentioned in (1), or (b) you are able at  $t_0$  to render false one of the generalisations mentioned in (2). But (a) is excluded because nobody can alter what has already happened, and (b) is excluded by definition. If this is a sound argument, then it shows that if a certain form of determinism is true, nobody ever enjoys two-way powers of action. The determinism in question is roughly this:

- (D1) For any event there is a set of true generalisations whose conjunction with a proposition describing the whole state of the universe well before that event entails that the event occurs when it does, and no member of which is true only because somebody does not exercise his ability to make it false.

There is an excellent statement and defence of an essentially similar argument in Peter van Inwagen's paper, 'The Incompatibility of Free-Will and Determinism'.<sup>1</sup> Van Inwagen's version differs mainly in these respects. (i) He operates not with a set of true generalisations none of which is true only because somebody does not exercise his power to do what would make it false, but rather with the conjunction into a single proposition of all laws of physics, taking it as a 'conceptual truth' that if anyone has it in his power to render some proposition false, then that proposition is not a law of physics.<sup>2</sup> (ii) His argument applies, not to all actions,

1 'The Incompatibility of Free-Will and Determinism', *Philosophical Studies*, 27 (1975), pp. 185-199 (reprinted in G. Watson (ed.), *Free-Will* (Oxford Readings in Philosophy 1982). Cf. James W. Lamb 'On a Proof of Incompatibilism', *Philosophical Review*, 86 (1977), pp. 20-35.

2 *Ibid.*, p. 193.

but only to actions which involve physical movements. (iii) The time-references are either different or less explicit. He conjoins the laws of physics with a proposition expressing the state of the universe at a time prior to the agent's birth, rather than simply with a proposition expressing the state of the universe at some time before the time at which the agent is supposed to have been able to do what he did not do. And he does not distinguish between being able to do something at a particular time, and being able at a particular time to do something at a particular time.

An earlier paper by Carl Ginet contains another essentially similar argument.<sup>1</sup> He begins with two definitions. 'If "A" and "B" are descriptions specifying kinds of events then A *contingently necessitates* B if and only if

- (1) "A" does not entail "B",
- (2) every instance of A is accompanied by an instance of B (in a manner indicated in the descriptions of "A" and "B"), and
- (3) no one ever has a choice as to whether or not an instance of A shall be accompanied by an instance of B.

... a *complete first-level description* of a person's behaviour during a certain period ... is a description that contains only the specification of the places occupied by all the externally observable parts of the person's body throughout the period, relative to each other and to the adjacent environment, as well as the specification of all the sounds emitted by his body during that same period' (ibid., p. 89). With the aid of these definitions, Ginet states the following hypothesis. '(H) Every temporal segment of every human being's behaviour has a complete first-level description, 'B', and a series of antecedent sets of circumstances having the descriptions "A<sub>1</sub>", "A<sub>2</sub>", ... "A<sub>n</sub>", such that

- (1) "A<sub>1</sub>" does not entail "B";
- (2) A<sub>1</sub> contingently necessitates A<sub>2</sub>, A<sub>2</sub> contingently necessitates A<sub>3</sub>, ... , A<sub>n-1</sub> contingently necessitates A<sub>n</sub>, A<sub>n</sub> contingently necessitates B; and
- (3) the human being in question clearly had no choice as to whether or not the antecedent instance of A<sub>1</sub> would occur' (ibid., p. 90).

He then shows that (H) entails that no one has any choice as to what

<sup>1</sup> 'Might We Have No Choice?' in K. Lehrer (ed.), *Freedom and Determinism* (New York, 1966), pp. 87-104.

first-level descriptions his behaviour satisfies, and hence that no-one has any choice as to what descriptions of any kind his behaviour satisfies. If we assume that the instance of  $A_1$  is an event which takes place before the person's birth, then the main difference between Ginet's argument and van Inwagen's argument is that Ginet cautiously refrains from calling (H) 'determinism' and from suggesting anything about the relation between contingent necessitation and what people have in mind when they talk of laws of nature or laws of physics.

VI is true just because it employs an openly anthropomorphic concept of natural necessity in its definition of determinism. Traditional incompatibilisms shrink from this step, and operate with an undefined notion of natural necessity. David Wiggins has recently defended such a traditional doctrine.<sup>1</sup> He offers an argument for incompatibilism which is, or at any rate could quite easily be made, formally similar to the initial argument I gave for VI. But he does not commit himself to saying exactly how the necessity of the relevant laws or generalisations is to be understood, and gestures merely at an ostensibly self-sustaining circle of modal terms. It will help us to appreciate VI if I can show how fatal the consequences of this reticence are for Wiggins's argument. His reasoning is this. Let R say that N moved his finger at  $t_1$ , and A say that he kept it still at  $t_1$ , and ' $\Box t_1$ ' be read as 'it is historically inevitable at time  $t_1$ '. Now suppose we have a law of nature from which it follows that

- (1)  $\Box t_1$  (if C at t then R at  $t_1$ )

We then have

- (2)  $\Box t_1$  (if R at  $t_1$  then not A at  $t_1$ )

and if ' $\Box t_1$  (if ... then —)' is transitive, it follows that

- (3)  $\Box t_1$  (if C at t then not A at  $t_1$ ).

Now suppose that

- (4) N could have kept still at  $t_1$

However this is to be analysed, it at least entails

- (5)  $\Diamond t_1$  (A at  $t_1$ )

<sup>1</sup> David Wiggins, 'Towards a Reasonable Libertarianism' in Ted Honderich (ed.), *Essays on Freedom of Action* (London, 1973), pp. 31–62.

But it seems necessarily true that

$$(6) \quad \Diamond t_1 (A \text{ at } t_1) \supset (\Box t_1 ((A \text{ at } t_1) \supset (\text{not } C \text{ at } t)) \supset \Diamond t_1 (\text{not } C \text{ at } t))$$

Now (5) is the antecedent of (6). So we can detach

$$(7) \quad (\Box t_1 ((A \text{ at } t_1) \supset (\text{not } C \text{ at } t)) \supset \Diamond t_1 (\text{not } C \text{ at } t))$$

The antecedent of (7) is the contraposition of (3). So we have

$$(8) \quad \Diamond t_1 (\text{not } C \text{ at } t)$$

But (8) is obviously absurd, since  $t$  is before  $t_1$ . So if (1) and (2) are true, we will have to reject (5) and hence (4).

Things would be simplified if Wiggins were willing to replace (6), 'derived from a principle of Diodorus Cronus', by

$$(6a) \quad (\Box t_1 (C \text{ at } t) \ \& \ (3)) \supset \Box t_1 (\text{not } A \text{ at } t_1)$$

He could then just argue that since  $t$  is before  $t_1$

$$(9) \quad C \text{ at } t \supset \Box t_1 (C \text{ at } t)$$

So if  $C$  at  $t$  then by (6a)

$$(10) \quad \Box t_1 (\text{not } A \text{ at } t_1)$$

which is incompatible with (5) and hence (4). Thus simplified, the argument would have some formal similarity to the initial argument for VI. For that relied on the principle that if one is now powerless to make false a proposition about the past, and now powerless to make false a generalisation whose conjunction with that proposition entails a future event, then one is now powerless to prevent that future event.

The resemblance would be closer, and the argument improved, if a time were assigned to the ability which (4) refers to, and corresponding changes made to the time-references elsewhere. If, as Wiggins supposes,  $N$  did move his finger at  $t_1$ , then we can infer from this alone, and without the aid of premisses about laws of nature, that  $N$  did not have the ability at  $t_1$  to keep still at  $t_1$ . So (4) should presumably be replaced by

$$(4a) \quad N \text{ could, at } t, \text{ have kept still at } t_{+n}$$

and (5) by

$$(5a) \quad \Diamond t (A \text{ at } t_{+n}).$$

(1) then becomes

$$(1a) \Box t \text{ (if } C \text{ at } t_{-n} \text{ then } R \text{ at } t_{+n});$$

and (6a) becomes

$$(6b) (\Box t (C \text{ at } t_{-n}) \& \Box t \text{ (if } C \text{ at } t_{-n} \text{ then not } A \text{ at } t_{+n})) \supset \Box t \text{ (not } A \text{ at } t_{+n})$$

From 'C at  $t_{-n}$ ' we derive ' $\Box t (C \text{ at } t_{-n})$ ' and hence, by (6b),

$$(10a) \Box t \text{ (not } A \text{ at } t_{+n})$$

What does Wiggins understand by 'historically inevitable'? He says 'By "it is historically inevitable at time  $t$  that  $p$ " is intended something like this—whatever anybody or anything does at  $t$  or thereafter it *can* make no difference to  $p$ ,  $p$  being either a law of logic or a law of nature or already history or being the logical or physical consequence of what is already history'. And he adds ' $p$  is historically necessary at  $t$  if, and only if,  $p$  is true whatever may happen at  $t$  or thereafter (consistently with the laws of nature)'.<sup>1</sup> There is no attempt here to explain what it is about laws of nature that allows us to suppose that it follows from a law of nature that

$$(1) \Box t_1 \text{ (if } C \text{ at } t \text{ then } R \text{ at } t_1)$$

and in particular no suggestion that that it is part of the definition of a law of nature that it is a true generalisation which people are unable to make false in precisely that sense of 'ability' which is exemplified in (4). Referring to the 'can' which I italicised, Wiggins remarks that his definition of historical inevitability 'already includes the notion of possibility'. His only aim, he assures us, is to fix 'from within the circle of modal notions a sense of necessity which satisfies (6) and yields a strict implication which is transitive (for the passage from (1) and (2) to (3))' (ibid).

The advantage of VI, as compared with Wiggins's modally non-committal doctrine, is that we can at any rate be reasonably sure that ' $\Box t_1$ ' stands for something meaningful, and something moreover which makes it absolutely plain that (4a) entails (5a). One might perhaps try to defend Wiggins's neutrality by means of a dilemma Either (4a) is meaningful or it is not. If it is not, then freedom is an illusion whether or not any form of determinism is true, and it is pointless to try to beat off incompatibilist attacks on freedom by

<sup>1</sup> Ibid., p. 45 (my italics).

showing that the incompatibilist is talking about a determinism which there is no reason to believe. If on the other hand (4a) is meaningful, then on any plausible conception of physical or historical possibility it obviously entails that it is physically or historically possible at  $t$  for  $N$  to have kept still at  $t_{+n}$ . But the dilemma fails. If (4a) is meaningful, and (5a) is actually equivalent to (4a), then of course (4a) entails (5a). But that is just what Wiggins does not want to commit himself to, with respect to the meaning of 'physical or historical possibility'. He wants to say that (4a) obviously entails (5a), even if (5a) does not entail (4a). And the difficulty is that we do not know that there is an intelligible sense of 'physical or historical possibility' on which (4a) and (5a) are thus related. It is of course easy enough to define an intelligible and non-voluntaristic sense of 'physical or historical possibility' if we are prepared to say that every concept of laws of nature, voluntaristic or otherwise, has correlative concepts of physical necessity and possibility. 'It is physically or historically possible at  $t$  for  $N$  to do  $x$  at  $t_{+n}$ ' will then mean whatever we mean by, 'There is no law of nature whose conjunction with a description of how things are at  $t$  entails "N does not do  $x$  at  $t_{+n}$ ".' And if we take laws of nature in a Humean way, as for example just true spatio-temporally unrestricted generalisations, or as true spatio-temporally unrestricted generalisations for which there is good inductive evidence, then the correlative concept of physical possibility is both non-voluntaristic and perfectly intelligible. But I do not think that this move will really help. Suppose we grant that, on a Humean concept of physical possibility, (4a) and (5a) are non-equivalent. For Wiggins's argument to work, it must also be true that (4a) entails (5a). But that will have paradoxical consequences. Suppose laws of nature are taken just as contingently true spatio-temporally unrestricted generalisations. Then

(5a) It was historically or physically possible at  $t$  for  $N$  to keep still at  $t_{+n}$

will be equivalent to

(11) It is false that

(12) Whenever someone like  $N$  is in circumstances like those obtaining at  $t$ , he will not do something like keeping still at  $t_{+n}$ .

But (11) entails

(13) Once at least someone like N in circumstances like those obtaining at  $t$  will do something like keeping still at  $t_{+n}$

So if (4a) entails (5a), then N was able at  $t$  to keep still at  $t_{+n}$  only if the type of ability he would have exercised if he had kept still at  $t_{+n}$  is sometimes exercised. This seems counter-intuitive. Nor are matters improved if laws of nature are taken as contingently true spatio-temporally unrestricted generalisations for which there is good inductive evidence. (5a) will then mean 'it is false that both (12) and

(14) There is good inductive evidence for (12)'

So if (4a) entails (5a), then, paradoxically, N was able at  $t$  to keep still at  $t_{+n}$  only if either the type of ability he would have exercised if he had kept still at  $t_{+n}$  is sometimes exercised or there is no good inductive evidence that it never is.

### III

VI, though true, is trivially true unless there is some good reason to believe in the voluntaristically defined determinism whose freedom-excluding consequences it brings out. In searching for such a reason, I turn first to the voluntaristic incompatibilists themselves.

Van Inwagen says quite specifically that he will not try to establish the truth or falsity of determinism as he defines it (*ibid.*, p. 185). But Ginet has a good deal to say about the truth or falsity of the voluntaristically defined hypothesis (H) from which, as he argues, it follows that we have no choice. Roughly put, (H) was that all human behaviour is contingently necessitated by some event as to the occurrence of which the agent obviously had no choice, where A contingently necessitates B only if nobody has a choice as to whether or not an instance of A shall be accompanied by an instance of B. By 'nobody has a choice as to whether, etc.' is meant not that nobody chooses or wills that an A shall or shall not be accompanied by a B, but that no such choice determines whether or not a B actually occurs (*ibid.*, pp. 92–93). According to Ginet (H) is not only logically capable of confirmation, but also not now known to be false. Although Ginet does not think that (H) either is, or is likely to become well-confirmed, his claims about the conceivable future

observations which would make it reasonable to believe (H) might seem at least to make the hypothesis plausible. 'As a first step in imagining such future observations,' he writes, 'we need to look at some cases where connections of contingent necessitation are already reasonably believed to hold, and to consider what sorts of observations these beliefs are grounded in. Such beliefs are numerous. Some are stated by the following conditionals: 'If a piece of copper surrounded by ordinary atmospheric conditions becomes hotter, then it must expand. If a small material body, such as a bowling ball, is left near but not in contact with a very large material body, such as the earth, without anything between them but air, then the small body must move into contact with the large one. If a baseball is hurled with the maximum force that a normal man can muster, in such a way that, after travelling 10 feet from his hand, it strikes at right angles the centre region of a piece of glass of just the same sort, and mounted in just the same way, as the glass inside my childhood home, then the glass must break'. We are justified in these beliefs, because 'quite a number of instances of the antecedent have been observed to be accompanied by an instance of the consequent, and . . . instances of the antecedent have been observed or produced under a variety of circumstances, including any with regard to which there seemed a choice that in them the consequent might not accompany the antecedent. In short, we have never been able to observe the antecedent unaccompanied by the consequent, in spite of our making every effort to do so' (ibid., pp. 96-97). Similar observations 'could . . . conceivably be made with respect to any given sort of behaviour-sequent as consequent and some set of physiological and/or environmental circumstances as antecedent'. And if we accumulate such observations for a great many different kinds of behaviour-sequences, that would constitute support for (H).

The trouble is, however, that each of the conditionals which Ginet cites is strictly speaking false, and only seems plausible because we tacitly take it to have a *ceteris paribus* clause. The piece of copper will not expand in normal atmospheric conditions if somebody puts both ends in a vice, the small body will not move into contact with the large body if it is made of iron and somebody puts a vast magnet on the other side of it at the moment of release, and if striking a piece of glass means touching it as distinct from actually breaking it, the glass will not break if somebody deflects the baseball along the plane of the glass at the moment when it touches the glass. No doubt true conditionals can be found in which *ceteris*

*paribus* clauses covering human intervention are redundant. For example: if a body is accelerating, then it does not reach a speed greater than the speed of light. And it is conceivable that any given kind of behaviour sequent should figure in such a conditional. But it is not plausible to suppose this, in the way that it would be plausible if any old rough-and-ready generalisation of the sort cited by Ginet were true even without a *ceteris paribus* clause covering human intervention.

#### IV

VI deduces our unfreedom from

(D<sub>1</sub>) For any event there is a set of true generalisations whose conjunction with a proposition describing the whole state of the universe well before that event entails that the event occurs when it does, and no member of which is true only because somebody does not exercise his ability to make it false.

And (D<sub>1</sub>) needs to be distinguished from the initially more plausible

(D<sub>2</sub>) For every event there is a set of contingently true generalisations whose conjunction with a proposition describing the whole state of the universe well before that event entails that the event occurs when it does.

Suppose we grant (D<sub>2</sub>). Might we then not just as well grant the hypothesis (D<sub>3</sub>) which we would get by substituting 'laws of nature' for 'contingently true generalisations' in (D<sub>2</sub>)? And then, if 'law of nature' is taken in any ordinary sense, will not (D<sub>3</sub>) actually entail (D<sub>1</sub>)?

There is perhaps a hint of support for this kind of argument in van Inwagen's main paper. He defines determinism as the conjunction of '(a) For every instant of time, there is a proposition which expresses the state of the world at that instant [and] (b) If A and B are any propositions that express the state of the world at some instants, then the conjunction of A with the laws of physics entails B' (ibid., p. 186). And although, as I said, he claims to be unconcerned with the truth or falsity of this determinism, van Inwagen does insist that it is 'a conceptual truth' that if anyone has it in his power to render some proposition false, that proposition is

not a law of physics. But why draw attention to this supposed fact about our ordinary concept of laws of physics, unless it would tend to make a determinism which employs our ordinary concept more plausible than a determinism which employed some new non-voluntaristic concept of laws of physics?

Van Inwagen distinguishes between laws of physics and laws of nature. He thinks it is an obvious conceptual truth that nobody has the power to render a law of physics false, but obviously not a conceptual truth that nobody has the power to render false a law of nature. There is in particular nothing at all odd about saying that psychological laws hold only because people do not exercise their power to make them false.<sup>1</sup> Anyone who thought that (D<sub>3</sub>) entailed (D<sub>1</sub>) when 'law of nature' was taken in any ordinary sense would have to reject van Inwagen's distinction. And there are various arguments for rejecting it. Is not 'true proposition describing natural necessities' part of what we ordinarily mean to convey by 'law of nature'? If so, how is it possible to assign a sense to 'It is naturally necessary that all A-type events are followed by B-type events' without taking it to mean that if an A-type event has occurred nobody has the ability to prevent its being followed by a B-type event? Or, if the modal content is not so immediately present, perhaps it can be extracted from the fact that someone who says that a proposition is a law of nature will ordinarily be willing to put forward a corresponding counterfactual. Is not this readiness most easily explained by saying that he is using 'law of nature' in a sense on which laws entail counterfactuals? But as Kneale points out, 'A contingent universal proposition can always be expressed in the form "There are in fact no  $\alpha$ -things which are not  $\beta$ ", and from such a proposition it is impossible to deduce that if something which was not in fact  $\alpha$  had been  $\alpha$  it would also have been  $\beta$ .'<sup>2</sup> So are we not obliged to say that laws of nature, taken in the normal counterfactual-entailing way, are necessary universal propositions? If so, then once again, 'necessary', if it were to have any clear sense, might have been interpreted in terms of human impotence. And there is another idea of Kneale's which someone might try to use against van Inwagen's restrictions. 'It is part of our ordinary idea of a law of nature that it is logically possible for there to be unrealised empirical possibilities in the sense of false propositions consistent

1 Peter van Inwagen, 'Reply to Narveson', *Philosophical Studies*, 32 (1977), pp. 91-92.

2 W. Kneale, *Probability and Induction* (Oxford, 1948), p. 75.

with all laws of nature, and of the form 'something is an F', where F is a spatio-temporally unrestricted empirical predicate. But 'Something is an F' cannot be both false and compatible with all laws of nature if 'Nothing is an F' is a law of nature.' What could prevent 'Nothing is an F' from being a law of nature, and thus allow for the logical possibility of unrealised empirical possibilities? If, as Kneale in effect suggests,<sup>1</sup> the only answer is that 'Nothing is an F' does not describe a natural necessity, then we could once again ask what 'necessity' could mean here if not interpreted in terms of human impotence.

I think it is unnecessary to assess these complex claims about the content of our ordinary concept of natural law. Even if it is true that (D<sub>3</sub>) entails (D<sub>1</sub>) when 'law of nature' is taken in any ordinary sense, it still remains to be shown that (D<sub>3</sub>) is as plausible as (D<sub>2</sub>) when 'law of nature' is taken in such a way. If the ordinary concept of a law of nature is voluntaristic, and used in (D<sub>3</sub>), then (D<sub>3</sub>) is stronger than (D<sub>2</sub>), and there can be no automatic equation of their relative plausibility. From the fact that a concept is ordinary, nothing follows about whether or how often it actually applies to the world. And even if the generalisations which we call laws of nature would not in reality be different in any other general way from the generalisations we call accidental, unless they had those features which we ordinarily suppose them to have when we call them laws, it still does not follow that they really have those ostensible differentiating features.<sup>2</sup>

## V

Finally, let us suppose again that

- (D<sub>2</sub>) For every event there is a set of contingently true generalisations whose conjunction with a proposition describing the whole state of the universe well before that event entails that the event occurs when it does.

If (D<sub>2</sub>) is true and the voluntaristic (D<sub>1</sub>) false, then at least one of the true generalisations which (D<sub>2</sub>) refers to will be true only

- 1 W. Kneale 'Natural Laws and Contrary-to-fact conditionals', *Analysis* 10 (1950), pp. 123-124. For a fuller statement see George Molnar, 'Kneale's Argument Revisited', *Philosophical Review*, 78 (1969), pp. 79-89.
- 2 For a defence of the 'heroic Humean' position that there is no real distinction see Tom L. Beauchamp and Alexander Rosenberg, *Hume and the Problem of Causation* (New York, 1981), ch. 4.

because people do not exercise their ability to make it false. But it is not utterly unlikely that any generalisation with that sort of contingency will be true at all, or at least utterly unlikely that it will be true if it has a vast number of actual instances? For then the chances are that on at least one occasion the falsificatory ability would be exercised. And even if, despite the existence of unexercised abilities to make them false, these generalisations were true, would not the existence of the falsificatory abilities at least prevent us from being reasonable in believing the generalisations? For we could be reasonable in believing them only if we had tried to falsify them and failed, and their falsification would be all too easy if there were a widespread ability to make them false. So even if (D<sub>2</sub>) were true and (D<sub>1</sub>) false, we would not be reasonable in believing (D<sub>2</sub>) if we were not also reasonable in believing (D<sub>1</sub>).

Let me take the second of these two arguments first. What makes for reasonable belief in (D<sub>2</sub>)? Is it enough to be reasonable in believing that for every event there is some set or other of contingently true generalisations of the kind referred to by (D<sub>2</sub>), and not necessary to be reasonable in believing any particular set of contingently true generalisations of the kind referred to by (D<sub>2</sub>)? If I have the ability to make false a particular set of generalisations S whose conjunction with a proposition of type T entails something about an event e, that may prevent me from being reasonable in believing S. But it does not necessarily prevent me from being reasonable in believing that there is some set or other of true generalisations whose conjunction with a proposition of type T entails the same thing about e. It is of course true that we do want to be reasonable in believing particular sets of generalisations of the kind referred to by (D<sub>2</sub>). We want this because we want to be able to predict the future in detail. And it is not easy to see how someone can have first-hand reasonable belief in any such particular set of generalisations, based on the failure of his own best efforts to falsify the generalisations, if he himself has the ability to make them false. But that is a further problem, distinct from the problem of whether there is any general freedom-excluding determinism which it is reasonable to believe, and I shall not try to discuss it here.

This leaves us with the argument that (D<sub>2</sub>) is plausible and (D<sub>1</sub>)'s falsity would make (D<sub>2</sub>) very unlikely to be true. I find this argument very difficult to assess. But I suggest that if it is successful then it is too successful. It is a reply to the suggestion that voluntaristic incompatibilism is trivially true. But it relies on a

principle which, if valid, would make every form of incompatibilism redundant.

Let me begin with a fuller statement of the argument itself. It exploits an Aristotelian thought. Aristotle held that whatever is always possible is at some time actual. The underlying idea, I assume, is that the more often there is a finite chance of a kind of thing's occurring, the greater the chance that it will occur on at least one occasion.<sup>1</sup> In the extreme case, when there is always or infinitely often a chance of a kind of thing's occurring, the chances of its occurring at least once are as high as you like. Suppose then that we have a true generalisation of the form 'All A's are B's', and this generalisation has a huge number of instances. According to the Aristotelian thought, if each of this huge number of A's were possibly not a B, then the chances that one of them actually is not a B are very high indeed. So, if the generalisation is true, and has a huge number of instances, it follows that probably not all of the A's are possibly not B's, i.e. that probably at least one A must be a B. Aristotle's thought, then, if there is anything in it, gives if the following general principle:

(R) Probably if all of a huge number of A's are B's, at least one A must be a B.

The modal terms in the argument are of course obscure. But let us interpret them in terms of human powers, so that 'this A must be a B' means 'nobody has the power to make it false that this A is a B'. Now if Aristotle's thought gives us (R), thus interpreted, it presumably also gives us

(R1) Probably if all of a huge number of A's are B's then quite a lot of A's must be B's

with 'must' interpreted in the same way. Suppose it does, and suppose also that every event falls under some true generalisation which links the type of circumstance in which it occurs with the type of event it is, and suppose also that all these generalisations have huge numbers of instances. It would then follow that probably, on quite a lot of occasions, people do not have the power not to do what they do. Or, putting it another way, that probably at least quite a lot of propositions of the form 'This was an A and this was a B' which instantiate the generalisations, are not true only

<sup>1</sup> For a discussion of the Aristotelian texts see C. J. F. Williams, 'Aristotle and Corruptibility', *Religious Studies*, 1 (1966), pp. 95-107, 203-215.

because people do not exercise their power to make them false. We would not in this case have found any reason to believe that if any determinism is true, a determinism is true which excludes all freedom. The most we would have found is that if any determinism is true, and if it is also true that all the regularities which make that determinism true have huge numbers of instances, then probably quite a lot of actions are not free. But let us pursue the argument a little: its proper conclusion does bear some resemblance to the thesis that if (D<sub>2</sub>) is true, so is (D<sub>1</sub>).

How can we hope to show that all the regularities which make (D<sub>2</sub>) true have huge numbers of instances? It would be a sufficient threat to freedom if all the regularities under which actions fall have huge numbers of instances. What grounds are there for thinking that even that much is true? Consider the hypothesis that all actions fall under just one true generalisation linking the type of circumstance in which they occur with the type of actions they are. If there is just one such true generalisation, it will have as many instances as there are actions. Suppose that an action *x* is performed whenever and only when the agent previously prefers some state of affairs to exist rather than not, and believes that *x* is sufficiently likely to lead to this state of affairs to outweigh any distaste he feels for anything else that *x* might lead to. Then instead of having to say that all actions belong to some type or other, and all circumstances of their occurrence belong to some type or other, such that whenever there are circumstances of that type there is an action of that type, we can say that all actions belong to a *single* type, and all circumstances of their performance belong to a *single* type, such that whenever there are circumstances of that type there is an action of that type. The type of circumstance can be specified by saying that it is one in which the agent has this type of preference and belief, the type of action by saying that it is one with respect to which the agent has this type of preference and belief. If this simple generalisation about a type of action and a type of circumstance is true, then it has a huge number of instances because a huge number of actions have been and will be performed. So if (R<sub>1</sub>) is also true, then there are quite a lot of propositions describing particular pairs of circumstances and actions which are not true only because the agent did not exercise his power to make them false.

This gives us a crude model of the argument. But it is of course wildly implausible to suppose that an action is performed whenever and only when the agent has the preferences and beliefs I specified.

Even if actions are performed only when the agent previously has these types of preferences and beliefs, they are surely not performed whenever the agent previously has these types of preferences and beliefs. Let us then alter our generalisation about actions and consider instead the hypothesis that (i) usually at least when someone prefers S to exist rather than not, and believes that doing x is sufficiently likely to lead to S to outweigh any distaste he feels for anything else his doing x might lead to, he does x; and (ii) in all those cases in which preference and belief are followed by the appropriate or 'rational' action some single type of event E occurs prior to the action. We then have the generalisation, 'All E's are followed by "rational" actions.' If this generalisation were true, and (R<sub>1</sub>) were also true, then there would be quite a lot of propositions describing particular pairs of E's and 'rational' actions which were not true only because the agent did not exercise his power to make them false.

Everything turns, however, on whether (R<sub>1</sub>) is true, and it seems to me that if (R<sub>1</sub>) really is true, then we can construct a simpler anti-freedom argument, with a stronger conclusion, which does not require us to suppose that human actions fall under regularities at all. To put it in another way, if (R<sub>1</sub>) could be used to vindicate a determinism which probably excludes some of the freedom which is excluded by the determinism of voluntaristic incompatibilism, there would be a surer way of excluding even more of this freedom without vindicating any thesis about regularities which we are powerless to alter. Which is to say: it is not necessary to defend any incompatibilism against the charge that its determinism is groundless, in order to threaten the freedom which that incompatibilism refers to. Here is the argument I have in mind. (R<sub>1</sub>) is not true unless it is also true that

(R) Probably if all of a huge number of A's are B's, at least one A must be a B.

So suppose (R) is true, and suppose I have the ability, in one and the same set of circumstances, both to do something and not to do it. If I have this ability with respect to an action, then there is some period of time, however short, prior to the action, at each of the infinitely many instants of which I have the ability to make the decision at the next instant, and also the ability not to make the decision at the next instant, and there is no relevant change in the circumstances over this period of time. But if (R) is given and its 'must' interpreted in

terms of human ability, it is highly unlikely that this should ever be the case. If I have an ability to do x, on each of as many occasions as you like up to my actual performance of x, then why is my ability not exercised on at least one of these occasions?<sup>1</sup> If my having the ability on such an occasion amounted merely to its being the case that I would have done x, if circumstances had been different, then there would be no implausibility. But this fails to capture the idea that it is precisely in the circumstances as they are, on the occasion as it is, that I have the ability.

## VI

It is one thing to be unconvinced by every seeming reason one can think of for believing something, another to conclude that there is no good reason to believe it. So for all I have so far said, there may still be a good reason to believe in determinism as VI defines it. But we can at least not be alarmed by VI until some such reason is discovered. Can we indeed confidently hope that every form of incompatibilism is either false or trivially true? Not quite, if 'incompatibilism' is taken broadly, to mean any doctrine according to which some species of determinism entails that we are less free than we normally think. For there is some plausibility in a voluntaristic version of the Aristotelian principle (R1), and if this principle is true, then we should perhaps believe in a determinism about rational actions which would exclude some of the freedom excluded by the determinism of VI. But even so, we can at least confidently hope that every non-redundant incompatibilism is either false or trivially true. For as I argued at the end of the last section, if a voluntaristic (R1) really is true, then we have a cogent argument for total unfreedom whose premisses do not include determinism of any kind.

SCHOOL OF ECONOMIC AND SOCIAL STUDIES,  
UNIVERSITY OF EAST ANGLIA,  
NORWICH, NR4 7TJ.

<sup>1</sup> For an essentially similar argument see F. Brentano, *The Foundation and Construction of Ethics* (1925), tr. Elizabeth Schneewind (London, 1973), pp. 265-267.