Book Section:
Steward, H orcid.org/0000-0003-1654-577X (2021) What Is Determinism? Why We Should Ditch the Entailment Definition. In: Hausmann, M and Noller, J, (eds.) Free Will: Historical and Analytic Perspectives. Palgrave Macmillan , pp. 17-43. ISBN 978-3-030-61135-4

https://doi.org/10.1007/978-3-030-61136-1_2

What is Determinism? : Why we Should Ditch the Entailment Definition

Helen Steward

University of Leeds

What is the thesis of determinism? Though it is obvious that in principle there is more than one possible thesis that might be given this name, it seems to be the case that philosophers working on the free will problem have gradually gravitated towards a more-or-less standard definition, minor variations on which can now be found widely scattered through the free will literature. I call it the 'entailment definition' and it states, roughly, that determinism is the thesis that for any given time, a complete statement of the facts about that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time. In this paper, I argue that acceptance of the entailment definition has been a mistake – and that we need a definition of determinism which, by contrast with the entailment definition, makes explicit mention of the notion of natural necessity.

Keywords: determinism, free will, natural necessity, laws of nature

What is the thesis of determinism? Since there are many possible theses which might warrant the label 'determinism' (amongst them, for example, sociological, economic, genetic and psychological determinisms), let me put my question somewhat more precisely: what is the *general* thesis of determinism, the thesis with which philosophers who worry about the free will problem mostly concern themselves – and which perhaps more often than any of the other versions tends simply to be given the general name, 'determinism'?[1] J.L. Austin once claimed that 'determinism' was "a name of nothing clear" (1970, 231). But I think someone might reasonably wonder whether this remains the case today. On the whole, philosophers working on the free will question tend to operate with a definition of determinism the main components of which are now fairly standard, and that appears to be pretty well accepted on all hands. I call it the *entailment definition* and it will be helpful, in order to render later discussion clearer and easier, to spell out one representative version of it explicitly here:

(ED) For any given time, a complete statement of the [nonrelational] facts about that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time.[2]

Of course, there are concepts contained within ED which it has generally been realised need considerable unpacking and which raise myriad philosophical issues of their own. What, for instance, is a 'complete statement' of the facts about a time and is there any reason for supposing that such a thing could be formulated, even in principle? Which, exactly, are the 'nonrelational' facts, and do we know how to distinguish them properly from the relational ones? What is a law of nature and how should such a law be stated? These are important questions to which it is not clear that we have satisfactory answers; and indeed I suspect satisfactory answers to some of them are not to be had. But for present purposes, I intend simply to allow the possibility that there may be ways of answering them which will permit (ED) at least to be considered a reasonably clear and coherent doctrine. My main focus here will be on a different question – namely, assuming for the sake of argument that (ED) is not objectionable for reasons to do with the unclarity of some of its main concepts, is the overall *form* of (ED) correct? By which I mean: is it right to think that the claim to

which determinism centrally commits itself should be expressed as a thesis concerning the entailment of certain true statements by certain sets of others?

In this paper, I want to argue that the answer to this question is 'no' and that acceptance of the entailment definition of determinism has been a historically understandable, but now recognisable mistake. In a sense, of course, one is free to define technical terms as one will, and it might be argued that even for non-technical concepts, the only court of appeal for definitions is generally established usage. One might wonder, therefore, how an argument for the inappropriateness of a particular definition of a quasi-technical concept (one which, moreover, I have conceded already is certainly the market leader) can hope to get off the ground at all. What I shall try to show is that (ED) does not permit us to set up the free will problem in the best and most illuminating way. I shall offer two, closely related reasons for thinking this is the case: (i) (ED) is too weak to capture the deterministic thesis that troubles most incompatibilists; and (ii) (ED) does not sort compatibilists from incompatibilists in the way they ought really to be sorted, leaving many people who would account themselves *incompatibilists* on the wrong side of the compatibilist-incompatibilist divide, according to the letter of the definition. At the root of the problem, I shall suggest, is the fact that *entailment* is a logical concept; it is about what is *logically* compossible with what. Whereas determinism as it often figures in discussions of the free will problem – and more importantly still, in our thinking about it - is an inescapably *metaphysical* thesis whose import cannot be captured by (ED). The proper expression of determinism needs a metaphysical connection at its heart. Entailment will not do the job.

The paper falls into four sections. In the first, I ask the question what the thesis is that we are really worried about when we worry about free will, and offer what I call a *metaphysical* definition of determinism as a potential rival to (ED), an alternative I call (MD) (for 'metaphysical determinism', of course), as a means of stating the potentially worrisome thesis. In the second, I examine the relationship between (ED) and (MD), asking first whether (MD) entails (ED), to which my answer is a tentative 'probably'; and then whether (ED) entails (MD), to which it is a definite 'no'. (MD) is thus shown to be a stronger thesis than (ED). In the third section, I try to say something about how we have arrived historically at (ED), in order to encourage the thought that it is no more than a contingent product of the philosophical thinking of a very particular way of viewing the world, one from which philosophers today increasingly dissent, and so is something we need not feel is written in stone. Then, in the final section, I try to argue the case for thinking that it would be better, at any rate for the purposes of the free will debate, if ED were replaced by something much more like MD – and try to respond to what I think are likely to be the two main objections to my argument.

(i) *What is the thesis we are worried about when we worry about free will?*

Let me begin by asking: what is the thesis which worries us, when we are concerned that the world might not permit us free will because something called 'determinism' might be true? Here are some rough first shots at saying what seems worrying: it is the idea that things that have already happened (and which we cannot therefore alter) might *settle* or *fix* the future, or that that the universe is so constrained by governing laws that *nothing can happen* other than what in fact does happen, or that the future is causally or physically *necessitated* by what has gone before it.[3] As William James puts it metaphorically, "[determinism] … professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be"(James 1968, 40). Note that concepts of natural modality (such as physical possibility, causal necessity, and what are plausibly derivative relatives of these such as the idea that the future might be 'settled' or 'fixed' by

the past)[4] are central to these thoughts. If anything of the sort expressed by such claims were true, we might worry, there could be no such thing as free will, no such thing as the power to do otherwise than one in fact does. In order to have a handy label for this general conception of determinism, let us formulate the following definition (modelled on the same formal lines as (ED), for ease of comparison), and call it (MD), for 'metaphysical determinism':

(MD): For any given time, the total set of [non-relational] facts[5] about that time, causally or physically or naturally necessitates every fact about what happens after that time.

In (MD), note, it is the relationships of causal, physical or natural necessitation, rather than that of entailment, which have pride of place.

A few clarificatory comments are immediately in order. First, I use the term 'metaphysical determinism', rather than, say, 'causal determinism' or 'physical determinism', and offer a disjunctive elucidation of what is meant by it, in an attempt to formulate a definition which will serve to include quite a wide variety of related doctrines. There may be those, for example who do not suppose that everything is describable in physical terms, or who do not believe that the physical is fundamental, and so cannot be called *physical* deterministsl; but who nevertheless believe that everything that happens in the world is causally settled by past facts. And there may be those who are suspicious of the parochiality of the concept of 'cause', and so would not want to be accounted *causal* deterministsl, but nevertheless believe that determinism can be perfectly well expressed by means of the idea that the laws of nature, conceived of as metaphysical realities, are both all-encompassing and deterministic. I use the term 'metaphysical', therefore, in order to allow for views of both these kinds – and perhaps also for other possible variants of what I regard as the idea, centrally important to the formulation of determinism, that there are necessitating forces, powers or causes in the world that make it the case that certain things *must* follow upon certain others.[6] The crucial point is that metaphysical varieties of determinism, as I conceive of them, are doctrines according to which the determinative relation holds between *worldly states of affairs* and is owed to *a worldly variety of constraint*, though perhaps one of which we have very little real understanding. The crucial contrast class by means of which to understand what counts as a metaphysical conception of determinism, for my purposes, will be the *entailment* conception, which formulates determinism as essentially a matter of the holding of certain entailment relations, and does not explicitly mention any relationship of natural necessity.

A second point that deserves comment is the fact that there is, in (MD), no explicit mention of the laws of nature. We have grown very used to formulations of determinism which explicitly mention the laws, as well as past conditions. But this, I believe, may be a corollary, at least in part, of our penchant for definitions along the lines of (ED). If one is trying to formulate the thesis of determinism using entailment as one's central concept, it is arguable that one simply *must* mention the laws, on the grounds that at least some of the relevant statements about the future will fail, otherwise, to be entailed by the relevant sets of non-relational facts about previous times.[7] But if we are aiming at a *metaphysical* formulation of determinism, it is by no means clear that we must mention laws explicitly. The relation of natural necessitation itself, I suggest, when thought of as a metaphysical relation, might rather be thought of as something which can hold simply between past and future facts or states of affairs. The idea would be simply that past combinations of states of affairs are such as to draw future ones inevitably in their train. They may do so, of course, *according to* laws – but that is, I suggest, a different matter.

One thing which has perhaps made it hard to get clear about this, is the fact that the idea that laws are, as it were, the ontological necessitators of the natural world is a powerful and very

influential one[8] (though it has come under increasing pressure in recent years from alternative, more Aristotelian ideas).[9] There are those, therefore, who will be wont to insist that it is in fact the laws (in the form, for example, of necessary relations amongst universals) that provide the ontological *ground* for necessitation relations in general. I do not wish to take a view here on the question whether this idea about laws is in any sense correct; though it seems important to note that the view that the existence of laws is the inevitable corollary of belief in natural necessity is by no means universally held.[10] But I do want to argue that *even if* it were true that (something like) relations between universals provided the ontological ground for natural necessity, it would not follow that the determinist must include the relevant laws in her account of the facts which (according to her thesis) together naturally necessitate each given future one. Of course, if we accept the existence of laws of nature, it is going to matter to what happens, given any particular set of non-relational facts, what the relevant natural laws are – but even if the laws are contingent, it doesn't follow that those laws need to be *added* to the list of necessitating circumstances before we can say that a natural necessitation relation obtains between some set of facts holding at t1 and a fact or set of facts holding at t2. To say that these natural laws exist *implies* that the relevant relations of natural necessitation obtain simply amongst the particular facts in question; it is not to say instead that the relevant relations of necessitation obtain between the particular facts about a given time t1 *in combination with* the laws, on the one hand, and the particular facts about any given later time, t2, on the other. This is to suppose that everything that matters to the holding of a relation, R, must be specified when we specify its relata. It would be to make an analogous error to that that would be made by someone who thought that the rule of *modus ponens*, say, always needs adding to the *premises* of any argument which exploits that form if the argument is to be valid. Someone who thinks that the laws must be *added* to the particular facts before we can suppose that those particular facts necessitate certain future ones, indeed, invites the question whether *further* laws of nature are required which govern the necessitation relations between circumstances plus first-order laws, on the one hand, and further, future circumstances, on the other – and whether those further laws then *also* need to be added to the account of the total list of necessitators … and so on. The regress should be avoided (as is usually best with regresses) by refusing to take the first step and insisting that the natural necessitation relation itself always holds between particular facts or circumstances, even if we believe in laws.

Someone might say that (MD) is immediately problematic because we do not really know what natural necessitation is. I have some sympathy with this thought, but for the purposes of this paper, it will not be relevant to address the question whether we can, in the end, make sense of the idea of natural necessitation. If it turns out to be the case, in the end, that the concept of natural necessitation contains some unresolvable incoherence, I would then want to argue that no worrying conception of determinism of the secular and science-compatible kind I am attempting to characterise here will make sense either, and we will simply be home free on the free-will problem. It will be relevant for my purposes here, though, only to establish that we are not entitled to *assume* without argument when we set up our definition of determinism that natural necessitation is an illegitimate notion, something I shall aim to argue in section (iii). Moreover, there are things that can be said which might encourage us to think that we *do* understand the idea of natural necessitation. When one billiard ball rolls into another and the second moves off, for example, it is very natural to think that the precise way in which the second moves off is necessitated by a large set of facts, including the velocity of the ball which struck it, the position in which it was struck, the composition and topology of the surface, etc. – meaning that given all these facts together, the precise result which actually follows *must* follow, in virtue of the constitution of the natural world. It is plausible, to be sure, that if the relation between the obtaining of these facts and the facts about the way in

which the second ball moves off is really to be *necessitation*, the set will, in addition, have to include facts which entail the *absence* of the huge range of things which might have interfered with the motion of the second ball, had they been present – e.g. a person positioned so as to be able to stop the ball in its tracks and bent on doing so, a meteorite poised to land at the crucial moment on the table, etc. And perhaps there are insuperable obstacles, in the end, to the idea that any such definite set of facts exists, even in principle - but still, it is not *obvious* that this must be so. For now, anyway, we may remark that whatever one thinks about the coherence of the idea of natural necessitation, the entailment definition seems not to mention it explicitly at all. It says nothing about the future being necessitated, fixed or settled by facts about the past. It says nothing whatever about natural necessitation. Rather, it speaks merely of certain statements being *entailed* by certain other statements. The entailment definition, then, seems to be at one remove from the central worry which motivates the free will problem (assuming I have characterised that worry more or less correctly to begin with).

One might, of course, think that this is all to the good – and that in so far as we have moved away from the original worry with which we began, we have done so only to render it clearer and more perspicuous. We might hope that with the entailment definition, we are substituting clearly defined logical concepts which are tractable and relatively free from obscurity, for murky metaphysical notions out of which not much sense can be made. But there is not much point substituting clear notions for murky ones if we thereby lose the essence of the very idea we were originally trying to express, the core idea about a possible way our world might be whose compatibility with free will seemed doubtful. The question is: will we have we lost our grip on that idea, if we exchange (MD) for (ED)? I shall be arguing that the answer is 'yes' – for the purposes of which argument I now turn to ask the question how precisely (MD) and (ED) are related.

(ii) *The relationship between (MD) and (ED)*

Let us ask first, then, whether (ED) follows from (MD): that is, whether it would follow from the claim that the total set of non-relational facts about a given time causally or physically or naturally necessitates every fact about what happens after that time that a complete statement of the nonrelational facts about a given time together with a complete statement of the laws of nature entails every truth about what happens (or is the case) after that time. I am happy to allow that given a suitably broad conception of what is allowed to count as a statement of a law of nature, it probably does follow –though perhaps only at the cost of robbing the concept of a law of nature of some of its usual connotations. To see this, consider a couple of imaginary possible worlds that might be thought, at first glance, to represent possible obstacles to the entailment of (ED) by (MD). First: one might wonder whether there could be a possible world in which the sum-total of facts holding at a given time naturally necessitated all future matters, but in a way that was extraordinarily complex and irregular, so that, for example, the facts at t1 necessitated the facts at t2 according to one set of general principles, but that the facts at t2 necessitated the facts at t3 according to another, alternative set (and so on). One might wonder whether there could, in such a world, be such things as 'a statement of the laws of nature', and accordingly wonder whether (ED) would be true. But of course, if determinism of the form encoded by (MD) is to be true in such a world, the question which laws are operative at any given moment are going to have to be settled *somehow* by the facts about previous times. And so we will always be able *somehow* to express the 'laws' of this world – including, as it were, the 'meta-laws' – the laws which will enable us to infer which laws will hold at each particular moment. These laws might be untidy and complex – perhaps too complex to be readily usable by a human being to predict what will happen, or to be recognised

in the human world as the 'laws of nature'. But they will still be in principle stateable. And if such laws as these are genuinely to count as laws, then determinism as characterised by the entailment definition will hold in this world – that is to say, it will be true of this world that at any given time, a complete statement of the nonrelational facts about that time, together with a complete statement of the laws of nature will entail every truth about what happens after that time.

Someone might think that this possible world fits the entailment definition only because we have not been sufficiently imaginative. We have remained within a world in which there are laws, even though they are laws which change moment-by-moment. But what about an entirely lawless world? One might perhaps wonder whether the natural necessitation of the character of one temporal world-slice by another must be lawful at all – and if it were not, one might think one had imagined a world in which settling of future by past obtained, but in which no relevant entailments could be found. But this suggestion does not seem to be coherent. What the suggestion invites us to imagine is an entirely particularistic form of settling such that one particular time-slice of world just somehow determines the next temporal world-slice, but not in virtue of any *general character* it has, of the sort we could attempt to encode in something that would suffice at least to serve as a potential *candidate* for a statement of law. It is hard to make sense of what this could mean; it would perhaps demand turning time itself into a causal factor. But even here, one might say, that if (MD) holds, there would still have to be statements of the form 'If p_1….. p_n at t1, then q_1 at t2'; 'if q_1…. q_n at t2 then r_1 at t3', etc. – statements which simply describe, as it were, the particularistic, but deterministic, chaos. And one might simply say that in such a world, these are the laws, and so that the entailment definition continues to hold in a kind of 'limit' form.

This point, or something very like it, is considered by Bertrand Russell, in his classic and highly influential paper, 'On the Notion of Cause' (Russell 1912/1986). Russell suggests that if we are to operate with a notion of determinism such that not every world turns out to be deterministic, it will be essential to insist that *time* must not enter explicitly into the formulae in terms of which the laws of nature are expressed (Russell, 1912/1986, 196). This would prevent the intuitively lawless world just imagined from counting as a deterministic world according to (ED). But could such a world nevertheless be deterministic according to (MD), and thus be an obstacle to the entailment of (ED) by (MD)? I am not sure. It depends on whether there could, as it were, be *brute* causal necessitation – an entirely particularistic drawing on of the future by the past of a form which could not be encoded at all in any way – an idea which I confess I find difficult to make much sense of. Certainly it is hard to see how one could ever have any reason to believe in natural necessitation in such a world.[11] My tentative conclusion, then, is that (MD) probably does entail (ED), provided we are armed with a sufficiently generous conception of a statement of a law of nature, such that the statements in question are permitted to be very great in complexity or very great in number, or both.

It seems possible, then, to defend the claim that (MD) entails (ED). But what about the reverse implication? Does the holding of determinism as characterised by (ED) imply the holding of determinism as characterised by those who advert to the murky notions of settling and necessitation? Here, I think, it is much more obvious that the answer is 'no'. What makes it evident that the answer is 'no' is brought out rather beautifully by Helen Beebee in her paper, 'The Non-Governing Conception of Laws of Nature' (Beebee 2000). In that paper, Beebee contrasts the positions held by what she regards as two main camps in the debate about the metaphysics of laws. On the one hand, there is the realist view that laws are relations of necessity between universals, various versions of which are developed by Dretske (1977), Tooley (1977) and Armstrong (1983). On the other hand, there is what she calls the 'Ramsey-Lewis' view.[12,13] On the Ramsey-Lewis view,

Beebee points out, laws do not *govern*. They are not that kind of thing. According to Lewis, for example, "a contingent generalization is a *law of nature* if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength." (Lewis, 1973: 73). Beebee's explanation of Lewis's idea is very neat and picturesque, and so I quote her here at length:

" … the idea is something like this. Suppose God wanted us to learn all the facts there are to be learned … He decides to give us a book – God's Big Book of Facts – so that we might come to learn its contents and thereby learn every particular matter of fact there is. As a first draft, God just lists all the particular matters of fact there are. But the first draft turns out to be an impossibly long and unwieldy manuscript, and very hard to make any sense of – it's just a long list of everything that's ever happened and will ever happen … Luckily, however … God has a way of making the list rather more comprehensible … he can write down some universal generalizations with the help of which we can derive some elements of the list from others …

… God … wants the list of particular matters of fact to be as short as possible – that is, he wants the axioms to be as strong as possible; but he also wants the list of *axioms* to be as short as possible – he wants the deductive system – the axioms and theorems – to be as simple as possible. The virtues of strength and simplicity conflict with each other to some extent; God's job is to strike the best balance. And the contingent generalizations that figure in the deductive closure of the axiomatic system which strikes the best balance are the laws of nature" (Beebee, 2000: 574).

On the Ramsey-Lewis conception, then, the laws are not rules written into the fabric of reality which the universe must follow. Rather, the universe just consists of "a vast mosaic of local matters of particular fact, just one little thing and then another" (Lewis, 1986, *ix*). Laws are fundamentally *descriptive* and fully knowable by us, as it were, only *post facto* – since of course the laws, on the Ramsey-Lewis conception, are to be the principles which achieve the best balance of simplicity and strength in an axiomatization which fits the whole history of the world, not just the portion of that history which is past. Laws are not *constraints* on the forward flow of reality at all.

Suppose, then, that one held the Ramsey-Lewis view of laws. And imagine a world in which the only laws are Ramsey-Lewis laws and which is also deterministic according to (ED). Must this world also be deterministic in the sense that the past in any way naturally necessitates or settles the future? I think the answer to that question is fairly obviously 'no', as Beebee indeed argues in her paper. With the laws thought of only as descriptive and non-governing, and the world consisting merely of one little thing and then another, there is no necessitation, no settling of the sort that an incompatibilist defender of free will need worry about. As Beebee says, "the sense in which I am constrained … is a purely logical one. And this logical sense surely cannot be an obstacle to free will" (Beebee 2000: 579).[14]

I have argued thus far, then, that even if (MD) implies (ED), (ED) does not imply (MD). There could perfectly well be worlds in which the entailment definition was satisfied, though the metaphysical definition was not – worlds where the laws did not govern – provided only that Ramsey-Lewis laws are permitted to *count* as statements of the laws of nature – a proviso to which I shall return in section (iv). Those like Beebee and Vihvelin who have recognised this rather clearly have both suggested that this gives us a ready route to compatibilism about free will and determinism. Vihvelin, indeed, makes it clear at the outset of her book, *Causes, Laws and Free Will*, that she regards any embrace of a metaphysical style definition of determinism as confusing and misleading, insisting that "determinism should not be confused with the view of laws that has been called "the governing conception of laws", "the pushy explainer view" and most commonly "the necessitarian view"(Vihvelin 2013: 4). It is easy, she says, to get confused, "because determinism is

often formulated in a loose and misleading way, e.g. as the thesis that the facts about the past 'metaphysically determine' or 'necessitate' or 'fix' all future facts" (Vihvelin 2013: 4). She recommends that these loose ways of talking be avoided. But my question is whether this is really the right response to what, I think, is agreed on all hands is an important distinction between different conceptions of what a law of nature might be. For having recognised that the entailment definition can be satisfied when the metaphysical definition is not (provided only that we are permitting "the contingent generalizations that figure in the deductive closure of the axiomatic system which strikes the best balance between simplicity and strength" to count as statements of laws of nature), ought we not to question whether the entailment definition really and truly captures what it was we thought we were worried about when we were worried, originally, about determinism? Does not the coming apart of the two definitions reveal, in fact, that the entailment definition is inadequate to capture the *worrying* doctrine of determinism? In the next section of the paper, I want to address this question.

(iii) *Deciding on the Definition of Determinism*

Vihvelin begins her discussion of determinism in *Causes, Laws and Free Will* by stating what she takes to be its definition – a definition which seems fairly clearly to be a version of the entailment definition: "determinism is the thesis that for every instant of time *t* there is a proposition that expresses the state of the world at that instant, and if P and Q are any propositions that express the state of the world at some instants, then the conjunction of P together with the laws of nature entails Q" (Vihvelin, 2013: 3) [15]. And then, on the basis of this definition, Vihvelin proceeds to argue that those who embrace metaphysical conceptions of determinism and use concepts such as necessitation are confused. But on what basis can she possibly justify this claim? Definitions in philosophy are negotiable. The particular definitions of the crucial concepts with which we operate have histories; and it may be instructive, sometimes, to reflect on why we have ended up with the ones we have, and in this particular case, why we have ended up with a definition of determinism that has the concept of entailment, rather than the concept of natural necessitation, at its heart.

It is an interesting historical question when precisely the doctrine of determinism understood by way of the entailment definition began to be substituted for the more metaphysical-sounding thesis that used to be called the 'doctrine of necessity', and why. Although Hume's (1975/1777, 1978/1749) views on the absence of any proper provenance for the idea of causal necessity must be a very important part of the story that needs to be told, the full adoption of the entailment definition of determinism does not seem to have occurred until rather later. Pierre-Simon Laplace's 'demon' who, knowing the location and momentum of every particle in the Universe at a given time would be able also to know its entire future, that future being, in Laplace's view, a mere consequence of the laws of classical mechanics, was doubtless an important stepping stone towards the entailment definition.[16] But it was not until after various developments in logic and mathematics encouraged philosophers to begin to appreciate the usefulness of the concept of a function that the entailment definition really came into its own. I suspect Russell was very likely a very powerful voice in the shift made by philosophers to definitions of determinism resembling (ED). According to the highly influential paper, 'On the Notion of Cause' which I've already had occasion to mention, Russell claims that:

> A system is said to be 'deterministic' when, given certain data, $e_1, e_2, \ldots, e_n$, at times $t_1, t_2, \ldots, t_n$ respectively, concerning this system, if E_t is the state of the system at any time t, there is a functional relation of the form

$E_{t} = f(e_1, t_1, e_2, t_2, \ldots, e_n, t_n, t)$ (Russell: 190)

On this conception of determinism, determinism holds wherever the way things are at a given time is a function of the way they are at another time – a function that will be encoded by some relevant law of nature. But there need be no implication at all that the law *constrains* anything; for all Russell's definition of determinism implies, the law which encodes the functional relation may just be a Ramsey-Lewis law. The functional relation is entirely descriptive, and merely expresses the fact that where such a relation exists, one can in principle infer the state of the world at one moment in time from the state of the world at another. Moreover, the relationship between past and future states is now completely symmetrical – one can infer past states from future ones, as well as the other way around. Determinism thus ceases to be merely a doctrine about the *development* of the world through time – and becomes instead a doctrine about the relationship between descriptions of the world at different times, a relationship which has no particular implications pertaining to *constraints* on the development of reality.

What motivated Russell to formulate the concept of determinism in this way? Earlier in the same paper, Russell expresses his suspicion of the idea that causes *compel* their effects – compulsion, he insists, is "a complex notion involving thwarted desire"(182) – a concept of psychology and not of metaphysics. One can see in this critique the influence of the Humean idea that we have no clear idea at all of necessitation, and that in so far as the idea has a source in our experience, that source is in the mind and not in external objects. Empiricism had no room for the unverifiable existence of necessary connections between states of affairs, and the concept of determinism developed by adherents of empiricism unsurprisingly followed suit. I believe, then, that we should see the entailment definition as the product of a long and important period of philosophical history in which realistic ideas about powers and causes faced an empiricism that refused to countenance them, and also at the same time, had to be encompassed within a mathematicization of parts of metaphysics that promised new ways of thinking about the concept of law. The doctrine of necessity was effectively, then, declared dead owing to a bad case of unintelligibility: and the entailment definition entered in its place as the respectable expression of the doctrine of determinism.

However, unless I am much mistaken, I think it is fair to say that philosophy has now entered a new phase of development, in which that empiricism itself faces serious challenges. Those who believe in the reality of such things as powers and natures and objective laws which constrain reality have been once again admitted into the fold of philosophical respectability.[17] Of course, it cannot yet be said that they have won the day. There are still those who do not think that the idea that laws truly constrain the development of reality can ultimately be made sense of; regularity theories of causation and Ramsey-Lewis theories of law remain quite widely held. But there are now many adherents of more realistic conceptions of natural necessitation. We cannot, then, simply assume as perhaps Russell once did, that determinism as defined by (MD), is the mere product of silly confusions of psychological concepts with metaphysical ones. There are plenty of philosophers these days who are absolutely happy to sign up to the existence of natural necessitation, either because they believe in constant objective forces, such as the fundamental forces postulated by physics, or in governing laws, or because they embrace the doctrine of intrinsic natures which dictate how objects and substances will behave with respect to one another. But if it is respectable once again to believe in real, producing causation and objective, metaphysical laws, it should be respectable once again to formulate determinism in ways closer to the ways in which the doctrine of 'necessity' used to be formulated – as a firmly metaphysical thesis about the natural necessitation of future facts or states of affairs, by past ones. And in that case, we are not entitled simply to *assume* that (ED) is the correct definition of determinism – as though that were written somewhere in an unchallengeable

philosophical dictionary. We are entitled to ask whether (ED) might not itself be the outdated relic of a long-gone empiricist era and whether we do not have an obligation, as the custodians of a rather different philosophical period, to ask whether it is really any longer fit for all the purposes to which we might wish to put it.

(iv) *Why (MD) is to be preferred to (ED) as a definition of determinism*

Thus far, then, I have argued that there is no clear reason that is agreed on all hands for regarding the thesis expressed by (MD) as illegitimate, and so that arguments such as Vihvelin's that those who adopt alternative definitions are obviously 'confused' should be rejected. In this final section of the paper, I want to suggest that, more than that, we should re-embrace (MD) as our preferred definition of determinism when it comes to the free will debate.

The main reason for thinking that (MD) offers a better definition of determinism than (ED) for the purposes of the free will debate is that it more exactly expresses the thesis we have reason to suppose might possibly be in tension with free will. Vivhelin points out that her own (ED)-like definition of determinism is strictly neutral concerning the concept of law and is consistent with the Ramsey-Lewis conception as well as what she calls the governing conception. She may even regard its neutrality as an advantage. But neutrality, I submit, is not what we should want from our definition of determinism, if we are seeking a definition suitable for formulating the free will problem. We should want a definition of determinism that makes determinism look, at least *prima facie*, as though it might present some sort of serious threat to the existence of free will. When we teach students to state problems in philosophy, we teach them to offer the best version of the problem that they can manage, even if – perhaps even *especially* if – their eventual intention is to argue that the problem can be solved. That is presumably because we think it is a methodological *desideratum* to do so. And it is this same methodological assumption on which I mean to rely here. For, I submit, there simply *is* no threat to free will from determinism as construed by the entailment definition, provided a 'complete statement of the laws of nature' is permitted to be a complete statement of the regularities which achieve the best combination of simplicity and strength in the axiomatic system that attempts to organise God's Big Book of Facts into a more digestible read. If laws are merely descriptive and have no constraining power, (ED) can be satisfied easily in a way which does not even begin to raise any worries about free will. The question should go away before we even have chance to raise it. And so to use (ED) to state the free will problem is to state it in a way which makes it much too easy to solve. To make the free will problem hard, to assure ourselves of a real issue, we need the metaphysical definition. We need to raise the spectre of the world as governed by real necessitating laws, or natures, in order to be confident we have described the problem in its most worrisome form.

A second, related reason for thinking that (ED) should be replaced by (MD) is that it would generate a more illuminating division of positions on the free will debate into compatibilist, on the one hand, and incompatibilist, on the other. If (ED) is the definition of determinism, indeed, I find it quite hard to see why *anyone* would remain an incompatibilist, once it had been pointed out to them how weak a thesis (ED) actually is. It is not very hard to see (as Beebee and Vihvelin argue) that Ramsey-Lewis laws do not constrain us – so provided these laws are permitted to count as the laws whose 'statement' figures in (ED), it should be clear that we ought all to be compatibilists. But do we really *want* a definition of determinism which leaves no one sensible in the incompatibilist camp? Surely we should prefer a definition which states a thesis which some, at least, believe creates problems for our belief in free will? – at any rate, if a coherent statement of such a thesis exists? But

if (MD) is coherent, as I have been assuming for the purposes of this discussion, then (MD) states such a thesis. (MD) ought therefore to be preferred to (ED) as the definition of determinism.

I want now to address two possible objections to the case I have made for replacing (ED) with (MD).

The first and I think most pressing objection is that my claim that the problem with (ED) is about its central use of the relation of *entailment* is mistaken – since we could perfectly well meet the desideratum I have proposed (that is, of ensuring that determinism remains a thesis strong enough to figure in the statement of a potential threat to free will) and yet retain (ED) as the definition of determinism, *provided* we are careful what we permit to count as a 'statement of a law of nature'. It might be conceded that mere Ramsey-Lewis descriptive generalisations might indeed need to be excluded if we are to represent determinism as a doctrine potentially threatening to free will. But provided law statements themselves somehow incorporate the idea of natural necessitation within them, it might be thought, it will not matter if we retain the entailment definition.

But the question I would like to ask is how exactly it is proposed to incorporate the idea of natural necessitation into the 'complete statement of the laws of nature' which (ED) refers to. What will the laws look like with the modality of natural necessitation built in? None of the laws we actually have in science contains a modality operator of any kind; it is perfectly consistent with the expression, for example, of Boyle's Law [18]or Coulomb's Law[19] that they should be merely generalised descriptions of relationships which have made it into the deductive closure of the Lewisian 'best system'. The statement of these sorts of laws, then, if we leave them as they are normally stated, will not do. Perhaps it might be suggested that we might introduce something like a natural necessity operator - call it 'N' – so that we could express law statements themselves, or at any rate, their deductive consequences, as natural necessities (e.g. N $\forall x[Fx \rightarrow Gx]$); or alternatively, that we might use strict conditionals (e.g. $\forall x\, N[Fx \rightarrow Gx]$)). But the trouble with these suggestions is that the strengthening of the premises which 'N' has been used to introduce is simply not *required* in order for the wanted entailments of future facts by past ones (together with laws) to go through. The entailments will go through anyway, with or without the modal operators. The modal operators are mere idle cogs in the mechanism of the deductions which yield the future facts as conclusions. How then can it be *expressive* of the distinctive thesis of determinism to say that these entailment relations hold (as the entailment definition proposes)? They hold *in any case* (given only that (ED) as I originally conceived it applies) whether the modal operators are there or not, which seems to reveal that mere commitment to the entailments is not properly expressive of the determinist's claim. What (worrisome) determinism is fundamentally about is the claim that the natural laws do indeed *need* to be stated using modal operators of some sort (along with the important additional claim that these laws are all-embracing) – that mere universal generalisations or specifications of functional relationships will not do. But this distinctive claim is not stated, and nor is it implied, by the entailment definition.

One might respond to this difficulty by suggesting that rather than attempting to incorporate modal operators into the laws, it should just be stated to be a *presupposition* of the wanted reading of (ED) that the complete statement of the laws of nature must be a complete statement of *genuine* laws which truly reflect natural necessities in nature, rather than just Ramsey-Lewis laws. That would, I think, ensure an effective equivalence between (MD) and (ED). However this result comes only at the cost of making it absolutely plain that a substantial part of the burden of ensuring that the doctrine of determinism is at least conceivably a threat to free will is borne by the presupposition, rather than by the explicit content of the so-called 'definition'. And surely if we are trying to *state* a thesis, the most perspicuous way to state it is to bring everything crucial, so far as is possible, into the *explicit content* of our statement. We need the definition of determinism to state

what it is that one believes in believing in determinism. And so it will not do to attempt to meet my concern that (ED) guarantees no genuine problem of free will by suggesting that we retain (ED), while noting alongside it that it will only do to specify the doctrine, given a certain additional presupposition. For this will not enable (ED) to meet the purpose for which it was wanted in the first place – that of stating as plainly as possible the exact content of what must be believed by someone who thinks determinism is true (rather than (at best) a mere *consequence* of that doctrine). We would need to *add* the content of the presupposition that the laws encode natural necessities to the content of (ED) in order to state the doctrine of determinism. But that seems to be as much as to say that what we *really* need is (MD) - a definition which makes it more perspicuous that what is really crucial to determinism is that there be *natural* necessitation relations between the facts at one time and the facts at another (perhaps, but not necessarily in virtue of the holding of laws, depending on one's view of the ontology of natural necessity). It is (MD), that is, which properly encodes what the worrisome deterministic thesis really is.

The second objection I want to consider is that I am simply arguing for a definition of determinism that will make incompatibilism true by fiat. But it should really go without saying that this is not the case. There are many interesting ways to argue for compatibilism, even if one believes in real, metaphysical necessitation and constraining laws – for example, by arguing that determinism doesn't entail inevitability (see e.g. Dennett, 2003); by endorsing 'ability plus opportunity' conceptions of free will (see e.g. Kenny, 1975); or by arguing that free will *requires* natural necessitation (see e.g. Hobart, 1934). The embrace of (MD) by no means brings incompatibilism inevitably in its wake; there will still be plenty of ground for a wide range of different kinds of compatibilists to occupy. We will simply have removed from the mix as misleading a view of the nature of law which can only muddy the waters – which confuses those I would be inclined to regard as *true* compatibilists, who really *do* believe that free will is compatible with real honest-to-goodness necessitation, with an ersatz sub-group, who secure their compatibilism only by effectively denying with their conception of law the truth of the metaphysical thesis of determinism. So far as the defender of (MD) is concerned, those who believe only in Ramsey-Lewis laws simply *deny* determinism; about (MD), though (supposing they allow its coherence), they might be either compatibilists *or* incompatibilists, for all we know.

Perhaps I should end by saying, in case anyone should be confused about this, that I am no believer in (MD) – nor even a confident believer in its coherence. As stated at the outset of the paper, I think it shares with (ED) a number of problematic concepts; and even if worries about these concepts were to turn out to be unfounded, I think it is vanishingly unlikely that anything like (MD) might conceivably be true – not because I think there is no such thing as natural necessitation, but rather because I do not believe that the reach of natural law is total, it being likely, in my view, that such laws as there are may *constrain* everything without *dictating* everything. I have defended (MD) here, then, not because I think it is true, but because I regard it as an honest attempt, at least, to convey a vision of reality which looks potentially troubling for free will. And if we are going to argue about whether free will is possible, given determinism, we owe determinism the courtesy of expressing it in a form which at least makes it *look* as though it might present some sort of problem for free will. (ED) does not meet this criterion; and nor can we readily adjust it in order to build in the additional assumptions we require. For the purposes of the free will debate, we should ditch the entailment definition of determinism and return to its venerable older relative, the doctrine of necessity.

Reference List

Armstrong, David. 1983. *What is a Law of Nature?* Cambridge: Cambridge University Press.
Austin, J.L. 1970. "Ifs and Cans." In his *Philosophical Papers*, 2nd edition, 205-32. Oxford: Oxford University Press.
Beebee, Helen. 2000. "The Non-Governing Conception of Laws of Nature", *Philosophy and Phenomenological Research* 61: 571-594.
Beebee, Helen, and Mele, Alfred. 2002. "Humean Compatibilism", *Mind* 111: 201-23.
Bird, Alexander. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.
Dennett, Daniel. 2003. *Freedom Evolves*. Penguin, London.
Dretske, Fred I. 1977. "Laws of Nature", *Philosophy of Science* 44: 248-68.
Ellis, Brian. 2001. *Scientific Essentialism*. Cambridge: Cambridge University Press.
Fine, Kit. 2005. "The Varieties of Necessity" In his *Modality and Time: Philosophical Papers*. Oxford: Oxford University Press, 235-260. (First published (2002) in *Conceivability and Possibility* ed. T. Gendler and J. Hawthorne, Oxford: Oxford University Press: 253-82).
Fischer, J.M. 1983. "Freedom and Foreknowledge", *The Philosophical Review* 92: 67-79.
Fischer, J.M. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
Harré, Rom and Madden, E.H. 1975. *Causal Powers: A Theory of Natural Necessity*. Oxford: Blackwell.
Hobart, R.E. 1934. "Free Will as Involving Determination and Inconceivable without it", *Mind* 43: 1-27.
Hume, David. 1975/1777. *An Enquiry Concerning Human Understanding* in *Enquiries* 3rd edition, edited by L.A. Selby-Bigge with text revised and notes by P.H Nidditch. Oxford: Oxford University Press.
Hume, David. 1978/1740. *A Treatise of Human Nature*, 2nd edition, edited by L.A. Selby-Bigge. Oxford: Oxford University Press.
James, William. 1968. "The Dilemma of Determinism". In his *Essays in Pragmatism*, 37-64. New York: Hafner.
Kane, Robert. 2002. *The Oxford Handbook of Free Will*. Oxford. Oxford University Press.
Kenny, Anthony. 1975. *Will, Freedom and Power*. Oxford: Blackwell.
Laplace, Pierre-Simon. 1951/1814. *A Philosophical Essay on Probabilities*, translated into English from the French 6th ed. by F.W. Truscott and F.L. Emory. New York: Dover Publications.
Lewis, David. 1973. *Counterfactuals.* Cambridge, MA: Harvard University Press.
Lewis, David. 1986a. *Philosophical Papers*, Vol II. Oxford: Oxford University Press.
Lewis, David. 1986b. "Are we Free to Break the Laws?". In his 1986a: 291-98.
Mumford, Stephen. 2002. *Laws in Nature*. London: Routledge.
Ramsey, Frank. 1978. "Universals of Law and Fact" in his *Foundations: Essays in Philosophy, Logic Mathematics and Economics*, edited by D.H. Mellor, 128-32. London: Routledge and Kegan Paul.
Russell, Bertrand. 1925/1986. "On the Notion of Cause" in his *Mysticism and Logic* 173-99. London: Unwin.
Tooley, Michael. 1977. "The Nature of Laws", *Canadian Journal of Philosophy* 7: 667-98.
Van Inwagen. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
Van Inwagen. 2017. *Thinking about Free Will*. Cambridge: Cambridge University Press.
Vetter, Barbara. 2015. *Potentiality*. Oxford: Oxford University Press.
Vihvelin, Kadri. 2013*. Causes, Laws and Free Will: Why Determinism doesn't Matter*. Oxford: Oxford University Press.
Wiggins, David. 2003. "Towards a Reasonable Libertarianism." In *Free Will* 2nd edition, edited by Gary Watson, 94-121. Oxford: Oxford University Press.

[1] Of course, it is possible to believe that some of these other, special varieties of determinism might also be threatening for free will; but I restrict myself, for the purposes of this paper, to the centrally important *general* thesis. It is also important to note here, at the outset, that my questions concern determinism *considered as a thesis important to the free will debate*. There may be other purposes for which philosophy (or science, or theology, or economics, say) require a concept of determinism; none of my arguments here need bear on the question what concept might best suit those other purposes.

[2] I borrow this particular formulation from Fischer (1998, 14); but many other philosophers offer very similar definitions. See, for example Van Inwagen (1983, 65); Kane (2002, 4); Wiggins (2003, 99); Vihvelin (2013, 1).

[3] In fact, one might think we could give an even more general description than any of these of what worries us when we worry that determinism might be true. One might say that what we worry about is simply that *something beyond our control* somehow completely dictates the whole future course of reality (and leave it entirely open whether the 'something beyond our control' is past facts together with the physical laws; or God and his wishes or dictates; or God's knowledge of what we will do; or the position of the stars …. or something else). But the most prevalent source of the worry for most modern philosophers stems from a thought that is essentially (i) secular and (ii) intended to be at the very least compatible with a scientific world view, if not actually supported by it. I therefore focus for present purposes on this specific version of the worry.

[4] Plausibly, the idea that something is 'fixed' or 'settled' is the idea that it is not possible for it to be otherwise (given that certain things – such as the laws of nature, for example - are held fixed).

[5] Ontological issues may raise their ugly heads here – there will be those who think of facts as essentially linguistic items and hence as things which could not be involved in worldly necessitating. So be it. Those who have these worries may substitute the term 'state of affairs' or 'event' (understanding by 'event' something like an exemplification of a property at a time) if it makes them feel easier. Nothing should turn on this.

[6] One possible confusion which might be introduced by the use of the word 'metaphysical' which I need to avert: metaphysical determinism is not committed to the metaphysical necessity of the laws of nature. It is perfectly consistent with metaphysical determinism as I have stated the doctrine that the laws are contingent.

[7] Though it may not follow that *no* such future facts are thus entailed, without mention of laws. To take an example of Fischer's, "Smith existed at T1" entails "It is not the case that Smith existed for the first time at T2", without the addition of any laws of nature as extra premises (Fischer, 1983: 75). Thanks to Marco Hausmann for reminding me about examples of this sort.

[8] In recent years, this view has been most closely associated with Dretske (1977), Tooley (1977) and Armstrong (1983).

[9] See e.g. Harré and Madden (1975), Ellis (2001), Mumford (2002), Bird (2007), Vetter (2015).

[10] See especially Mumford (2002), for a very explicit repudiation of the idea that belief in natural necessity commits one to the existence of laws.

[11] The view that this idea makes no sense also seems to be in accordance with a suggestion made by Fine (2002) that:

> "there are no distinctive *de re* natural necessities. Let us suppose that *x* and *y* are two particles and that it is a natural necessity that they attract one another … Then it is plausible to suppose that this should follow from (1) its being a metaphysical necessity that each of the particles is of the kind that it is and (2) its being a natural necessity that particles of this kind attract one another. Thus the *de re* natural necessity will reduce to a *de re* metaphysical necessity and a *de dicto* natural necessity; and it might be thought that something similar should be true of any *de re* natural necessity" (243).

[12] Ramsey (1978).

[13] Lewis (1973) and (1986).

[14] Beebee's conception of laws enables her to respond to Van Inwagen's well-known Consequence Argument (which of course uses the entailment definition of determinism) in a similar way to David Lewis (1986b) by arguing against the premise that claims that no agent can render the laws of nature false (Beebee and Mele, 2002). Whilst Beebee and Mele concede that it is true that no agent can render false what *in fact* turns out to be the final set of laws (any more than I can render false what *in fact* turn out to be the truths about the future), it is not true that no agent can contribute with their actions to the description of the Universe which needs codifying as neatly as possible by such laws – and hence it is not true that no agent can indirectly change with their actions what those finalised laws might eventually turn out to be. This seems to me the right response to the Consequence Argument for someone who does not believe in natural necessity.

Van Inwagen himself claims that even if something like Lewis's Humean view of laws as mere exceptionless regularities was correct, the Consequence Argument holds up. His reasoning is that it would not

follow from the fact that laws were no more than exceptionless regularities that any human being would be able to render any such regularity false:

> "Suppose, for example, that the most massive star is 260 times as massive as our sun. 'All stars have masses less than or equal to 260 solar masses' may well be a mere exceptionless regularity: it may well be that there could have been a star with a mass of 261 solar masses. But, no doubt, no human being is (or ever has been or ever will be) able to cause a counterinstance to this regularity to exist. I would suppose that, even if Lewis's Humean conception of laws is right, it would be an even more difficult task to produce a counterinstance to a law – in the sense in which it would be "even more difficult" for me to lift an object weighing 10,000 kilograms than it would be for me to lift an object weighing 1,000 kilograms (Van Inwagen, 2017: 228, fn.36).

It seems true that there may be some mere exceptionless regularities that cannot be rendered false by a mere human being and hence that it does not follow from a law's merely *being* such an exceptionless regularity that a human being could render it false. But this is not an adequate defence against Beebee. For the Consequence Argument requires that no human being could render false the statement which states *the sum total* of the laws of nature, so it will not do as a response to Beebee to point out merely that *some* exceptionless regularities cannot be rendered false. It must be the case that *none* of those which are at any point in time potential candidates to become laws, should the Universe go in this or that direction, can be rendered false. The Consequence Argument will have a false premise if *any* of those which would otherwise have made it into the Big Book of Laws might have failed to make it onto the statute book after all, should a given agent have acted differently; and the onus is on Van Inwagen to show that this is definitely *not* going to be the case. He admittedly appears confident that it is *not* going to be the case ("I would suppose that … it would be an even more difficult task to produce a counterinstance to a law"). But the question is what justifies this confidence. If we are assuming Humeanism about laws, it cannot be faith in natural necessity. And Beebee would doubtless reply to Van Inwagen's assertion that it is going to be a difficult task to produce a counterinstance to a law, that if determinism is true, it is going to be plausible that breaking the laws cannot be any more difficult than raising one's hand (for this action, which most of us assume we generally have the ability to undertake whenever we wish, would be bound to amount to law-breaking on many occasions in a deterministic world). Thanks to Marco Hausmann for alerting me to the relevance of Van Inwagen's response to Beebee.

[15] This definition is based on that offered by van Inwagen (1983: 65).

[16] Laplace (1951/1814).

[17] A huge range of work from the last fifty years or so might be mentioned here. Amongst some of the most prominent work in the tradition I have in mind are Harré and Madden (1975), Dretske (1977), Tooley (1977), Armstrong (1983), Ellis (2001), Mumford (2002), Bird (2007), Vetter (2015).

[18] Usually stated as P1V1 = P2V2, Boyle's law states that at constant temperature, the pressure of a given quantity of gas is inversely proportional to its volume.

[19] Coulomb's Law states that the electrical force between two charged objects is directly proportional to the product of the quantity of charge on the objects and inversely proportional to the square of the separation distance between the two objects.