

Nihil Obstat: Lewis's Compatibilist Account of Abilities

Helen Beebee, Maria Svedberg, and Ann Whittle*

ABSTRACT

In an outline of a paper found amongst his philosophical papers and correspondence after his untimely death in 2001—“*Nihil Obstat*: An Analysis of Ability,” reproduced in this volume—David Lewis sketched a new compatibilist account of abilities, according to which someone is able to *A* if and only if there is no obstacle to their *A*-ing, where an obstacle is a ‘robust preventer’ of their *A*-ing. In this paper, we provide some background context for Lewis’s outline, a section-by-section commentary, and a general discussion of the account’s main features.

INTRODUCTION

In his paper “Are We Free to Break the Laws?” (Lewis 1981), David Lewis famously argues that the ability to do otherwise is compatible with determinism. Suppose a judge just put her hand down on her desk. Was she able to do something else instead, for instance, to raise her hand? Provided that she was not tied up, paralyzed, hypnotized, brainwashed, or anything like that, it seems natural to think that she was. Suppose further that determinism is true. In that case everything that happens is completely determined by the remote past together with the laws of nature. This means that if the judge was able to raise her hand, then she was able to do something such that had she done it, the remote past or the laws of nature would have been different. A certain type of argument for incompatibilism—and in particular Peter van Inwagen’s Consequence Argument (Inwagen 1975, 1983)—builds on the assumption that no one is ever able to perform acts like that.

In response to the Consequence Argument, Lewis (1981) defends the idea that agents in a deterministic universe are sometimes able to perform an act such that had they done it, the laws of nature would have been a little different. Our judge, for instance, was able to raise her hand and if she had done so, then something that is in fact a law of nature would have been broken and thus no law. But the law-breaking event would not have been her raising her hand. Had she raised her hand, a law would have been broken beforehand by a so-called divergence miracle. And this

*HB: University of Manchester; MS: Uppsala University; AW: University of Manchester

divergence miracle would not have been any act of hers. Nor would it have been caused by any act of hers. So while she was able to act so that a law would have been broken, she was not able to break a law. And this was Lewis's aim in that paper: to show that compatibilists do not have to accept the thesis that deterministic agents are sometimes able to *break* the laws of nature.

What Lewis did not attempt there, however, was to provide a positive account of what it *is* for a deterministic agent to be able to do otherwise. He simply assumed that deterministic agents do indeed sometimes have that ability, and argued that this ability does not amount to the ability to perform (or cause) any event that constitutes a violation of the laws of nature.

In the Fall semester of 2000, Lewis wrote an outline—which appears in this volume (Lewis 2020)—of a paper, “*Nihil Obstat: An Analysis of Ability*.” His death in October 2001 unfortunately prevented him from writing the paper up, but its purpose is clear: it is an attempt to make good on the assumption in “Are We Free to Break the Laws?” that deterministic agents are sometimes able to act otherwise by providing—as he puts it—“a compatibilist analysis of ability.” In §1 he offers a “simple proof of compatibilism,” thus showing why a compatibilist account of abilities is needed. In §2 he presents various reasons why the traditional ‘conditional analysis’ of abilities is unsatisfactory. In §§3 and 4 he provides his own analysis of abilities. Very crudely, according to that account *S* is able to *A* iff there are no “obstacles” to *A*-ing, where an obstacle is a “robust preventer”: a preventer (that is, something that would cause *S* not to *A*) that “would not go away if things were just a little different.” In §5, he discusses “removable” obstacles, and argues—in effect—that our ambivalence about whether or not *S* is able to *A* in a range of cases (e.g., where *S* lacks the skill to *A*) is due to our ambivalence about whether or not to count removable obstacles as obstacles to *A*-ing. Finally, in §6 he argues that preferring not to *A* is sometimes an obstacle to *A*-ing, and sometimes not.

In §§1–6 below, we go through Lewis's outline section by section, highlighting some of his more interesting moves, attempting to flesh out what he may have had in mind at various points in the outline where his intentions are not entirely clear, and flagging up some interpretive difficulties and potential problems with the position he sketches. (We recommend that the reader keep the outline [Lewis 2020] to hand in reading these sections, since we do not reproduce it in full below.) In §7 we turn to a more general discussion of the view Lewis presents in the outline, picking up on some of our earlier remarks.

It is unfortunately impossible to say how closely related to the outline a final version of “*Nihil Obstat*” would have been. Lewis did not, apparently, present a version of the paper anywhere (he tended to present from fully written papers, and no such version of this paper exists); nor did he discuss it in his correspondence, aside from briefly summarising it in a letter to Daniel Dennett.¹ The interpretative problems would undoubtedly have been resolved in a final version of the paper, but it is of course hard to say what Lewis would have done in connection to the substantive problems with the views expressed in the outline that we flag up.

1. A SIMPLE PROOF OF COMPATIBILISM

This argument has not, to the best of our knowledge, been presented by anyone in the existing literature. It is an argument for compatibilism in a broad sense since the conclusion is that free will is compatible both with determinism and with indeterminism. It can be reconstructed the following way:

1. It's a Moorean fact that we often have a choice what to do. (That is, we're often able to do what we actually do, but also what we don't do.)
2. We do not know whether our world is deterministic or indeterministic.
3. If (1) and (2), then having a free choice is *epistemically* compatible both with determinism and with indeterminism.
4. If having a free choice is *epistemically* compatible both with determinism and with indeterminism, then having a free choice is compatible *simpliciter* with both.
5. Thus, having a free choice is compatible *simpliciter* both with determinism and indeterminism.

We take it that A and B are compatible *simpliciter* if and only if there is some possible world where A and B are both true, and that A and B are *epistemically* compatible if and only if *for all we know* there is some possible world where A and B are both true. Thus, where F is 'agents often have a free choice', D is 'determinism is true', and I is 'indeterminism is true', the argument first establishes that for all we know, there is a possible world where $(F \& D)$ is true (indeed, for all we know, the actual world is just such a world), and likewise for $(F \& I)$; and then concludes on that basis that, for each of $(F \& D)$ and $(F \& I)$, there really is some possible world where that conjunction is true.

The contentious premises are (1) and (4). Let's take them in turn. Premise (1) repeats a claim that Lewis makes in a paper written a couple of years earlier, "Void and Object," where he says that it is a "Moorean fact that we make free choices" (2004, 79). Elsewhere, he characterises a Moorean fact as "one of those things that we know better than we know the premises of any philosophical argument to the contrary" (1996, 549) and "one of the many facts that even philosophers should not deny, whatever account or analysis they give of such facts" (1983, 351). In Lewis's view, then, the only philosophical views that remain on the table given this constraint are compatibilism and libertarianism: hard determinists and impossibilists about free will deny what "philosophers should not deny." This claim, of course, is controversial—but we shall not discuss it further.

Premise (4) is an inference from the epistemic possibility to the possibility *simpliciter* of both $(F \& D)$ and $(F \& I)$. In support of this premise, Lewis claims that none of the "alleged" counterexamples to this kind of inference—cases of epistemic possibility that are not possible *simpliciter*—are analogous to the free will case. It is not immediately clear, however, exactly *why* he thinks these cases are disanalogous. In each of the "alleged precedents" that Lewis lists, we have (allegedly) a necessary truth M that we do not (or at least some people do not) know; hence both M and $\sim M$ are

epistemically possible (for such people) despite one of them ($\sim M$) being impossible: there is no possible world at which it is true. (We assume that Lewis's example of the "geography of the pluriverse" refers to facts about the space of possible worlds that, if true, are necessarily true. In a letter to Brian Skyrms from 1975, he says: "Modal questions appear, and we don't know the answers, and we don't even know how to find out. Yet sometimes [when the answers don't vary from one world to another] the answers are necessary. For instance, the answer to the question how many worlds there are is a necessary truth, but it's a necessary truth which I don't know.")²

In the case of free will, too, there are two mutually exclusive epistemic possibilities: (a) agents have a free choice and determinism is true; and (b) agents have a free choice and indeterminism is true. Unlike the "alleged precedents," neither of these epistemic possibilities is a necessary truth. It does not follow, however, that neither (a) nor (b) is impossible; after all, the incompatibilist will insist that (a) is indeed impossible: there are no possible worlds where agents have a free choice and determinism is true. So why is it that the "alleged precedents are irrelevant to this case"?

We suggest that Lewis's answer is as follows. The cases he lists as alleged precedents are all cases where we know that some proposition—either M or $\sim M$ —is impossible *simpliciter*, but we don't know which it is (which is why both M and $\sim M$ are epistemically possible). The free-will case is not like that; it isn't as though we know that either (a) or (b) is impossible but we don't know which it is. Granted, incompatibilists think they have an argument (such as the Consequence Argument) for thinking that (a) is impossible *simpliciter*; but to take such an argument to be sound, as incompatibilists do, establishes (by one's own lights) not just the impossibility *simpliciter* but also the *epistemic* impossibility of (a): once we have an argument that establishes, to our own satisfaction, that p , $\sim p$ becomes an epistemic impossibility for us. But Lewis already takes himself to have established that (a) is epistemically possible; and if we grant him that, we cannot *also* coherently claim that (a) is metaphysically impossible. That would be akin to claiming both that there is a convincing proof of the continuum hypothesis *and* that it is nevertheless epistemically possible that the continuum hypothesis is false—and that circle cannot be squared, since that epistemic possibility would, precisely, be ruled out by a convincing proof of the continuum hypothesis.

The conclusion of the argument of this first section, then, is that having a free choice is compatible both with determinism and with indeterminism. Lewis ends the section by saying: "So we need a compatibilist analysis of ability." But in what sense do we 'need' such an analysis? In a sense, no such analysis is needed to shore up his argument for compatibilism; as we've seen, that argument depends on the Moorean fact that we are able to do otherwise, which Lewis takes to require no justification (after all, that's the point of Moorean facts). On the other hand, asserting that it is a Moorean fact that p does nothing to dispel the puzzlement that arises from the convincing-looking arguments whose conclusion is $\sim p$ that have led some philosophers to deny p .

Compare Lewis's approach in "Elusive Knowledge" (Lewis 1996): he does not think that merely asserting that it is a Moorean fact that I know a lot, e.g., that I have

hands, constitutes a fully adequate response to the sceptic. The Moorean fact establishes that scepticism is false, but a puzzle remains: how can it be that we know a lot despite being unable to rule out sceptical scenarios? To resolve the puzzle, Lewis needs a theory of knowledge. The case of the ability to do otherwise, we suggest, is analogous: Lewis (1981) has already shown that the Consequence Argument fails to undermine the Moorean fact that we are sometimes able to do otherwise, the epistemic possibility that the world is deterministic notwithstanding. But a puzzle remains: how can it be that we are able to do otherwise even if determinism is true? To resolve that puzzle, Lewis needs a compatibilist theory of abilities.

2. THE CONDITIONAL ANALYSIS IS UNSATISFACTORY

A simple conditional analysis of ability was popular among compatibilists up until the 1960s.³ Due to various counterexamples, some of which Lewis mentions here, it is nowadays universally dismissed. Nonetheless, Lewis lists four objections; we'll expand briefly on the first three since the final objection is already covered in the literature (for discussion see Vihvelin [2013, 201–208]).

The first two objections to the conditional analysis concern finkish dispositions. A finkish disposition is a disposition that something has but, were it to be placed in circumstances that would normally trigger a manifestation of the disposition, something would prevent it from manifesting (Lewis 1997; Bird 1998; Choi and Fara 2018). Lewis's stammering example of a finkish disposition to succeed is supposed to fit this model: normally when one has a disposition to *A*, trying to *A* is likely to result in *A*-ing. In the stammering case, this is not so: indeed, trying, or rather trying "too hard" (as opposed to just *A*-ing spontaneously), will prevent one from *A*-ing. Thus, Lewis is claiming, the stammerer does indeed have the ability to say "hello" (or whatever), but it is not true that they would say "hello" if they tried (or chose) to. Hence finkish dispositions to succeed are counterexamples to the conditional analysis.

Lewis's example of a finkish disposition *not* to succeed is rather curious, since in Frankfurt's original case (Frankfurt 1969) Jones—the victim of the neurosurgeon—is not such that if he tried, he would gain an ability he now lacks. On the contrary: assuming that he does have the disposition to decide otherwise than the neurosurgeon wants him to decide (see [Fara 2008]), the neurosurgeon finks (or masks—see Fara [2008, 894, n.15]) that disposition by preventing it from manifesting. So the original Frankfurt case looks like a case of a finkish disposition to succeed, and not—as Lewis defines it—a finkish disposition not to succeed. It may be, however, that Lewis had a different kind of Frankfurt-style case in mind, akin to Neil Levy's *Phobia* case (Levy 2008, 234), where the neurosurgeon is ready to step in and enable Jillian to conquer her pathological fear of deep water should she try to rescue the drowning child, so that, despite her phobia, 'if Jillian were to try, she would succeed' is true. Lewis's implied view is that Jillian lacks the ability to rescue the child (since finkish dispositions not to succeed are supposed to be counterexamples to the conditional analysis of ability) and this is controversial; Levy himself holds that the "presence of the counterfactual intervener has given her a capacity that she would otherwise have lacked" (2008, 234).

The third objection to the conditional analysis concerns luck. According to Lewis, an accomplished Australian Rules footballer like Lloyd is able to kick a goal even though he sometimes tries but fails because he is unlucky (a ‘behind’ is a near miss). On the occasions when he turns out to fail, he still had the ability—even though it is false on those occasions that if he tried to kick a goal (which he did) he would succeed (which he didn’t). Nor (we take Lewis to be saying) can we amend the conditional analysis so that the requirement on ability is that *S* would *probably* succeed if they tried (or chose or whatever), since one “might have the ability to succeed even if the chance of success is low, if others’ chance was much lower still.” So a not-very-good amateur footballer with a much lower success rate than Lloyd (Dave, say) is still able to score, since—for example—their chance of success is still much higher than Lewis’s.

Note that Lewis here is claiming that even if he has a “minute chance of scoring a goal,” he is “not able to do it.” One might be tempted to take this to imply that, while someone (e.g., Dave) can have an ability despite having a low chance of success, there is a lower limit: nobody who has only a “minute” chance of *A*-ing is able to *A*. We suggest, however, that this is not what Lewis is claiming. Rather, the claim that Lewis is not able to score should be read in contextualist terms. The question at issue here is whether *Dave* is able to score, to which Lewis’s answer is ‘yes’; and we suggest that the reason why the answer, for Lewis, is ‘yes’ is that Dave (despite his low chance of success) compares favourably on the ability front to Lewis. In that context, Lewis is indeed unable to score a goal; but there may well be other contexts in which Lewis *is* able to score. Case 1: you’re playing football in the park, and you can pass the ball to Lewis or to Dave. Who do you choose? No contest: Lewis can’t score for toffee! In other words, he isn’t able to. Case 2: you can pass the ball to Lewis or a two-year-old. Who do you choose? Lewis, of course. He’s a terrible footballer, but he is—just about—able to do it, even though he’s very unlikely to succeed; and in particular in this context there is a relevant party, the two-year-old, whose chance of success is “much lower still” and hence who is unable to score.

Some evidence that Lewis has contextualist inclinations when it comes to abilities comes from “The Paradoxes of Time Travel.” In his discussion of whether or in what sense Tim can kill his own grandfather, he says:

To say that something can happen means that its happening is compossible with certain facts. Which facts? That is determined, but sometimes not determined well enough, by context. An ape can’t speak a human language—say, Finnish—but I can. Facts about the anatomy and operation of the ape’s larynx and nervous system are not compossible with his speaking Finnish. But don’t take me along to Helsinki as your interpreter: I can’t speak Finnish. My speaking Finnish is compossible with the facts considered so far, but not with further facts about my lack of training. What I can do, relative to one set of facts, I cannot do, relative to another, more inclusive, set. (Lewis 1976, 150)

Given his minute chance of scoring a goal, Lewis’s doing so is certainly compossible with a set of facts concerning his current position relative to the goal, his physiology,

his desire to score, and so on—and so, in an appropriate context, it's true that he can score. And in §§5 and 6 of the outline, Lewis repeatedly uses 'can' and 'can't', which strongly suggests that he takes them to be equivalent to 'is able' and 'is unable'. But if we were instead to hold Lewis to the anti-contextualist claim that, given his minute chance of success, Lewis is (context-independently) unable to score a goal, that would open up a gap between what someone *can* do (according to Lewis [1976]) and what they are *able* to do.

We therefore propose that Lewis is here gesturing towards a contextualist view of ability claims, such that what an agent is *able* to do is context-dependent in such a way that keeps it in line with what—holding the context fixed—they *can* do.

3. FRESH START: ABILITY IFF NO OBSTACLES

Lewis's proposed analysis of 'ability' is disarmingly simple: an agent is able to perform a basic action if and only if there are no obstacles to her doing so (where defining 'obstacle' is deferred to the next section). He goes on to define abilities for "generated" actions: "You're able to ? iff for some basic action (1) Doing it would be ?ing; (2) No obstacle to doing it."

He then asks whether or not we must know which basic action the generated action is in order to be able to perform the generated action, and says that the answer is sometimes yes and sometimes no. He gives the example of a safe: opening the safe is a generated action, where the basic action is keying in the right combination. Ordinarily, we would say that someone (Helen, say) who doesn't know the combination—and hence doesn't know which basic action opening the safe would be—is not able to open the safe. But, he says, "Sometimes we think no: The secrets could be even safer if not only didn't you know the combination but also you were unable to open the lock, say, because it's on a timer." We take it the idea here is that the fact that the secrets would be even safer if the lock was on a timer entails that, in the absence of the timer but in a context in which the possibility of installing a timer is salient, Helen *is* able to open the safe despite not knowing the combination. Thus Lewis seems—as, we argued, he was in the previous section—to be pointing towards a contextualist view about abilities: it can be true in one context, but false in another, that you are able to open the safe.

A question that Lewis does not explicitly answer is whether (in a context in which it's true to say that one lacks the relevant ability) not knowing which basic action one's generated action would be constitutes an *obstacle* to doing it. On the one hand, not knowing which basic action it is does not appear on his list in the next section of kinds of thing that can count as obstacles (though clearly this list is not supposed to be exhaustive). On the other, his definition of ability seems to require that he *does* count lack of knowledge as an obstacle (again, in the relevant context). Without that assumption, his definition of ability for generated actions is satisfied—independently of context—in the safe case, since (1) there is indeed some basic action that would be the opening of the safe (namely keying in the right combination) and (2) by hypothesis there is no obstacle to Helen performing that basic action. So for his analysis to yield the verdicts he indicates in the safe case, this kind of ignorance must—in some contexts but not in others—count as an obstacle.

4. WHAT'S AN OBSTACLE?

An obstacle, Lewis tells us, is a preventer and a preventer is something that would cause an agent not to perform an action. He then tells us something about what obstacles are not. For starters, an obstacle does not have to be incompatible (either nomologically or *simpliciter*) with the agent's performing the action since there are chancy preventers. Imagine an agent who is just about to try to open a door. A controller surveilling the door pushes a button that activates a chance mechanism, which often but not always results in the door being locked. In this particular case the door locks. Then the controller's act of pushing the button was an obstacle to the agent's opening the door although the agent's opening the door was nomologically compatible (and hence also compatible *simpliciter*) with the controller pushing the button.

Neither—Lewis says—is nomological incompatibility (or incompatibility *simpliciter*) with the agent's action *sufficient* for something to be an obstacle. We assume that by 'nomologically incompatible' Lewis means 'nomologically incompatible in the circumstances'. Lewis's view in "Are We Free to Break the Laws?" (Lewis 1981) is that an agent *S* is never able to do anything, *A*, such that doing *A* would itself constitute a violation of the laws. Call the conjunction of the actual laws *L*. It would seem that Lewis does—or should—think that if the truth of *L* is, on its own, incompatible with *S*'s *A*-ing then the truth of *L* is sufficient to constitute an obstacle to *A*-ing. *S* is unable to run faster than the speed of light. If the inability is to be understood in terms of the presence of an obstacle, then the obstacle in question must surely simply be the truth of *L*. By contrast, on Lewis's (1981) view, incompatibility with the laws *in the circumstances that actually obtain* is no obstacle to, say, raising one's hand; so if we understand 'nomological incompatibility' as incompatibility with the laws in the circumstances, nomological incompatibility is indeed, on Lewis's view, no obstacle to raising one's hand (in normal circumstances).

This, however, is not the reason Lewis gives here for nomological incompatibility not being sufficient for something to count as an obstacle; rather he appeals to the possibility of reliable traces of obstacles, which are not themselves obstacles. A huge rock blocks Ann's way; it prevents her from leaving, and indeed its presence is nomologically incompatible (in the circumstances) with her doing so. Its presence is captured on a CCTV camera. We take it the idea here is that the existence of the footage implies the presence of the rock (given some background assumptions) and is thus also nomologically incompatible with Ann's leaving. But it is not a preventer of her leaving, and is hence not an obstacle. Similarly, the fact that Maria will not, in five minutes' time, raise her hand is not an obstacle to her raising it then, since it is not a preventer: the future fact is not a cause of her failing to raise it.

Not all preventers are obstacles. So which are? Lewis's answer is: obstacles are preventers that would not go away if things were just a little different. As an example of a preventer that is *not* an obstacle, Lewis mentions being unlucky. Consider again the skilled footballer Lloyd. There are occasions where he tries to score a goal and misses because he is unlucky. For instance, an unexpectedly strong gust of wind throws the ball off its course as it travels through the air after he kicked it. This causes Lloyd not to score so it constitutes a preventer. But this preventer would have

been absent, had things been just a little different. So it is not an obstacle. Thus, Lewis's analysis accommodates his previously stated intuition that Lloyd's ability does not desert him when he is being unlucky. Moreover, even if the "minute details" of the preventer determine that Lloyd actually fails, it is still something that would go away if things were just a little different. So Lewis's analysis also accommodates the result of the compatibilist argument in the first section of his notes: agents are sometimes able to act otherwise even if determinism is true, because sometimes there are no obstacles to their acting otherwise.

Now, consider the list of examples Lewis offers of preventers that *are* obstacles. Being in shackles, paralyzed, or dead, etc. are preventers that we normally take to deprive agents of their ability to perform certain acts. But are there not examples of some of these preventers where it is true that had things been just a little different, the preventer would have been absent? Take not-being-on-the-spot. An agent was not able to board a train because she was not at the platform. Is it not conceivable that had things been just a little different, she would have been there? For instance, had she not followed an impulse to check the stove before she left home, she would have arrived at the platform thirty seconds earlier. A response on Lewis's behalf might be that this does not count as being "just a little different." The deviation from the actual course of events lies a while back in time relative to the moment of the action in question, so the past would have to have been more than just a little different. And in cases where the deviation would have been close enough in time, it is bad luck rather than not-being-on-the-spot that is the preventer (and hence is not an obstacle). For instance, the agent slipped when she was just a couple of steps from the train. In that case she misses the train not because she is not on the spot, but because she is unlucky. And perhaps this is so in every case of a preventer on the list: for it to not be just bad luck the deviation from the actual course of events has to have happened too far back in time. And then it is not true that things would have been just a little different relative to the time of the action.

A notable exception from Lewis's list of preventers that are obstacles is the Frankfurt-style counterfactual intervener. This would seem to be a curious omission. While his mention of "finkish dispositions to succeed" in §2 might seem to suggest that—since Jones does indeed have the finkish (or masked) disposition to decide otherwise—Jones is indeed able to do otherwise, it seems equally clear that a Frankfurtian counterfactual intervener is a robust preventer: no matter how Jones's deliberation had gone, Black would have stepped in and ensured that he decided exactly as he did. (We briefly discuss the implications of Lewis's account of abilities for questions concerning moral responsibility in §7.)

5. REMOVABLE OBSTACLES

Having just presented the hypothesis that an obstacle is a robust preventer, Lewis now appears to present a caveat. Some robust preventers are removable. Do such robust preventers count as obstacles or not? According to the hypothesis (and according to Lewis's claim that "something is an obstacle as long as it persists") the answer is yes. But he goes on to claim that we are in fact often ambivalent about whether such robust preventers are obstacles. In particular, he claims that we are often

ambivalent about whether the lack of a skill that I am able to acquire (that is, lack of a skill such that there is no robust preventer of me acquiring it) is an obstacle.

There seems, then, to be a tension in Lewis's account at this point. On the one hand, an obstacle is a robust preventer, and something is an obstacle as long as it persists. But that implies that the right answer to the question, "can I shed my accent?" is not—as Lewis says—"Yes, but only after long practice," but rather "No, but I *will* be able to if I practice a lot." The "Yes, but . . ." answer can only be the right answer—in at least some contexts—if, in such contexts, we are prepared to judge that removable obstacles are not obstacles at all. But in that case the hypothesis that an obstacle is a robust preventer must be false: at least in some contexts, some robust preventers are not obstacles. We return to this issue in §7 below.

6. IS PREFERRING NOT TO AN OBSTACLE?

Preferring not to do something is a preventer. It causes agents not to do the thing they prefer not to do. But is it an obstacle? We take it that Lewis's strategy at the start of this section is to list some commonsense claims, explain the underlying reason (in terms of the "balance of pros and cons") why those claims are all true, and then use that as confirming evidence for his hypothesis that an obstacle is a robust preventer. When the reasons against doing *A* are overwhelming (because one would lose one's job or get shot or whatever), then one's preference for refraining is indeed an obstacle to *A*-ing (although in some circumstances those same reasons might not be overwhelming: losing one's job might be an overwhelming reason not to call in sick so that one can spend the day at the zoo, but it is not an overwhelming reason to obey the boss's command to shoot some hostages). When the reasons are finely balanced, however, preferring not to is not an obstacle. Lewis takes this to confirm his hypothesis, claiming that an overwhelming reason to refrain is a robust preventer and hence an obstacle, while a more finely-balanced reason to refrain is nonrobust. (Perhaps one might have made the opposite decision on the basis of the same reasons had circumstances been slightly different; or perhaps the reasons themselves might easily have been slightly different.)

It is worth briefly comparing Lewis's view here to that of one of his incompatibilist adversaries, Peter van Inwagen. Van Inwagen writes that agents are faced with a genuinely free choice only on "occasions on which an agent is confronted with alternatives and it is not clear to him what to do—not even when all the facts are in" (van Inwagen 1989, 415). According to van Inwagen, these occasions are limited to two types of cases: "cases of an actual struggle between perceived moral duty or long-term self-interest, on the one hand, and immediate desire, on the other; and cases of a conflict of incommensurable values" (ibid. 417). These cases undoubtedly exemplify a delicate balance between pros and cons. Van Inwagen considers a third type, namely cases where the alternatives are incompatible but seemingly interchangeable. In this category he also includes cases where the alternatives are not interchangeable in the sense that the agent cannot tell the difference between them, but the properties that constitute the difference between them are the objects of conflicting desires. Van Inwagen is hesitant to say that the agents in such cases have a

choice about what to do since what occurs in those situations is rather “something like an internal coin-toss” (ibid.).

Lewis’s analysis would appear to be rather more permissive than van Inwagen’s. (It should be noted that while Lewis is here elucidating an account of the conditions under which agents are able to do otherwise rather than an account of free action, he does seem in §1 to treat being able to do otherwise and having a “free choice” as equivalent.) Van Inwagen’s “internal coin-toss” cases would seem to be straightforward cases of the ability to do otherwise for Lewis. Moreover he is, he claims, able to order lager even though he prefers bitter (so long as his preference is not overwhelming), and that case does not fit any of van Inwagen’s criteria.

On the other hand Lewis’s account is considerably less permissive than traditional conditional analyses of the ability to do otherwise, since it implies that agents are often unable to do otherwise even if it’s true that they would have done otherwise had they chosen or wanted or tried to. In a case where someone has an overwhelming preference for bitter, say, Lewis thinks they are not able to order lager. It is surely nonetheless true that if they *had* chosen (or wanted or tried) to order lager, they would have succeeded, since had they been in that situation, the obstacle to their ordering the lager (namely their overwhelming preference for bitter) would not have obtained.

We end our section-by-section commentary by noting that Lewis’s story about when and why preferences constitute obstacles is at best incomplete and at worst rather problematic. The notion of an obstacle is a modal notion, and it is unclear how well such a notion will line up with facts about “pros and cons.” First, even given an internalist understanding of reasons—pros and cons—many actual agents’ preferences do not line up especially well with their reasons because they are not acting rationally. In cases of irrationality an agent’s preference for *A*-ing may be overwhelming despite finely balanced pros and cons, or—conversely—their preference may be mild but their reasons for *A*-ing vastly outweigh their reasons for refraining. Lewis’s analysis of the ability to do otherwise ought to count such an agent as unable to do otherwise in the first case and not the second, but—if so—it cannot be that whether or not a preference is an obstacle “depends on the balance of pros and cons.”

Second, even in rational agents it seems implausible to claim that an “overwhelming balance” of cons vs. pros with respect to *A*-ing always constitutes a robust preventer of *A*-ing, since—contrary to what Lewis assumes—it may not take much of a departure from actuality to overturn or render otiose an overwhelming balance. Suppose Helen’s preference for bitter over lager is overwhelming—she really doesn’t like lager that much at all and adores bitter—but, as it happens, she really does want a drink, this establishment only sells two kinds (bitter and lager), and the bitter barrel is very nearly empty. In a close possible world, the barrel runs out just before she gets to the bar and her choice is: lager or nothing. In that possible world she orders the lager, her strong preference for bitter notwithstanding. Or: Maria is on a hiring panel, and the evidence before her is that Candidate *X* is a better candidate than *Y* by an overwhelming margin, so she votes for Candidate *X*. As it happens, the administrator accidentally left one of *X*’s references on his desk just before the meeting—a damning one, which provides ample compelling reasons not to hire *X*—so Maria

didn't read it. In a close possible world Maria's overwhelming preference for *X* is overturned: the balance of pros and cons shifts dramatically. Thus it would seem that an overwhelming preference need not be a sufficiently robust preventer (in this case, a preventer of Maria's voting to hire *Y*) to count as an obstacle.

7. CONTEXTUALISING LEWIS ON ABILITIES

How does Lewis's proposal regarding abilities fit into the contemporary debate? At first glance, not easily. Broadly speaking, two main approaches to the analysis of abilities have been developed. The first takes as its inspiration the simple conditional analyses of Hume (1955, §VIII), Moore (1912) and Ayer (1954). For reasons discussed earlier (§2), such analyses are now widely thought to be untenable in their simple form. However, their central insight has been embraced and developed by dispositional analyses of abilities. These argue that, fundamentally, abilities are dispositional properties instantiated by agents.⁴

The second approach looks, instead, to Angelika Kratzer's analysis of 'can' (1977). What we are able to do, or what we can do, is analysed in terms of restricted possibilities. *S* is able to *A* if and only if *S*'s *A*-ing is compossible with certain facts, where which set of facts is relevant is determined by the context of the ascriber. As we saw in §2, in his discussion of time travel (1976) Lewis embraces Kratzer's view of 'can'. How, then, does Lewis's new proposal regarding abilities cohere with this view? Or is it supposed to be, as advertised, a "fresh start," with no links to his earlier view or other analyses in the literature?

The answer to this question is rather unclear. Certainly, Lewis is not drawn to the view of abilities as dispositions, as no attempt is made to employ his own analysis of dispositions (a move made explicitly, for instance, by Vihvelin [2004, 2013]). It is noteworthy, however, that one might take some of his reasons for rejecting the conditional analysis of ability (§2 above) to motivate just such a dispositional analysis: after all, several of the cases he mentions are, precisely, cases where the agent has a disposition—and, Lewis thinks, the ability—to *A*, but it is not true for some appropriate *X* that if *X* had been the case, the agent would have *A*-ed. Lewis's own account of dispositions is well placed to explain why such cases are indeed cases where the agent has the disposition, the falsity of the relevant counterfactual notwithstanding. Still, for whatever reason this is not a line that Lewis chooses to pursue.

On the other hand, nor is there much by way of direct evidence that Lewis's proposal is supposed to be an extension of a broadly Kratzerian contextualist account of 'can' claims, in line with his remarks about Tim the time traveller (Lewis 1976). However, we think that Lewis's proposal fits quite naturally within this broad approach, especially given the contextualist element that seems to be implicit in various moves that he makes in the outline. We suggested in §3 that a contextualist reading is the best interpretation of Lewis's remarks about his inability to score a goal, and also that it is a natural interpretation of his claims in a number of places about the 'ambivalence' of our ability claims.

If that's right, then it looks as though we should take Lewis's considered view to be that whether or not something is a robust preventer is a context-sensitive matter—since I am able to *A* if and only if there is no obstacle to my *A*-ing, and there is

no obstacle to my *A*-ing if and only if there is no robust preventer of my *A*-ing. But it is worth noting that Lewis's distinction between a preventer and an obstacle (that is, a robust preventer) is analogous to his distinction between sensitive and insensitive causation. Consider, for instance, Lewis's example of insensitive causation, shooting someone at point-blank, causing their death (1986, 186). Given the cause, the effect is relatively insensitive to the circumstances in which it pertains. This contrasts to a case of sensitive causation, where a slight change in the circumstances would have resulted in the effect not occurring. Nowhere within this discussion does Lewis suggest that this distinction is relative to the context of utterance.⁵ Hence this provides some, albeit defeasible, evidence in the other direction—evidence, that is, that Lewis regards the question whether something is a robust preventer as one whose answer is invariant across different contexts.

On the other hand, it does look as though Lewis needs whether something is a robust preventer or not to be context sensitive in order to make sense of the goal-scoring and safe cases. Consider the question of whether or not Helen is able to open a safe when she doesn't know the combination—to which, we suggested, Lewis thinks the answer is context-dependent. (Normally, no; but if the possibility of installing a timer is salient, then yes: the secrets would be even safer that way.) We take it that neither this nor the goal-scoring case is supposed to be a case where we are ambivalent about whether *removable* obstacles count as obstacles (as per §5): the issue isn't whether Helen might *find out* the combination to the safe or whether Lewis might *acquire* the ability to score a goal. Rather, it looks as though these are cases where we are ambivalent about whether the relevant lack of knowledge or skill constitutes a candidate robust preventer in the first place. So perhaps the thought here is supposed to be: in some contexts we do not count (for example) 'fluky' worlds where Helen opens the safe or Lewis scores as worlds that are "just a little different," and in others we do.

Another place where ambivalence and hence, potentially, context-sensitivity, comes in is in Lewis's remarks on removable obstacles. Those remarks are somewhat puzzling. He says that something is an obstacle as long as it persists, but also that we sometimes (and, implicitly, legitimately) don't count an obstacle that one is able to remove as an obstacle at all. How can both of those claims be true? After all, an obstacle that one is able to remove is still an *obstacle*, so it can't be that one's ability to remove it renders it not an obstacle after all. Of course there is a perfectly good common-or-garden sense of 'obstacle' according to which a removable 'obstacle' in *that* sense is not an obstacle in Lewis's technical sense: in the common-or-garden sense, there may be plenty of obstacles on the path to achieving *X* that do not prevent me from achieving *X*—in fact, or in nearby possible worlds. Thus an 'obstacle' in this common-or-garden sense need not be a robust preventer, or indeed a preventer at all: 'There were many obstacles along the way but I did it!' is perfectly coherent. But clearly this sense of 'obstacle' is not what Lewis has in mind when he says that we're ambivalent about whether removable obstacles are obstacles.

Might an appeal to the context-sensitivity of 'robust preventer' proposed above help here? The idea here would be that what counts as a removable obstacle (and hence an obstacle) in one context might not count as an obstacle at all in a different

context because in the first context it's counted as a robust preventer (albeit one that one is able to remove) and in the second context it isn't. That proposal, however, doesn't look very promising. Take insufficient funds: can Ann buy this expensive car? In one context yes (but only after saving up for it), but presumably—since Lewis claims that we are ambivalent—in another context, no (she doesn't have enough money). But it just doesn't seem plausible to claim that the reason why 'yes' is the right answer in some contexts is that in those contexts Ann's lack of funds is not counted as a robust preventer. A robust preventer is something that prevents one from *A*-ing and would still exist (and prevent one from *A*-ing) if things were just a little different. But if Ann's current financial state is such that she needs to save up for the car in order to buy it, then it looks as though things would have to be *a lot* different in order for her *not* to *now* lack sufficient funds. She would have had to have started saving some time ago, or just won the lottery or inherited some money, or whatever, and those are all rather large departures from actuality.

Moreover, this proposed explanation for our ambivalence fails to connect in the right way with what Lewis says: the reason for our ambivalence is supposed to be the fact that one is able—by starting now or in the near future to save money or take a course of exercise or whatever—to remove the obstacle eventually. By contrast, whether lack of funds or puniness counts as a robust preventer of one's *A*-ing depends on how the actual world is—and hence how nearby possible worlds are—*right now*, and not how they will or might pan out in the future.

In fact, we want to propose a different contextualist explanation of our (alleged) ambivalence in these cases. We are not confident that it is what Lewis had in mind, but it does at least make sense of the ambivalence in a way that is consistent with what Lewis says here. The proposal is that what explains our ambivalence is that the ability claims in question are ambiguous with respect to time-indexing in a way that is (often) settled by context. Is Ann able, right now, to buy the expensive car right now? No: she lacks the funds. Is she able, right now, to buy the expensive car in a few months' time? Yes—but only after saving up for it. Thus the question 'Is Ann able to buy the car?' remaining silent, as it does, on what the time of action is, gets different answers depending on how we fill in the time of action. Correspondingly, whether or not Ann's lack of funds (or skill, or strength, or whatever) counts as an obstacle to *A*-ing depends on whether the question is whether she is able to *A* right now (or soon)—in which case it does count as an obstacle—or instead whether she is able to *A* in the future, in which case it does not count as an obstacle (though she will have to make some effort—by acquiring the necessary funds or skill or strength—to ensure that it does not at some point *become* an obstacle). As Lewis says, "I am not able to do the thing right away. But I'm able to remove the obstacle and then do it."

Our proposal, then, is that we take this claim at face value, and explain the ambivalence by appealing to the fact that it is generally settled by context whether what is in question is one's ability to do it right away, or rather one's ability to remove the obstacle (that is, the obstacle to doing it right away) and then do it. (And where context does not settle the matter, we are genuinely ambivalent: there is no fact of the matter about whether one has or lacks the ability.)

We end by considering whether or to what extent Lewis's discussion of abilities connects with debates about the criteria for moral responsibility, where one perennial locus of disagreement has been whether those criteria include the ability to do otherwise. Lewis's notes on preferences are likely to strike at least some of those compatibilists who think that the criteria for moral responsibility do include the ability to do otherwise as a move in the wrong direction, because Lewis in effect eschews any necessary connection between the ability to do otherwise and reasons-responsiveness. On Lewis's analysis, our own preferences count as obstacles if and only if they are modally robust preventers. So if Helen's preference for bitter would lead her to choose bitter in a range of not-too-distant counterfactual circumstances, that constitutes an obstacle to her choosing lager. But this, we might think, is mistaken—and we will be especially inclined to think that if we want moral responsibility and the ability to do otherwise to line up. The fact that Helen's preference for bitter is strong enough to count as an obstacle to choosing—and hence something that renders her unable to choose—lager, would not seem to render her nonresponsible for choosing the bitter. Having a strong—as opposed to a finely balanced—preference for theft or kidnap or telling a lie is not generally thought to count as a reason to let someone off the hook! (Such a person is, of course, failing to be as sensitive to distinctively moral reasons as they ought to be. Nonetheless they might well count as sufficiently reasons-responsive to meet the reasons-responsiveness requirements on moral responsibility proposed by, for example, Fischer and Ravizza [1998] and Sartorio [2016].)

Many compatibilists might therefore want to resist imposing a Lewis-style ability-to-do-otherwise requirement on moral responsibility. On the other hand, there are moves that a Lewisian might make at this point. One would be to insist that the project here is to defend what Kadri Vihvelin (2013, 19) calls “metaphysical compatibilism” (the thesis that determinism is compatible with free will) and not “moral compatibilism” (the thesis that determinism is compatible with moral responsibility), and to claim, further, that the ability to do otherwise is a constraint on acting freely but not on moral responsibility. Another move would be to endorse a weak connection by accepting that to be a morally responsible agent at all, one must be the kind of agent who sometimes or often has the ability to do otherwise—but this need not be present on every occasion (cf. Vihvelin 2013, 90; van Inwagen 1989, 418–19). Or they might avail themselves of a standard ‘tracing’ move, so that *S* can be morally responsible for *A*-ing despite lacking the ability to do otherwise, so long as the causal history of *S*'s *A*-ing on this occasion can be traced back to some previous action(s) or decision(s), *B*, such that (i) *S* could have done otherwise than *B* and (ii) by virtue of moulding *S*'s character or their habits of thought or action, *B* has made *S* the kind of person who is now unable to do otherwise than *A* (cf. for example Kane [1999]).

We close with an observation about the general strategy that Lewis appears to adopt in the outline, which connects with the first of the moves described in the previous paragraph. Granted, he starts out with an argument for the compatibility of the ability to do otherwise both with determinism and with indeterminism, and, as he says, we “need” a compatibilist theory of abilities. But free will and compatibilism largely disappear from view in the rest of the outline. So we might think of Lewis's

broader aim as simply that of providing an all-purpose analysis of ability that comports with our commonsense talk and thought. After all, talk of abilities pervades our everyday interactions (“Are you able to make the meeting tomorrow?,” “Can she bench-press a hundred kilos?,” “Can I buy that car?”), where philosophical considerations concerning free will and determinism are rarely pertinent. That the account must be compatibilist-friendly operates as a constraint, given the argument of §1, but it is not the main purpose of the account.

If we conceive the outline in that light, we might think of it as an attempt to shift the focus of the debate about the role of the ability to do otherwise in judgments about free will and moral responsibility. As we said, abilities—and their absence—are part and parcel of our ordinary talk and thought, and hence a conceptual analysis of ‘ability’ is something any philosopher might want to have at their disposal, whether or not they are interested in either free will or moral responsibility. Indeed a philosopher who is interested in those topics ought (or at least so Lewis would surely think) to be availing themselves of an account of ‘ability’ that meshes with our ordinary ascriptions and denials of ability, otherwise they might reasonably be accused of changing the subject when presenting views about what it is to be able to do otherwise in the specific context of defending or attacking the view that the ability to do otherwise is compatible with determinism (or with indeterminism).

So if we *start* with such an account, that may well place constraints on what counts as a legitimate move in debates about whether or not the ability to do otherwise is compatible with determinism, and about whether or not that ability is required for acting freely or morally responsibly. It may turn out, for example, that many fully reasons-responsive and hence morally responsible agents are unable to do otherwise—not because determinism is true, but because the strength of their preferences constitutes an obstacle. And it may turn out that some views about free will are untenable because they are inconsistent with an account of ‘ability’ that comports with our ordinary talk and thought. The view according to which a deterministic agent is only able to do what they actually do would fall into that category, since “I don’t know—you’ll just have to wait and see whether I show up” is never the right answer to the question, “Are you able to make the meeting tomorrow?”⁶

NOTES

1. Letter 115. To Daniel Dennett, 26 June 2000, in [Beebe and Fisher \(forthcoming\)](#).
2. Letter 141. To Brian Skyrms, 26 November 1975, in [Beebe and Fisher \(forthcoming\)](#).
3. Among its early advocates were Thomas [Hobbes \(1962\)](#) and David [Hume \(1955\)](#), and during the twentieth century its proponents included G.E. [Moore \(1912\)](#) and A.J. [Ayer \(1954\)](#).
4. For views of this kind, see, for instance, [Smith \(1997\)](#), [Fara \(2008\)](#), and [Vihvelin \(2004, 2013\)](#).
5. This has proved controversial. [Woodward \(2006\)](#), for instance, argues that this distinction is relative to context.
6. We are grateful to audiences at workshops on free will at Manchester and Cologne, to Jens Johannson and Erik Carlson for helpful comments on early drafts of this paper, and to Anthony Fisher for finding the outline of *Nihil Obstat* amongst Lewis’s papers. We are also grateful to the Arts and Humanities Research Council for supporting some of the research that went into this paper (grant no: AH/ N004000/1).

REFERENCES

Ayer, A.J. 1954. “Freedom and Necessity,” in his *Philosophical Essays*, London: McMillan, 271–84.

- Beebe, Helen and Anthony Fisher, eds. forthcoming. *Philosophical Letters of David K. Lewis: Volume I*, Oxford: Oxford University Press.
- Bird, Alexander 1998. "Dispositions and Antidotes," *The Philosophical Quarterly* 48: 227–34.
- Choi, Sungho and Michael Fara 2018. "Dispositions," in E.N. Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition). <https://plato.stanford.edu/archives/fall2018/entries/dispositions/>.
- Fara, Michael 2008. "Masked Abilities and Compatibilism," *Mind* 117: 843–65.
- Fischer, John Martin and Mark Ravizza 1998. *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge: Cambridge University Press.
- Frankfurt, Harry 1969. "Alternate Possibilities and Moral Responsibility," *The Journal of Philosophy* 66: 829–39.
- Hobbes, Thomas 1962. *The English Works of Thomas Hobbes*, vol. 5, ed. W. Molesworth, London: Scientia Aalen, orig. 1654.
- Hume, David 1955. *An Enquiry Concerning Human Understanding*, ed. L.A. Selby-Bigge, Oxford: Clarendon Press, orig. 1748.
- Kane, Robert 1999. "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism," *The Journal of Philosophy* 96: 217–40.
- Kratzer, Angelika 1977. "What 'Must' and 'Can' Must and Can Mean," *Linguistics and Philosophy* 1(3): 337–55.
- Levy, Neil 2008. "Counterfactual Intervention and Agents' Capacities," *The Journal of Philosophy* 105: 223–39.
- Lewis, David K. 1976. "The Paradoxes of Time Travel," *American Philosophical Quarterly* 13: 145–52.
- . 1981. "Are We Free to Break the Laws?" *Theoria* 47(3): 113–21.
- . 1983. "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61: 343–77.
- . 1996. "Elusive Knowledge," *Australasian Journal of Philosophy* 74(4): 549–67.
- . 1997. "Finkish Dispositions," *The Philosophical Quarterly* 47: 143–58.
- . 2004. "Void and Object," in J. Collins, N. Hall, and L.A. Paul, eds. *Causation and Counterfactuals*, Cambridge, MA: MIT Press, 277–90.
- . 2020. "Outline of *Nihil Obstat*: An Analysis of Ability," *The Monist* 103(3): 241–44.
- Moore, G.E. 1912. *Ethics*, Oxford: Oxford University Press.
- Sartorio, Carolina 2016. *Causation and Free Will*, Oxford: Oxford University Press.
- Smith, Michael 1997. "A Theory of Freedom and Responsibility," in his *Ethics and the A Priori*, Cambridge: Cambridge University Press, 84–113.
- van Inwagen, Peter 1975. "The Incompatibility of Free Will and Determinism," *Philosophical Studies* 27: 185–99.
- . 1983. *An Essay on Free Will*, Oxford: Clarendon Press.
- . 1989. "When Is the Will Free?" *Philosophical Perspectives* 3: 399–422.
- Vihvelin, Kadri 2004. "Free Will Demystified: A Dispositional Account," *Philosophical Topics* 32: 427–50.
- . 2013. *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*, Oxford University Press.
- Woodward, James 2006. "Sensitive and Insensitive Causation," *The Philosophical Review* 115(1): 1–50.