This is a repository copy of *Local Miracle Compatibilism*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/183835/

Version: Accepted Version

Local Miracle Compatibilism

Helen Beebee

1. Introduction

To those people who have not spent much time thinking about the problem of free will, it seems that we frequently have the ability to do otherwise than what we do. Provided I am not subject to any compulsions or coercive forces (addiction, bondage, someone holding a gun to my head, and so on), I am able now to raise my hand, drink some water, read the newspaper, and so on – and I am able to do those things irrespective of what I in fact end up doing. It further seems that there are certain things that we are unable to do. I am unable to throw a stone faster than light, or to build a machine that, when operated, would cause some particles to violate the principle of the conservation of angular momentum. And I am unable to do these things not just because my arm is not strong enough, or because I am no physicist or engineer, but because such acts would violate the laws of nature. It would be nice if such an intuitively plausible story about our abilities were true – not least because, if free will and moral responsibility require the ability to do otherwise, it would (in suitable circumstances) render us free to perform, and hence morally responsible for, our everyday decisions and actions, yet not free to throw a stone faster than light.

Standard libertarians hold that this intuitive story is correct. The laws of nature plus past facts, they claim, leave it genuinely open whether or not we will make the decisions and perform the actions we actually make and perform, so we are able, in a quite uncontroversial

sense, to do otherwise. But, since the laws plus past facts do <u>not</u> leave it open whether we run faster than light or do something that violates the principle of the conservation of angular momentum, we are unable to do those things. Of course, the libertarian's view, <u>qua</u> claim about the actual world, is something of a hostage to scientific fortune. If the libertarian is right about what is <u>required</u> for the ability to do otherwise, it might turn out that we never actually have the ability to do otherwise (determinism may yet turn out to be true, so that laws plus past facts leave only one possible course of events genuinely open), or that we hardly ever have that ability (if, say, the laws that are relevant to action are mostly deterministic), or that we sometimes act freely and sometimes do not, but that we have no privileged access to this information. (Suppose that some of the relevant laws are deterministic and some are not. Then whether or not I am able to raise my hand depends on very subtle facts that would (at best) only be available to a fantastically knowledgeable scientist. And if the world is indeterministic enough, it might turn out that common-or-garden cases of compulsion or coercion are after all cases where we are able to do otherwise – even if the laws grant only a tiny chance that the compulsive or coercive forces will fail.)

Still, even if the actual world does not turn out to be the way the libertarians say it is, they do at least have a plausible theory, insofar as it respects the intuitions described above, of what it would <u>take</u> for agents to be able to do otherwise – and a theory which, for all we now know, might apply to the actual world. They can draw a natural distinction between, say, the ability to raise one's hand (in normal circumstances) and the ability to raise it faster than light: one has the first kind of ability just if (as, for all we know, might be the case) the laws plus the past do not guarantee that one does not do so; but no agent at this or any other

possible world has the ability to break that world's laws, since the laws absolutely rule out any such act.

Compatibilists have a considerably harder time of it when it comes to distinguishing between the two kinds of ability because, according to a standard incompatibilist argument, determinism would render us unable, ever, to do otherwise than what we do. Never mind running faster than light: if the laws and past facts together entail that I will not raise my hand, then I am unable to do even that. Raising my hand would require a miracle – and nobody is able to perform miracles. So I am unable to raise my hand.

One way that compatibilists have tried to escape the unwelcome conclusion that we are no more able to raise our hands (when it turns out that we do not in fact do so) than we are able to run faster than light is to offer a 'conditional analysis' of ability claims: to say that I was able to do X, even though I did not in fact do so, is just to say that, had I chosen, or decided, or wanted to do X, I would have done it. Conditional analyses seem, at first sight, to capture the intuitive distinction between what we are and are not able to do. It seems plausible to suppose that, had I chosen or decided or wanted to, I would have raised my hand. (After all, when not impeded or coerced, I do generally do things like raise my hand when I choose or decide or want to do so.) And it seems equally plausible to suppose that, however much I wanted to, or however sincerely I decided, or however hard I tried, I would not run faster than light.

The familiar incompatibilist retort, however, is that, given determinism, and given that I did not in fact choose or decide or want to raise my hand, I was not able so to choose or decide or want. Grant that I would have raised my hand if I had chosen to; still, choosing to

raise it, and hence raising it, would have required a miracle, and nobody is able to perform miracles.[1]

A different compatibilist move – one that has been made by David Lewis (1981) – is to claim that the mere fact that my raising my hand would require a miracle does not render me unable to do it. Lewis claims that he <u>is</u> able to raise his hand, even though a miracle would have to occur in order for him to do so. However, according to Lewis, while (1) deterministic agents are able to do things, the doing of which would <u>require</u> a miracle to happen, (2) they are not able to <u>perform</u> miracles: they are not able to do things, the doing of which would <u>be</u> a miracle.

On Lewis's compatibilist view, then, the rather quick argument advanced above on behalf of the incompatibilist – 'raising my hand would require a miracle – and nobody is able to perform miracles. So I am unable to raise my hand' – rests on a fallacy of equivocation. True, given determinism, raising my hand <u>would</u> require a miracle. And, true, nobody is able to perform miracles. I am nonetheless able to raise my hand, since one can perform an action that <u>requires</u> a miracle without thereby <u>performing</u> a miracle.

Lewis's view seems to provide a compatibilist position that does justice to the intuitions mentioned at the beginning of the paper: I am able to raise my hand, but I am not able to raise it faster than light. Following John Martin Fischer, I shall call this view – the conjunction of claims (1) and (2) above – 'local miracle compatibilism' (so called because the view appeals to Lewis's theory of counterfactuals, and in particular on the notion of a 'local miracle').[2] The central claim of this paper is that local miracle compatibilism is not a sustainable position. After some scene-setting in sections 2 (a brief rehearsal of Lewis's analysis of counterfactuals and his response to van Inwagen's 'consequence argument') and 3

4

(a discussion and proposed definition of the notion of a 'law-breaking event'), I present, in section 4, a concrete argument against local miracle compatibilism. I argue that, given (1), nothing in Lewis's analysis of counterfactuals provides him with the resources necessary to justify (2).

Finally, in section 5, I offer a diagnosis of the failure of local miracle compatibilism. My claim will be that there is no conception of laws of nature according to which both (1) and (2) can be made to sound simultaneously plausible. A Humean conception of laws makes (1) plausible but not (2), while a necessitarian conception of laws makes (2) plausible but not (1). On a Humean view of laws, the laws do not constrain what happens, while on a necessitarian view, they do. Either way, there is no relevant difference between miracles that one <u>performs</u> by doing <u>X</u> and miracles that are merely <u>required</u> in order to do <u>X</u>.

Compatibilism therefore does not appear to be able to provide what our pre-theoretical intuitions seem to require: a substantive distinction between what we are typically able to do (raise our hands, for example) and what we are never able to do (perform miracles).

<u>2. Are we free to break the laws? (Yes and no.)</u>
First some scene setting. According to Lewis, a counterfactual $\underline{A} \rightarrow \underline{B}$ is true if and only if $\underline{B}$ is true at the $\underline{A}$-world (or worlds) which is closest – that is to say, most similar – to the actual world. For example, suppose I don't in fact raise my hand at time $\underline{t}$. To evaluate the counterfactual, 'if I had raised my hand at $\underline{t}$, then $\underline{B}$ would have been the case', we need to see which out of all the many and diverse possible worlds where I raise my hand at $\underline{t}$ is the most

similar to the actual world, and see if <u>B</u> is true at that world (or worlds, if there is a tie for the closest such world).

For Lewis, the similarity of one world to another is a function of two factors: the extent to which the two worlds' laws of nature are the same, and the extent to which the 'matters of particular fact' are the same. The respects of similarity and dissimilarity are traded off against one another:

(1) It is of the first importance to avoid big, widespread, diverse violations of law.

(2) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.

(3) It is of the third importance to avoid even small, localized, simple violations of law.

(4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly. (Lewis 1979, 47-48)

In general, assuming determinism and assuming that event <u>e</u> does not in fact occur, the closest world <u>w</u> at which <u>e</u> occurs will be one whose past up until either the time of <u>e</u> or a little before the time of <u>e</u> is exactly the same as the actual world. At some time <u>t</u> at or before the time of <u>e</u>, a small 'miracle' occurs at <u>w</u>: something happens at <u>w</u> that is incompatible with the conjunction of the prior-to-<u>t</u> matters of particular fact at <u>w</u> and the <u>actual</u> laws of nature.

Two features of Lewis's account are particularly important in the context of a discussion of local miracle compatibilism. One is that the notion of a 'miracle' is an inter-world notion. To say that the closest world <u>w</u> where <u>e</u> occurs contains a miracle is <u>not</u> to say that some event that occurs at <u>w</u> violates <u>w</u>'s laws: there is no world at which anything happens which violates that same world's laws. Rather, the claim is that an event occurs at <u>w</u>

that violates the <u>actual</u> laws of nature – the laws that obtain at our, actual world. (By the same token, plenty of things that occur at the actual world are miraculous relative to other worlds' laws.)

The second important feature is that the spatio-temporal region of perfect match of matters of particular fact between <u>w</u> and the actual world is to be traded off against the 'size' of the miracle required to bring about <u>e</u>. For example, suppose that in fact I remain seated at my desk for the next half hour. Which is the closest world in which I arrive at the pub in ten minutes' time (event <u>e</u>)? At one <u>e</u>-world, $\underline{w}_1$, I remain seated at my desk for, say, 9 minutes and 59.9 seconds, so that the whole of the past up until that moment is exactly the way it is at the actual world, and then spontaneously disappear, reappearing in the doorway of the pub a tenth of a second later. At another <u>e</u>-world, $\underline{w}_2$, something happens to me after, say, 5 minutes – the whole of the past up until that (earlier) moment being exactly the way it is at the actual world – which induces me to walk in the usual fashion to the pub, arriving exactly 5 minutes later. Let's suppose the thing that happens to me after 5 minutes at $\underline{w}_2$ is that I suddenly acquire a strong desire for beer.

Assuming determinism, both $\underline{w}_1$ and $\underline{w}_2$ contain miracles relative to the actual world. Facts about the distant pasts of both worlds are exactly the same at they are as the actual world, so those facts plus the <u>actual</u> laws of nature are incompatible with <u>e</u>'s occurrence. $\underline{w}_1$ preserves perfect match of matters of particular fact for several minutes longer than does $\underline{w}_2$; on the other hand, the miracle that occurs at $\underline{w}_1$ is arguably much bigger than the miracle that occurs at $\underline{w}_2$. At $\underline{w}_2$, all that is required by way of a miracle is that I acquire a strong desire for beer – and perhaps this can be brought about by a few extra neurons firing, say. The miracle required at $\underline{w}_1$, however, is much bigger. Extremely large numbers of molecules need to

7

vanish from one place and reassemble, in just the right configuration, in another.[3] Hence (arguably), since a large miracle occurs at $\underline{w_1}$ and only a small one occurs at $\underline{w_2}$, $\underline{w_2}$ is closer to the actual world – despite the slightly longer match of matters of particular fact at $\underline{w_1}$.

With Lewis's analysis of counterfactuals in place, let us define 'weak' and 'strong' abilities thus:

Weak ability: The ability to perform some act A, such that, were you to do A, an actual law of nature would have been broken – but not by your act or by any effects of your act.

Strong ability: The ability to perform some act $\underline{A}$, such that, were you to do $\underline{A}$, $\underline{A}$ itself would be, or would cause, a law-breaking event.[4]

For now, we can think of a 'law-breaking event' as follows. Let $\underline{L}$ be the conjunction of all the actual laws of nature. Then $\underline{e}$ is a law-breaking event if and only if, necessarily (that is to say, in all possible worlds), if $\underline{e}$ occurs then $\underline{L}$ is false.[5]

How does the distinction between weak and strong abilities carve up the set of possible acts? Well, consider first the ability to raise your hand faster than the speed of light. If in fact you have such an ability, then it is a strong ability: the closest world where you raise your hand faster than the speed of light is a world where that act itself is (relative to the actual world) a law-breaking event. The ability to build a machine that would cause some particles to violate the law of the conservation of angular momentum is also (if you have it) a strong ability: the act in question (building the machine) is such that, were you to perform it, your act would cause an event that is incompatible with the actual laws of nature. (There is no possible world at which $\underline{L}$ – which includes the principle of the conservation of angular momentum – is true and there are particles that violate that principle.)

8

On the other hand, the ability to raise your hand in a moment's time (if you have that ability) is a weak ability, on Lewis's view. The closest world at which you perform that act is one at which there is a small miracle just <u>prior</u> to the hand-raising; thereafter, the course of events (including the hand-raising itself) proceed according to the actual laws of nature. So the act of hand-raising itself, were you to do it, would neither be, nor cause, a law-breaking event.

Given these examples, it seems eminently plausible to think that deterministic agents (at least sometimes) have weak abilities, but never have strong abilities. This is precisely Lewis's view. And fortunately, not only does the view seem plausible, but it also allows Lewis to rebut Peter van Inwagen's 'consequence argument' for incompatibilism.[6] Here is a toy version of van Inwagen's famous argument:

(1) If determinism is true, then facts about what I do are the consequences of propositions stating the laws of nature plus true propositions about the past.

(2) I am unable to render propositions about the past false.

(3) I am unable to render propositions stating the laws of nature false.

So

(4) I am unable to render the consequences of those propositions false.

(5) Free action requires the ability to do otherwise than what one actually does (that is, the ability to render false propositions stating facts about what we actually do).

So

(6) If determinism is true, I never act freely.

Lewis in effect accuses van Inwagen of equivocating in premise (3). Lewis distinguishes between two readings:

(3a)   I am unable to do anything such that my act would be or cause a law-breaking event

(3b)   I am unable to do anything such that were I to do it, an actual law of nature would be false (and hence not a law).

Given the notions of weak and strong ability, (3a) amounts to the claim that I do not have strong abilities, while (3b) amounts to the claim that I have neither strong nor weak abilities.[7]

Lewis claims that (3b) is false: he claims that he is, for example, able to raise his hand, and that this ability is a weak ability. So replacing (3) with (3b) would render the consequence argument valid but unsound. And, while (3a) is true, replacing (3) with (3a) would render the consequence argument invalid: (4) does not follow from (1), (2) and (3a). For (assuming determinism), since we <u>are</u> able to do some things such that were we to do them, an actual law would be false, and since our acts are indeed the consequences of true propositions about the past and the laws, we <u>are</u> able to render <u>some</u> true propositions about what we are going to do false. By (2), we are unable to render true propositions about the past false – I am unable to render false the true proposition that I had toast for breakfast this morning, for example. And by (3a) we are unable to do anything such that our acts would be or cause a law-breaking event. I am unable, for example, to run faster than light. But my inability to do either of those kinds of thing does not entail my inability to, say, raise my hand within the next minute – even if facts about the past and the laws entail that I will not do so. Hence the consequence argument fails to establish incompatibilism on either reading of premise (3).

Someone who accepts a Lewis-style analysis of counterfactuals is perfectly within her rights to disambiguate premise (3) in the way Lewis does. However, I shall argue that Lewis's disambiguation does not, in the end, succeed in undermining the consequence argument. In order for Lewis's strategy to succeed, what is needed is not just the mere fact that (3) equivocates between (3a) and (3b). In addition, the local miracle compatibilist needs to <u>motivate</u> the claim that, assuming determinism, (3a) is true while (3b) is false. As Fischer says:

> Whereas it is true that there is a gap in the incompatibilist's argument, the incompatibilist has at least shown that the compatibilist (of a certain sort) is committed to a distinction which might appear to be rather "fine". There is an incompleteness in the incompatibilist's argument, and thus it is <u>open</u> to a person to reject it, but in rejecting it, he may be committed to a distinction between claims about our abilities which is hard to explain and justify. (1988, 249)

I shall argue that the prospects for any such justification are extremely bleak: I cannot see any hope for motivating a position that accepts (3a) while denying (3b).

3. Three ways to break a law

The distinction between weak and strong abilities trades on a distinction between two different ways for a non-actual act <u>A</u> to break a law. <u>A</u> might be such that, were you to do it, a law would be broken, but not by <u>A</u> or by any of its effects (in which case, if you are able to <u>A</u>, that ability is weak). Or <u>A</u> might be such that, were you to do it, a law would be broken either by <u>A</u> itself or by some of its effects (in which case, if you are able to <u>A</u>, that ability is strong).

11

Examples of the latter kind of act that are typically given include doing something (raising one's hand, throwing a stone) faster than light, and building a machine that would cause a particle to violate the principle of the conservation of angular momentum.[8] Such events – or rather, events described in this way – wear their miraculous nature on their sleeve: in each case, the fact that the event would itself be or cause a law-breaking event can be read off merely from the way the event is described (together with knowledge of the actual laws). Indeed, Lewis's definition of an event's falsifying a proposition requires law-breaking events to have this feature. He says:

> Let us say that an event would falsify a proposition iff, necessarily, if that event occurs then that proposition is false. For instance, an event consisting of a stone's flying faster than light would falsify a law. So would an act of throwing in which my hand moves faster than light. So would a divergence miracle. (1981, 297)

We can extract from Lewis's definition the following definition of a law-breaking event:

(LBE1) Event $\underline{e}$ is a law-breaking event (relative to world $\underline{w}$) iff, necessarily, if $\underline{e}$ occurs then the proposition $\underline{L}$ stating the conjunction of $\underline{w}$'s laws of nature is false.

Examples like the stone's flying faster than light ($\underline{e}$) clearly satisfy this definition: given that it is a law at the actual world that nothing travels faster than light, $\underline{e}$ is a law-breaking event relative to the actual world. However, plenty of events that are manifestly law-breaking events (including plenty of divergence miracles, contrary to what Lewis says in the passage quoted above) fail to satisfy Lewis's definition. Recall the earlier example of my arriving at the pub at $\underline{t}$ (event $\underline{e}$). We saw that there is a possible world, $\underline{w_1}$, at which matters

of particular fact are the same as the actual world right up until t, whereupon I spontaneously disappear from my desk and reappear in the doorway of the pub. Clearly e, as it occurs at w_1, is, relative to the actual world, a law-breaking event. By contrast, e at w_2 – where 5 minutes before t, I decide to walk to the pub and then, several minutes later, arrive in an entirely non-miraculous fashion – is not a law-breaking event.

Lewis's definition fails to make e-at-w_1 a law-breaking event, because the conditions he imposes on law-breaking events are too strict: for Lewis, a law-breaking event (relative to the actual world) must be one such that necessarily, if it occurs, L (the conjunction of the actual laws) is false. So, by definition, e is a law-breaking event at one world at which it occurs if and only if it is a law-breaking event at all worlds at which it occurs. Events that are intuitively law-breaking events, however, need not be like that. Events like raising one's hand faster than light are like that – any world at which I do that will be a world at which L is false. But plenty aren't.[9]

What is needed is a definition of a law-breaking event that is sensitive to the circumstances that obtain, in worlds where the event occurs, when (or perhaps immediately before) the event occurs. Whether an event is a law-breaking event depends not just on the nature of the event itself, but also on the circumstances under which it occurs. What is needed, then, is the following definition:

(LBE2) Event e is a law-breaking event at world w_a, relative to world w_b, iff e, together with the circumstances under which it occurs at w_a, is incompatible with L, where L is the conjunction of all of w_b's laws.

In other words, e is a law-breaking event at w_a (relative to w_b) iff the proposition E that e occurs, and the proposition C specifying background circumstances at w_a around the

time of e's occurrence, together entail that L is false. [1011] (We need to add that C by itself does not entail that L is false, otherwise if background circumstances are themselves incompatible with L, any event occurring at the same time will trivially be a law-breaking event.)

We are now in a position to distinguish between the status of my arriving at the pub – e – at w₁, where I spontaneously appear there, and the status of e at w₂, where I walk there. Assuming determinism, and that e does not occur in the actual world, there will be a law-breaking event (relative to the actual world) at both w₁ and w₂. The difference is that at w₁, e itself is a law-breaking event, whereas at w₂, it isn't. At w₂, a law-breaking event occurs prior to e – neurons fire in my brain, say, in such a way as to get me to desire beer. Thereafter, matters at w₂ evolve in accordance with the actual laws of nature L. So, at the time when e occurs at w₂, e's occurrence together with background circumstances at w₂ are perfectly compatible with L.

Why is getting the definition of a law-breaking event right important? Well, obviously, since the local miracle compatibilist is committed to a distinction between weak and strong abilities, and since that distinction depends on the notion of a law-breaking event, one needs an adequate definition of that notion. But there is a more general point to be made here. Incompatibilists who argue for premise (3) of the consequence argument – the premise that says we are unable to render the laws of nature false – do so by appealing to examples of the faster-than-light-travel variety (see Fischer 1988, 245-6) – that is, to the kinds of law-breaking event that satisfy (LBE1). It does indeed seem implausible to suggest that we have the ability to perform such acts. Fischer notes that, given Lewis's disambiguation of premise (3) into (3a) and (3b), such examples do nothing to motivate (3b), the disambiguated premise

which Lewis claims to be false and which the incompatibilist needs to be true in order for the consequence argument to be sound. In failing to distinguish between different ways of rendering a law false (the differences that distinguish (3a) from (3b)), the incompatibilist illicitly makes the claim that we are able to render laws false (in any sense) seem more implausible than (according to local miracle compatibilists) it really is.

So far so good for the local miracle compatibilist. But, on the other hand, Lewis's (implied) definition of a law-breaking event, (LBE1), illicitly makes it seem more plausible that we are able to render laws false (in the appropriate sense) than it really is, for it suggests a neater division of non-actual actions into law-breaking and non-law-breaking varieties than there really is. Given (LBE1), it follows straight away that no action we ordinarily take ourselves to be able to perform – raising our hands, going to the pub, deciding on the salad rather than the lasagne – will be a law-breaking event, since manifestly no such action is an event such that, necessarily, if it occurs then an actual law is false. Given (LBE1), actions that we ordinarily take ourselves to have the ability to perform will <u>never</u> turn out to require us to have strong abilities, so the ban on strong abilities – one of the central theses of local miracle compatibilism – does not appear to restrict our freedom to do the sorts of things we ordinarily take ourselves to be free to do.

Given (LBE2), however, there is no immediate reason why <u>all</u> the mundane actions we ordinarily take ourselves to be able to perform (even though we do not in fact perform them) should turn out to be actions that would only require us to have weak abilities. For there need not be anything in the nature of the action itself that determines whether or not the ability to do it would be a weak ability or a strong ability. So the recognition that it is (LBE2), and not (LBE1), that provides the correct definition of 'law-breaking event' provides

15

at least prima facie reason to be sceptical about the local miracle compatibilist's claim that deterministic agents sometimes have weak abilities but never have strong abilities. In the next section I provide a counter-example that shows this scepticism to be well founded.

4. Why are strong abilities impossible?

In 'Are We Free to Break the Laws?', Lewis claims that his project is merely to show that (1) determinism is compatible with the ability to raise one's hand (when one does not in fact raise it), and (2) this ability does not entail the ability to perform acts that are or cause law-breaking events. It might seem, then, that Lewis's aim is merely to show that the following three claims are consistent, which is to say, all true together in at least one possible world:

(D)   Determinism is true

(W)   Agents (sometimes) have weak abilities: that is, they are sometimes able to perform an act A such that, were they to do so, an actual law would be broken – but not by A or by any of its effects.

(S)   Agents never have strong abilities: that is, they are never able to perform an act A such that, were they to do so, an actual law would be broken either by A or by some of its effects.

I do not doubt that there are indeed possible worlds in which (D), (W) and (S) are all true together. However, this is not enough to vindicate local miracle compatibilism. The problem (as I shall show) is that local miracle compatibilism cannot rule out the existence of possible worlds at which (S) is false: local miracle compatibilism does not provide us with any reason to suppose that it is impossible for agents at deterministic worlds to break the laws of nature – nor even any reason to suppose that, assuming determinism, we are unable to

16

break the laws. And it is this that both undermines Lewis's alleged refutation of the consequence argument, and renders local miracle compatibilism untenable.

Why is it such a problem if local miracle compatibilism cannot rule out the possibility that some (deterministic) agents have strong abilities? Well, if it cannot rule out such a possibility, then it is committed to the claim that it is merely a contingent matter whether or not agents are able to perform law-breaking acts. But it is very implausible to claim that this is a contingent matter. The view that agents never have strong abilities is a view according to which the laws of nature place absolute, inviolable constraints on what we are able to do. As such, it does not seem to me to be a claim about how our world happens to be, but a conceptual claim about the nature of laws. Why do laws place constraints on what we are able to do? Because that's just what laws <u>do</u>: not what they in fact do, but what they do as a matter of conceptual necessity.

To see that local miracle compatibilism cannot rule out the possibility of deterministic agents having strong abilities, consider the following example. Suppose that determinism is true, and suppose I am at a real estate auction. My opponent has made a bid slightly above the price I had decided would be my maximum bid, but I really want the house. I have to come to a quick decision about whether or not to raise my hand, since the auctioneer's hammer is about to fall. I decide not to raise my hand. Was I nonetheless able to raise it? It seems that the local miracle compatibilist ought to say yes: this is just the kind of ordinary ability that, according to local miracle compatibilism, deterministic agents generally possess. And, as we've seen, the divergence miracle – the law-breaking event – would not, in this case, have been my act of hand-raising itself, but rather some prior event. So my ability to raise my hand is a weak ability. So far so good.

According to Lewis, 'the divergence miracle [would not] have been my act of raising my hand. That act was altogether absent from the actual course of events, so it cannot get under way until there is already some divergence … Nor would it have been any other act of mine.' (1981, 294) I claim that Lewis is not entitled to this last claim: he gives us no reason to suppose that the divergence miracle required for me to raise my hand would not have been any act of mine.

Consider the (non-actual) event <u>M</u>, my <u>deciding</u> to raise my hand. There is no reason to suppose that the closest world at which I raise my hand cannot be a world where <u>M</u> is the divergence miracle, and hence a law-breaking event.[12] In that case, if I have the ability to do <u>M</u> then that ability is a strong ability, and there is no reason to suppose that this is an ability that I do not possess. Hence there is no reason to suppose that (S) is true. And, of course, if there is no reason to suppose that (S) is <u>in fact</u> true, then <u>a fortiori</u> there is no reason to suppose that (S) is true at all possible worlds.

That's the counter-example. It exploits a fairly straightforward feature of Lewis's analysis of counterfactuals. When we enquire about the nature of the closest possible world <u>w</u> at which some non-actual event <u>e</u> occurs, we usually find that matters of particular fact at <u>w</u> begin to diverge from the facts that obtain at the actual world at some time prior to the time of <u>e</u>. Roughly speaking, the less conducive the actual world's circumstances are, around the time in question, to <u>e</u>'s occurrence, the further back in time the divergence starts. If <u>e</u> is my arriving at the pub at <u>t</u>, given that at the actual world I am, at <u>t</u>, a good five-minute walk away, the divergence miracle at the closest <u>e</u>-world will happen several minutes before <u>t</u>: the loss of a few minutes' perfect match of matters of particular fact is compensated by the fact that, given that loss, we only need a small miracle rather than a big one to get <u>e</u> to happen.

18

But it cannot be the case that for every non-actual event e, the divergence miracle at the closest e-world occurs before the time of e. For divergence miracles are themselves events, and, by definition, the closest world where a divergence miracle f occurs is a world whose facts do not diverge from those of the actual world until f itself occurs. Lewis's claim that deterministic agents never have strong abilities entails that divergence miracles can never be acts. But there is no reason to suppose that this claim is true.

There are some possible escape routes available to the local miracle compatibilist, and they need to be plugged. There are two basic strategies that the local miracle compatibilist might employ in order to block the conclusion that I have (or at least that it is metaphysically possible that I have) the strong ability to decide to raise my hand (M). One strategy is to deny that my ability to M is really a strong ability. The other is to concede that if I have that ability, it is indeed be a strong ability – but to deny that the ability to M is an ability that I could have. I shall consider the two strategies in turn in the following two subsections.

Is the ability to M really a strong ability?

The first strategy – denying that the ability to M (to decide to raise my hand) is, or could be, a strong ability – is suggested by two things Lewis himself says in 'Are We Free to Break the Laws?'. The first is the passage quoted above: 'the divergence miracle [would not] have been my act of raising my hand. That act was altogether absent from the actual course of events, so it cannot get under way until there is already some divergence … Nor would it have been any other act of mine' (1981, 294). Lewis says that the divergence miracle could not be the raising of his hand, since, as that event was absent from the actual course of events, a divergence miracle must already have occurred. One might attempt to object that the same

holds of <u>any</u> act that I do not in fact perform, including <u>M</u>: since <u>M</u> is absent from the actual course of events, it, like Lewis's hand-raising, cannot get underway until there is already some divergence. Hence my ability to <u>M</u> must be a weak ability rather than a strong one.

But this response cannot possibly be right, because it leads to a vicious regress. Suppose we grant that the response is correct: that there must indeed be a divergence miracle prior to <u>M</u>. But what goes for <u>M</u> goes for the alleged divergence miracle too: since <u>that</u> event is (by definition) absent from the actual course of events, there must be some divergence miracle <u>N</u> prior to <u>its</u> occurrence, to which, of course, we could then apply the same reasoning, <u>ad infinitum</u>. Indeed, if this line of reasoning were correct, divergence miracles would be impossible: for any alleged divergence miracle, there must be a prior divergence miracle – in which case our alleged divergence miracle would not <u>be</u> a divergence miracle, but would merely be the (non-miraculous) result of some prior divergence miracle.

The second way of denying that my ability to <u>M</u> is a strong ability runs as follows. I have taken it for granted that there is a unique closest possible world in which I do <u>M</u> (decide to raise my hand, say), where <u>M</u> is itself a law-breaking event. But this assumption, it may be argued, is false: as Lewis says, 'there is no particular divergence miracle that definitely would have occurred, since the divergence might have happened in various ways' (1981, 295). Since there is no divergence miracle that definitely would have occurred – that is, since there is a tie for the closest <u>M</u>-worlds, and since <u>M</u> is not a divergence miracle in all such worlds – it is false that, had I done <u>M</u>, I would have performed a law-breaking act. Hence the ability to do <u>M</u> does not count as a strong ability after all. (Perhaps it counts as an ability that is neither determinately strong nor determinately weak.)

Unfortunately, however, Lewis goes on to note that whether or not there is a tie for the closest M-world is a contingent matter: 'we can imagine a world of discrete processes at which one divergent history in which I raise my hand clearly takes less of a miracle than any of its rivals.' (1981, 295, footnote 4). And, of course, what goes for Lewis's raising his hand goes equally for my doing M. I claim that there is a world w of 'discrete processes' – indeed, a world which, for all we know, could be the actual world – that is so arranged that the unique closest world to w where I do M is a world where M itself is the divergence miracle and hence a law-breaking event. Local miracle compatibilism simply does not have the resources available to deny that there are possible worlds in which my ability to M is a strong ability.

Is the (strong) ability to M an ability I could have?

The second strategy that the local miracle compatibilist might try is to agree that the closest world at which I do M is a world where M is a law-breaking event, but to deny that the ability to M is an ability that I could possibly have. What is needed here is a conception of ability that explains why agents cannot have strong abilities. Lewis himself provides no such conception: he simply asserts that we are unable to perform law-breaking acts. ('A marvellous power indeed! Can you also bend spoons?' (1981, 292)) As mentioned in section 3, this claim might not seem to stand in need of justification if all law-breaking acts were of the faster-than-light travel variety – perhaps we could just appeal to common sense intuition to settle the question. But once it is recognised that law-breaking acts can easily be of the more mundane variety – the sorts of act that we ordinarily take ourselves to be able to perform – appeal to common sense intuition is unpersuasive.[13] We therefore need to look

elsewhere for a conception of 'ability' that provides us with a good reason to think that agents cannot have strong abilities.

A compatibilist conception of 'ability' is offered by Michael Smith (1997), who says: … what exactly does it mean to say that you were able to think of a better response to an argument than the one you in fact thought of? Does your possession of that ability require anything weird or transcendental of you? No it does not. Indeed, … we can spell out the meaning of this claim in terms of possible worlds. To say that you were able to think of a better response … means, <u>inter alia</u>, that the possible worlds in which you think of the better response are near by, or very similar to, the actual world in which you don't … By contrast, when we say that you thought of the very best response that you could, and so weren't able to think of a better response, what we mean is <u>inter alia</u> that the possible worlds in which you think of a better response are remote from, or rather dissimilar to, the actual world. (1997, 300)

Smith's suggestion, then, is that the issue of which abilities we have on a given occasion is to be decided on the basis of how far away from the actual world the worlds in which we do otherwise are. Given this analysis of 'ability', it is open to the local miracle compatibilist to make the following claim: if <u>A</u> would, were I to do it, be a law-breaking event, then the closest world where I do <u>A</u> will be too remote from the actual world to make it true that I am able to <u>A</u>.[14]

At first sight, this seems like a promising suggestion since, intuitively, the closest world where I raise my hand isn't very far away from the actual world, while the closest world where I raise it faster than light is considerably further away. But, unfortunately, the suggestion seems less plausible when we consider the case in hand, namely the question of

whether I am able to <u>M</u> (decide to raise my hand). My claim was that the closest world <u>where I raise my hand</u> might easily be a world at which <u>M</u> is the divergence miracle. But in that case, the closest world where I raise my hand just <u>is</u> a world where I perform a law-breaking act, namely <u>M</u>. So, given Smith's conception of ability, I cannot have the ability to raise my hand unless I also have the ability to do <u>M</u>, since the closest world where I do the former is the same distance from the actual world as the closest world where I do the latter.

More generally, assuming determinism, the closest world where I perform non-actual act <u>A</u> – where <u>A</u> is a perfectly ordinary act like raising my hand or choosing the soup instead of the salad – will contain a small local miracle. Granted that the ability to do <u>M</u> counts as a strong ability, the closest <u>M</u>-world is, similarly, a world that contains a small local miracle, namely <u>M</u> itself. There is no reason whatever to think that this miracle need be any bigger than the miracle that occurs in the closest world where I, say, choose something different from the menu or raise my hand or do anything else that I ordinarily take myself to be able to do.

What grounds the distinction between weak and strong abilities is not the <u>size</u> of the miracle required in order to get my closest-possible-world counterpart to perform the act in question, but whether or not that miracle is (or is caused by) an act of mine. Since whether the required miracle is an act of mine does not itself play any part in how similar the relevant possible world is to the actual world, it is not surprising that Smith's conception of 'ability' does not rule out the possibility that some agents at some deterministic worlds have strong abilities.

Local miracle compatibilism fails to rule out the metaphysical possibility of deterministic agents (agents who are very much like us, and who may indeed, for all we

know, be us) having strong abilities. This refutes local miracle compatibilism, granted the assumption that (S) – the claim that agents are never able to perform law-breaking acts – is a necessary truth.

5. The prospects for compatibilism

The driving force behind incompatibilism is the idea that the laws of nature constrain our abilities. The laws are inviolable, not just in the rather trivial sense that it is conceptually necessary that no law of nature is ever falsified, but in a more robust, metaphysically substantive sense: they provide what Simon Blackburn calls a 'straitjacket'[15] – something 'on which the regular course and succession of objects totally depends', as Hume puts it.[16] On such a view of laws, of course, the laws constrain not just what we agents are able to do, but also what goes on outside the realm of human action: they constrain which neurons fire, how fast stones travel, how fundamental particles behave, and so on. I'll call this view of laws the 'necessitarian' view.[17]

The necessitarian view of laws, however, is not mandatory. According to a competing view – the Humean view – laws are merely facts about regularities. (On a naïve regularity view, the laws are all and only the regularities. On a more sophisticated view – the Ramsey-Lewis view – the laws are a subset of the regularities: roughly speaking, the laws are the deepest, most pervasive, most explanatory regularities.) On the Humean view, laws place no constraints whatever on what happens. Laws are not that 'on which the regular course and succession of objects totally depends'. Quite the reverse, in fact: it is the regular course and succession of objects that determines what the laws are.[18]

A broadly Humean view of laws has been thought by some compatibilists to guarantee deterministic agents the kind of freedom which, according to incompatibilists, determinism deprives them of. A.J. Ayer, for example, says:

> But if all that is meant [by determinism] is that it is possible, in principle, to deduce [facts about the future] from a set of particular facts about the past, together with the appropriate general laws, then, even if this is true, it does not in the least entail that I am the helpless prisoner of fate … What it does entail is that my behaviour can be predicted: but to say that my behaviour can be predicted is not to say that I am acting under constraint. (1954, 284)[19]

For all their differences, the Humean and necessitarian views share one crucial feature: according to each view, laws constrain the actions of agents to precisely the same extent as they constrain other events: on neither view do the laws possess a kind of efficacy, or power to constrain, or jurisdiction, over actions that they do not possess over other events (or vice versa). The underlying problem with local miracle compatibilism, I believe, is that it embodies an inherently unstable attitude towards laws of nature, since is requires that laws have a special kind of jurisdiction over events that are acts, which they do not have over events that are not acts.

Consider first the local miracle compatibilist's claim that deterministic agents sometimes have weak abilities – that is, are sometimes able to things such that, were they to do them, a law of nature would be broken (but not by the action itself). This is a claim that looks plausible from a Humean perspective: if laws of nature are simply regularities, then the violation of a law (that is, the violation, in some nearby possible world, of an actual law of

nature) isn't such a big deal. We happen to live in a world that is such that one set of regularities obtain; different sets of regularities obtain at other possible worlds. These are brute facts about the worlds in question, that are not to be explained by reference to some extra thing upon which those regularities depend. So there is no reason to think that the fact that my raising my hand would require a violation of the actual laws compromises my ability to do it.

From a necessitarian perspective, however, the fact that my raising my hand would require a violation of the actual laws does seem to provide sufficient reason to think that I am unable to do it. If the laws <u>constrain</u> what happens in the actual world, and determinism is true, it really does seem that I am constrained to do what I actually do, and am thereby constrained not to raise my hand. Thus the claim that deterministic agents have weak abilities goes naturally with, and only with, a Humean view of laws.

The opposite is true of the other half of the local miracle compatibilist creed, the denial that agents ever have strong abilities (that is, the ability to do something such that, were you to do it, a law would be broken by your act itself, or by some of its effects). The necessitarian will have no difficulty in agreeing with the local miracle compatibilist about this. But from a Humean perspective it is very hard to see what reason there could be to believe it. If miracles are relatively cheap when it comes to events that are <u>not</u> actions of mine, or effects of actions of mine – as they must be if we are to hold that deterministic agents have weak abilities – why should they not be equally cheap when it comes to events that <u>are</u> actions of mine? Actions are, after all, merely a species of event, and the laws of nature manifestly do not care about the distinction between events that are and are not actions. So how could the laws of nature constrain my abilities in the sense that they render

me unable to violate them <u>by my acts</u>, yet <u>not</u> constrain my abilities in the sense that they do <u>not</u> render me unable to perform acts that <u>require</u> them to be violated?

In short, local miracle compatibilism needs a Humean view of laws to motivate the claim that agents have weak abilities, and a necessitarian view of laws in order to motivate the claim that agents never have strong abilities. That is why I said that local miracle compatibilism embodies an unstable attitude towards laws of nature, and that, in the end, is why it fails to be a viable version of compatibilism.

Where does all this leave the consequence argument? Well, recall Lewis's claim that premise (3) of the argument can be disambiguated into

(3a)   I am unable to do anything such that my act would be or cause a law-breaking event

and

(3b)   I am unable to do anything such that were I to do it, an actual law of nature would be false (and hence not a law),

and that (assuming determinism) while (3a) is true (on which reading the consequence argument is invalid), (3b) – the premise that makes the argument valid – is false. What I hope to have shown is that this claim lacks motivation. Lewis sketches a position according to which determinism is compatible with the denial of (3b). But that position does not satisfactorily rule out the claim that determinism is likewise compatible with the denial of (3a). Once we grant that a deterministic agent is sometimes able to do something such that, were he to do it, a law would be broken, there is no reason to think that he is never able to do something that would itself be a law-breaking event.

I conclude that the consequence argument is not severely damaged by Lewis's criticism. Read the disputed premise as 'we are unable to render false, in any sense, propositions stating the laws of nature' – or, equivalently, 'we are unable to do anything that would require a violation of the laws of nature'. Then the argument is valid. The compatibilist cannot respond merely by insisting that many agents in deterministic worlds are perfectly well able to raise their hands even when they do not in fact do so. That response amounts simply to a denial of the conclusion of the consequence argument, not a well-motivated reason to reject the above premise.

Given the validity of the consequence argument thus construed, and given the implausibility of the combination of (3a) and the denial of (3b), the compatibilist has two clear options. One is to accept (3b), and to claim that free action does not require the ability to do otherwise. This line has been taken by Fischer, who calls the position 'semicompatibilism'.[20] The other option is to deny both (3a) and (3b), and thus to hold that deterministic agents are able to perform law-breaking actions. As I remarked earlier, a Humean conception of laws has been put to use by compatibilists in the past. But they have not, to my knowledge, embraced or defended the obvious consequence of their view of laws – namely, the denial of (3a). Part of the point of this paper has been to show that that consequence must be embraced by someone who wishes to deny (3b): one cannot plausibly maintain that we have weak but not strong abilities. To incompatibilist ears, the claim that agents have strong abilities will sound like the first line of a rather obvious reductio of compatibilism. That is a charge that the compatibilist needs to rebut rather than dodge.[21, 22]

Bibliography

Aune, B. (1967): "Hypotheticals and 'Can': Another Look", Analysis 27, 191-195

Austin, J.L. (1961): "Ifs and Cans", in his Philosophical Papers (Oxford: OUP)

Ayer, A.J. (1954): "Freedom and Necessity", in his Philosophical Essays (London:
    Macmillan)

Beebee, H. (2000): "The Non-Governing Conception of Laws of Nature", Philosophy and
    Phenomenological Research 56, 571-594

Beebee, H. & Alfred R. Mele (2002): "Humean Compatibilism", Mind 111, 201-223

Blackburn, S. (1993): "Hume and Thick Connexions", in his Essays in Quasi-Realism (New
    York: OUP)

Fischer, J.M. (1988): "Freedom and Miracles", Nous 22, 235-252

Fischer, J.M. (1994): The Metaphysics of Free Will (Oxford: Blackwell)

Horgan, T. (1984): "Compatibilism and the Consequence Argument", Philosophical Studies
    47, 339-356.

Hume, D. (1975): Enquiries Concerning Human Understanding and Concerning the
    Principles of Morals, ed. L.A. Selby-Bigge (Oxford: OUP)

Lehrer, K. (1966): "An Empirical Disproof of Determinism?" in K. Lehrer, ed., Freedom and
    Determinism (New York: Random House).

Lehrer, K. (1968): "Cans Without Ifs", Analysis 29, 29-32.

Lewis, D.K. (1979): "Counterfactual Dependence and Time's Arrow", Nous 13, 455-476.

Lewis, D.K. (1981): "Are We Free to Break the Laws?" Theoria 47, 113-121, reprinted in his
    1986.

Lewis, D.K. (1986): Philosophical Papers, volume II (New York: OUP)

Lewis, D.K. (1986a): "Postscripts to 'Counterfactual Dependence and Time's Arrow'", in his
      1986, 52-66.

Loewer, B. (1996): "Humean Supervenience", <u>Philosophical Topics </u>24, 101-127.

Moore, G.E. (1912): <u>Ethics</u> (London: O.U.P.).

Smith, M. (1997): "A Theory of Freedom and Responsibility", in G. Cullity and B. Gaut,
      eds., <u>Ethics and Practical Reason</u> (Oxford: Clarendon Press).

Swartz, N. (1985): <u>The Concept of Physical Law</u>. New York: Cambridge University Press.

van Inwagen, P. (1975): "The Incompatibility of Free Will and Determinism", <u>Philosophical</u>
      <u>Studies </u>27, 185-199

van Inwagen, P. (1983): <u>An Essay on Free Will</u>. Oxford: Clarendon Press.

<u>Endnotes</u>

[1] For exposition and defence of conditional analyses, see for example G.E. Moore 1912,

Chapter 6; A.J. Ayer 1954, 282; Bruce Aune 1967. For objections, see J.L. Austin 1961;

Keith Lehrer 1966 and 1968.

[2] See Fischer 1994, Chapter 4

[3] See Lewis 1986a, 55-56, for more on big and little miracles.

[4] Lewis himself does not frame the issue in terms of weak and strong abilities, but rather in

terms of a 'weak thesis' ('I am able to do something such that, if I did it, a law would be

broken') and a 'strong thesis' ('I am able to break a law', that is, 'I am able to do something

such that, if I did it, my act either would cause or would be a … law-breaking event') – see

Lewis 1981, 293-294. The strong thesis entails the weak thesis. By contrast, the possession of

a strong ability does not entail the possession of a weak ability. Weak and strong abilities are mutually exclusive: if you have the ability to A on a particular occasion, that ability will be either weak or strong, but not both.

[5] This is a consequence of Lewis's definition of what it is for an event to falsify a proposition (see his 1981, 297). That definition is rejected in section 4 below, but it will suffice for present purposes.

[6] Versions of the argument appear in van Inwagen 1975 and 1983 (Chapter 3).

[7] (3b) amounts to this claim, and not just to the weaker claim that I do not have any weak abilities, since (3b) implies (3a), whereas the lack of weak abilities does not entail the lack of strong abilities.

[8] See for example van Inwagen 1983, Chapter 3 and Fischer 1994, Chapter 4.

[9] Note that Lewis's definition of a law-breaking event is incompatible with his claim that divergence miracles are (always) law-breaking events, for there is no reason why a divergence miracle should be of the faster-than-light variety rather than the less exotic-sounding miracle-in-the-circumstances kind.

[10] In his 1984, Terence Horgan, who is broadly sympathetic to Lewis's response to the consequence argument, also notes that Lewis's definition is too strong, and proposes a different alternative definition of 'event e falsifies proposition P'. However, as he notes (349-350), his definition does not distinguish between what I have called weak and strong abilities. So Horgan's definition will not serve any compatibilist attempt to motivate a distinction between the two kinds of ability.

[11] One might try to save (LBE1) by building the background circumstances into the event itself, so that we can distinguish the event e (which happens to occur in circumstances C) from the event e_c, which is e's-occurring-in-circumstances-C. Then for any event e that satisfies (LBE2) but not (LBE1), there will be an event e_c which satisfies (LBE1); hence (LBE1) is not inadequate after all. However, the focus of this paper is on whether or not law-breaking events could be <u>actions</u> that we are able to perform. The kind of event that renders (LBE2) unnecessary is not the kind of event for which this issue can arise. My raising-my-arm-in-cricumstances-C (three feet from a light bulb, say, or in the presence of oxygen) might count as an event, on some suitably relaxed conception of events – but it is not an <u>action</u> of mine: only the arm-raising itself is an action.

[12] Of course, this is not to say that $\underline{M}$ is a law-breaking event at <u>all</u> worlds at which it occurs. There are plenty of worlds where it isn't – for example, a world where, thanks to a prior divergence miracle, I am richer, and thus have more money to spend on buying a house, and consequently have no reason not to put in a higher bid. I am claiming only that the <u>closest $\underline{M}$-</u>world might easily be a world where $\underline{M}$ itself is the divergence miracle.

[13] Spoon bending, of course, is an in-the-circumstances kind of law-breaking event, rather than the faster-than-light travel kind. In the right circumstances, or with the right kind of spoon, there is nothing remotely marvellous about the ability to bend spoons.

[14] This was suggested to me by Michael Smith.

[15] See Blackburn 1993, especially 97-99.

[16] Hume 1975, 55 (quoted by Blackburn 1993, 98).

[17] For a classic statement of the necessitarian view, see Armstrong 1983, especially Chapters 6 and 11.

[18] See Loewer 1996 and Beebee 2000.

[19] For a similar compatibilist line, see Swartz 1985.

[20] See Fischer 1994, Chapters 7 and 8.

[21] The implications of a Humean view of laws for the consequence argument are discussed in Beebee & Mele (2002).

[22] Many thanks to Michael Smith, Al Mele and Julian Dodd. Thanks also to audiences in Bristol and Sussex, especially Jessica Brown, James Ladyman, and Murali Ramachandran.