

MAKING Ethics

by

Thomas Byrne

MPhil Stud King's College London, 2016

B.A. King's College London, 2014

Submitted to the Department of Linguistics and
Philosophy in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

©2021 Thomas Byrne. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author:.....

Department of Linguistics and Philosophy

May 28, 2021

Certified by.....

Caspar Hare

Professor of Philosophy

Thesis Supervisor

Accepted by:.....

Bradford Skow

Laurence S. Rockefeller Professor of Philosophy

Chair of the Committee on Graduate Students

MAKING Ethics

by

Thomas Byrne

Submitted to the Department of Linguistics and Philosophy on
May 28, 2021 in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

ABSTRACT

A window broke and Annie was involved. What's of moral importance in a situation like this? Not whether Annie caused the window to break and not whether the window wouldn't have broken if it weren't for Annie. What's morally important is whether Annie *broke* the window. In this thesis, I first generalise and argue for that claim; afterwards, I put it to work in ethics, applied ethics, and legal theory.

Thesis Supervisor: Caspar Hare
Title: Professor of Philosophy

Acknowledgements	4
Introduction	5
MAKING Metaphysics	9
The Bystander Doesn't Kill	34
Legal Causation	43
On Offsetting	64

Acknowledgements

This thesis was written and many other people were involved. What's of moral importance in a situation like this? Immediately undermining this work's central claim, it's not whether any of them wrote it, but that it's better than it would have been without them—and better it is. I'm grateful to all of them, but will single-out only a few.

Kevin Dorst, Cosmo Grant, Agustín Rayo, Kieran Setiya and Judy Thomson each read and commented upon much of what's written here.

In addition to the five above, David Builes, Joseph Byrne, Brendan de Kenessey, Haley Schilling, Jack Spencer and Jake Wojtowicz discussed many of these ideas with me.

Judy Thomson inspired me to do ethics and she will long remain the voice inside my head: “think carefully, write clearly, do right by your reader—no nonsense.” I miss her.

David Balcarras, Allison Balin, David Builes, Alex Byrne, Kevin Dorst, Nicole Garcia, Cosmo Grant, Michele Odisseas Impagnatiello, Milo Phillips-Brown, Haley Schilling, Jack Spencer and Kirsi Teppo each contributed to my time at MIT in the most valuable and sustaining of ways: making it FUN.

And I save the best to the last. Caspar Hare (my supervisor) and Brad Skow (who could well have been my supervisor) were fantastic and I had a blast—thanks guys.

Introduction

If an ethicist had been the first person to notice that, for example, *causing to die*, *causing to flood*, and *causing to break* all shared something in common—namely, that they are all instances of the same metaphysical relation causation—that ethicist might well have argued for a causation-centric theory of ethics. That theory would hold that, by paying close attention to matters of causation, we can achieve serious progress on various ethical problems.

I am not that ethicist, but this thesis follows a parallel path. I argue that *killing*, *flooding*, *breaking*, *melting*, etc., all share something in common—namely, that they are all instances of the same metaphysical relation that I call *MAKING*. I then argue for a *MAKING*-centric theory of ethics. By paying close attention to matters of *MAKING*, this theory achieves serious progress on various ethical problems. I call this theory “*MAKING* Ethics.”

In the first half of “*MAKING* Metaphysics,” I lay the groundwork for all this. My objective there is to get *MAKING* on a firm-footing such that readers are comfortable with *MAKING* and, importantly, confident that they know *MAKING* when they see it. To that end, I show that *MAKING* is distinct from causation (if, for example, Annie merely instructs Ben to break the window, then while Annie would have caused the window to break, she wouldn’t have broken it herself—that is, she wouldn’t have *MADE* it break). I then show that *MAKING* doesn’t reduce to causation, either—that is, *to MAKE* is not simply *to cause-in-a-particular-way*. I also argue that, despite us being unfamiliar with the general, metaphysical relation *MAKING*, we are nonetheless extremely familiar with each of its various instances: with what is and isn’t a *killing*, or a *flooding*, or a *breaking*, etc. “*MAKING* Metaphysics” ends by arguing that *MAKING* cuts at certain joints in ethics and in legal theory and, therefore, metaphysicians (and ethicists, and legal theorists) should pay attention to it.

In arguing that *MAKING* cuts at certain joints in ethics, I argue for a surprising claim regarding Judith Thomson’s famous “bystander” trolley case. In that case, a bystander diverts a runaway trolley away from

five works and towards one worker. Thomson claimed, and everyone else agreed, that it was permissible for the bystander to divert the trolley even though, in doing so, the bystander killed the one worker in order to save the five others. Thomson therefore claimed, and everyone else agreed, that the bystander case is a counterexample to the principle that, other things equal, its impermissible to kill one person in order to save the lives of five others. The surprising claim I argue for is that, in diverting the trolley, the bystander does *not* kill the one worker, after all; and, in turn, that the case is not a counterexample to the aforementioned principle. The argument followed quickly from some basic facts about MAKING. “The Bystander Doesn’t Kill” offers another argument for that same conclusion—one that doesn’t appeal to MAKING.

In “Legal Causation,” I shift focus from ethical theory to legal theory. Compare:

Wanting to break Ben’s leg, Annie sabotages the lintel of his front door. When Ben returns home and opens the door, the lintel collapses upon him, breaking his leg.

As above, but this time Ben notices the sabotage and instead of opening the door, he heads to buy a new lintel. On the way, he is hit by a car, breaking his leg.

In both cases, Ben’s leg wouldn’t have broken but for Annie’s actions, but only in the first case will the courts (rightly) find Annie guilty of breaking Ben’s leg. Lawyers say that this is because only in the first case is Annie a *legal cause* of Ben’s broken leg.

A definition of legal causation has eluded legal theorists. Instead, the standard approach (the one found in legal text books and in the rhetoric of court judgements) first asks whether the defendant was a cause of the harm in question and then provides various principles (/“tests”) for determining whether the defendant was additionally a legal cause of that harm. The unhappy result is a piecemeal theory of legal causation.

I argue that the standard approach is mistaken since legal causation is, in fact, just MAKING such that, e.g., Annie is a legal cause of Ben’s broken leg if and only if Annie breaks Ben’s leg. Unlike the standard, piecemeal approach, my account of legal causation is principled and unified.

In “On Offsetting,” I shift focus again—this time from legal theory to applied ethics. I consider the problem of environmental offsetting. Starting with the assumption that those affected by climate change have a complaint against us polluters today in virtue of our carbon footprints, this problem asks whether

that complaint remains even if we offset our CO₂ emissions—e.g. by capturing an equivalent amount of CO₂ from the atmosphere. John Broome claims that the complaint doesn't remain and that those affected by climate change have no complaint against us under those circumstances. However, his view is not popular; I defend it. The first part of that defence requires demonstrating where the various arguments against Broome's view go wrong: I argue it's because each of those arguments is blind to MAKING and, once we are live to it, it's clear that those arguments are unsound. The second part of the defence requires presenting a positive argument for Broome's conclusion. That argument introduces new ideas on the ethics of risk. In particular, I consider cases in which someone increases the risk of some harm, but also reduces that risk by the same degree. For example:

Gunslinger is determined to play Russian roulette with Vasha. Gunslinger has four bullets and a revolver with eight chambers. Elena swaps his eight-chambered revolver for one with six chambers. She then takes one of Gunslinger's bullets. Gunslinger loads the three bullets into the six-chambered revolver, vigorously spins the cylinder and, when it stops spinning, he aims at Vasha and pulls the trigger: Vasha is shot dead.

And I ask whether Vasha has a complaint against Elena in virtue of her swapping Gunslinger's revolver, given that Elena also took one of his bullets. I argue (a) that Vasha has no complaint against Elena and (b) that offsetting our emissions is structurally analogous to that case and therefore (c) that those affected by climate change have no complaint against us in virtue of our (offset) emissions.

Finally, it's worth highlighting two virtues of MAKING that are repeatedly put to work throughout this thesis. The first is that with MAKING in hand we can neatly organise classes of cases that previously looked untameable—seemingly too varying and too multifaceted to be organised in any unified way. This is most obviously at work in “Legal Causation,” but it shows up in the other chapters, too.

The second virtue of MAKING is that it enables us to move from facts about certain cases to conclusions about others cases—even though, on first glance, those cases might seem to share little in common. For instance, we could move from the fact that Annie doesn't break the window here to the conclusion that Ben doesn't kill the victim there. We can do so because *breaking* and *killing* actually share something

very important in common—namely, they are both instances of MAKING. This virtue is most obviously at work in the final sections of “MAKING Metaphysics,” but it also shows up in other chapters, too.

This thesis really represents just the beginning of “MAKING Ethics.” I have high hopes for what’s to come next...

MAKING Metaphysics

1. When the window breaks, it might be the case that

Annie caused the window to break,

and it might be the case that

Annie broke the window.

Often the two will go hand-in-hand—for instance, when Annie throws a rock at the window—but they needn't. If Annie only instructed Ben to throw a rock at the window, then while it wouldn't be the case that Annie broke the window, it would be the case that she caused the window to break (by causing Ben to break it).¹

Similarly, when the house floods, it might be that Annie caused the house to flood and it might be that Annie flooded the house. These can also come apart: floodwaters are heading towards the open doorway of Annie's house; Annie closes her door and so the waters instead enter Ben's house, one door down. Annie doesn't flood Ben's house, even though she causes it to flood.

And this is nothing special about Annie, since the same goes more generally. When the chocolate melts, it might be that the fire caused the chocolate to melt and it might be that the fire melted the chocolate. These can come apart, too: if the fire cuts power to the air-conditioning, then the fire doesn't melt the chocolate, even though it causes the chocolate to melt (by causing the summer air to melt it).

There's a pattern here. *Break*, *flood* and *melt* are all verbs that can be used both *intransitively* and *transitively*. A verb *V* is used intransitively in a given sentence when it doesn't allow a grammatical object (e.g. “the window broke”), while it's transitive when it does (whether explicitly, e.g. “Annie breaks the window,” or implicitly, e.g. “stop stealing [cars]!”). If we mark these intransitive and transitive uses with the subscripts “I” and “T,” then we can capture the pattern as follows.² For thing *X* and verb *V*, when

$X V_I$'s (e.g. the window breaks),

¹ Not only does *Annie* cause the window to break_I (an instance of so-called *thing* causation) but so too does *Annie's instructing Ben* cause the window to break_I (*event* causation). I focus on the former, but, in doing so, I make no claim about which kind is fundamental. I also focus on whether *A causes X to V_I* as opposed to whether *A is a cause of X's V_I-ing* (although, this distinction crops up briefly in §5); again, this is not a substantive commitment on my part.

² I say, for instance, that the verb *melt* can be used both transitively and intransitively: *melt_T* and *melt_I*. I might instead have said that there are two verbs, *melt_T* and *melt_I*, that share a lexeme (as linguists call it). Nothing rests on this choice, except easier expression. Also note that while *melt_T* and *melt_I* are homographic and homophonic—looking and sounding the same—this won't always be the case. For example, when Annie kills_T Ben, it's not that *Ben kills_I*, but that *Ben dies_I*. Irregular verbs like these are common and, for my purposes, uninteresting. See Parsons (1990) for more on *kill* and *die*.

it might be the case, for actor *A* (where *actor* ranges over both agents and things), that

- (i) *A* caused *X* to V_I (e.g. Annie caused the window to break)

and it might be the case that

- (ii) *A* V_T 'd *X* (e.g. Annie broke the window).

Philosophers have spent much time thinking about instances of (i) and the language is familiar: causes, causing, etc.

Not so with instances of (ii). For lack of a better term, I will say that, when $X V_I$'s:

A *MAKES* $X V_I$ just in case *A* V_T 's *X*.

Annie *MAKES* the window break_I just in case Annie breaks_T the window; the fire *MADE* the chocolate melt_I just in case the fire melted_T the chocolate; and so on. I will also talk of the things that *MAKE*—*viz.* *MAKERS*—and of the activity that *MAKERS* partake in, *MAKING*.

To what end? Well, we might, across various domains of inquiry, think that it's important to know what causes what to happen; or we might think it's important to know what *MAKES* what happen. Here are two examples from the normative domain.

When the ethicist is seeing to settle whether Annie acted wrongly, she might start by seeing what the various *consequences* of her action were (where *C* is a consequence of *E* just in case *C* counterfactually depends upon *E*). But then she might think it's important to additionally know, of those consequences, which ones Annie caused. Alternatively, she might think it's important to additionally know, of those consequences, which ones Annie *MADE*. Since causing and *MAKING* come apart, the upshot will be moral theories with conflicting verdicts.

Similarly, to convict *A* of *B*'s homicide, the lawyer must prove both that *A* possesses the requisite mental states for that crime and that *B*'s death is a consequence of *A*'s actions. But that is not sufficient for a just conviction (imagine: *A* poisons *B*'s gin, *B* notices, heads to the shop to buy a replacement bottle and is hit by a car while crossing the road) and so there must be a further component of the crime. That further component might be that *A* caused *B*'s death or it might be that *A* *MADE* *B* die—that *A* killed *B*. The upshot may be different people going to jail.³

³ Two points to avoid confusion. Firstly, it is the mental state required that distinguishes the various degrees of homicide. Roughly: intention and premeditation for first degree; intention for second degree; recklessness for involuntary manslaughter; and so on. Secondly, the law has brute provisos that deal with cases of preemption and overdetermination *vis-a-vis* the consequence requirement.

For now, it suffices to say that it's an open, empirical question whether ethicists and lawyers in fact think that what matters is who caused what or whether they think that what matters is who MADE what. (Although, for what it's worth, I don't think it's a coincidence that ethicists have spent much time with principles like "it's impermissible to *kill* one to save five," yet no time with causing.⁴ Nor does it count for nothing that the gin example, above, is itself a counterexample to the idea that the lawyer's further component is causing, since A does cause B's death, yet nonetheless should not be convicted of B's homicide. More on these points later.)

I take that same open, empirical question to apply to other domains too. Do oncologists want to discover just what things cause cells to mutate or do they want to discover just what things mutate cells? Does the crash investigator want to know what caused the boat to sink or does she want to know what sunk the boat? This time, the upshots will be different programmes of inquiry.

Since these are open questions—certainly open for the metaphysician—we might expect metaphysicians to have thought about both causing and MAKING. But they haven't. They have spent endless time thinking about causing—"what's its nature?", "is it transitive?", "can omissions cause?", and so on—but almost no time thinking about MAKING.⁵ Why is that? As far as I know, no one has ever defended the disparity (nor, even, thought that it needed defending), but, asking around, the rationale seems to look like this:

"MAKING reduces to causing so even if MAKING features in various theories in various domains it only does so in virtue of the fact that it reduces to causing."

And, indeed, when metaphysicians have paid attention to MAKING it has been in search of that reduction. In §5, I show that every proposed reduction of MAKING to causing fails. The rationale continues:

"Even though we haven't found the correct reduction, we nonetheless have good reason to think that there must be such a reduction."

In §6, I show that we do not have good reason for thinking that there must be such a reduction. And it continues still:

"There are in-principle objections to MAKING being an independent, interesting notion."

⁴ Carolina Sartorio (2010) is one exception. She considers whether causation might play an interesting role in ethics. Her conclusion is largely negative.

⁵ As far as I know, the only time that MAKING has been discussed as a serious peer of causing was in a talk by Brian Weatherston (2007). Unfortunately, the talk didn't progress to a paper. All other talk of MAKING has been oblique: concerning its grammar (see §2), concerning the specific instance of MAKING X die_i (see §10), concerning MAKING as a specific type of causing (see G.E.M. Anscombe (1971)), and so on.

In §7, I consider those objections and argue that none of them is compelling.

I conclude that if theorists, from various domains, are interested in what MAKES what, then metaphysicians should pay attention to MAKING. In §§9-11, I focus on the ethical and legal domains and argue that both ethicists and lawyers are interested in what MAKES what; and so metaphysicians should pay attention to MAKING. I end, in §12, by discussing just how that attention should be paid.

2. First, some more about MAKING, itself.

Above, I said that, when $X V_I$'s

A MAKES $X V_I$ (e.g. Annie MAKES the window break)

just in case

$A V_T$'s X (e.g. Annie breaks the window).

While this schema neatly captures the instances of MAKING we saw above, it might appear to not capture all possible instances. For starters, it might seem to not capture those instances of A MAKING $X V_I$ for verbs that have no transitive use in English. *Smile*, *laugh*, *sneeze*, and *jump* are all such verbs and for each of them, “ $A V_T$ 'd X” is ungrammatical (e.g. #A smiled X).

Thankfully, that those instances of “ $A V_T$ 'd X” are ungrammatical is no barrier to our determining whether they are true. Judith Thomson considers *smile*:

...suppose you sign on as apprentice to the chief photographer. As he is about to photograph a batch of children for their school yearbooks, he notices that a child in the back row looks gloomy, this being undesirable in such photographs. “Smile that child up there in the back row,” he says to you; “I want to see if you can.” You know perfectly well that he does not want to see merely whether you can cause the child to smile—as, e.g., by paying a more experienced apprentice to cause the child to smile. You know he wants *you* to smile the child (1977: 129).

Accordingly, I see no problem with coining the transitive use of *smile* and stating that, when X smiles_I

A MAKES X smile_I just in case A smiles_T X.

And, in turn, doing the same for *laugh* (e.g. “Annie *laughed* the baby”), for *sneeze* (e.g. “the nurse *sneezes* the patient”) and for those other verbs that lack transitive uses.⁶

Just as we can coin transitive verbs from intransitive ones, so too can we coin intransitive verbs from certain transitive ones. For example, English recently started using *microwave* as a transitive verb —“Annie microwaved_T the potato”—and we might, in turn, coin the intransitive use without any loss of sense: “the potato microwaved_I in three minutes.” That means that we can say that, when the potato microwaves_I

A MAKES the potato microwave_I just in case A microwaves_T the potato.

The same goes for other verbs too. In particular, it's worth mentioning—since it's a central concept in ethics—that the same goes for *harm*. Like *microwave*, *harm* has only a transitive use in English (e.g. Annie harms_T Ben), yet we can happily coin its intransitive use and say that, when X harms_I

A MAKES X harm_I just in case A harms_T X.

(And, indeed, *harm* was historically used intransitively: e.g. “The men is fresh, too, and won't harm_I for a bit of exercise.”⁷)

But what goes for *microwave* and *harm* doesn't go for all those verbs that have no intransitive use in English. That is, for some transitive verbs, not only is “X V_I's” ungrammatical, but it's also nonsensical. For example, it might be the case that Annie steals_T the money or that Annie ponders_T the universe, yet “the money steals_I” and “the universe ponders_I” are nonsense. Since “Annie MAKES the money steal_I” and “Annie MAKES the universe ponder_I” are similarly nonsensical (what *could* they mean?), verbs like this are not counterexamples, but cautionary tales: we shouldn't coin with abandon.

Lastly, just as we can happily coin transitive verbs from certain intransitive ones—and *vice versa*—in order to capture certain instances of MAKING, so too can we coin verbs from certain adjectives in order to capture others. Repurposing a different example of Thomson's (*ibid.*), consider *blue*:

If your customer says, “I like the work you do better than the work your assistants do; so would *you* blue the inside of this bluebottle they just tattooed on my arm?” you would know perfectly well that he does not merely mean that you are to cause the inside of the bluebottle to be blue—as, e.g., by getting one of your assistants to blue it.

⁶ English also provides its own workaround for these instances of MAKING: A *makes*_T X V_I. Reconsider Thomson's story, replacing “smile the child” with “make the child smile” and notice how her conclusion still holds: you might make a child smile by telling them a joke, but not by having someone else tell them that joke. This workaround inspired my terminology.

⁷ “harm, v.” *OED Online*. Oxford University Press, March 2020.

And so we might say that, when X blues_I

A MAKES X blue_I just in case A blues_T X

(just as, when X reddens_I, A MAKES X redden_I just in case A reddens_T X).⁸ The same goes for other adjectives, too: the script doctor might *funny*_T the play—and, thus, MAKE the play funny_I—by adjusting the dialogue, but not by having someone else adjust it; and so on.

This is all to say that while there are certain linguistic barriers to our accepting the schema, above, those barriers do not run deep. We should ignore them.

Are there other barriers to our accepting the schema? Perhaps there are brute counterexamples to it, ones that aren't confined by the caprice of our language. For example, I said earlier that the fire doesn't melt the chocolate when it merely cuts power to the air-conditioning. But can't we simply stipulate that the fire nonetheless *schmelts*_T the chocolate (where, roughly, A *schmelts*_T X just in case A cuts power to whatever was preventing X from melting); and, having done so, doesn't that make it the case that the fire MAKES the chocolate *schmelt*_I? We sure can and it sure might, but this is no counterexample. After all, the relevant instance of the schema is that, when X *schmelts*_I

A MAKES X *schmelt*_I just in case A *schmelts*_T X,

and the case accords with it.

Ultimately, we should accept that, when X V_I's

A MAKES X V_I just in case A V_T's X.

3. By now, certain patterns, as to when this might MAKE that, have started emerging. For instance, if A causes X to V_I merely by causing B to (willingly) V_T X, then that is sufficient for it to be the case that A does not MAKE X V_I. But these patterns do not amount to a theory of MAKING. Moreover, I provide no theory in this paper.

Without such a theory you might think you are ill-equipped to determine exactly when this MAKES that. Not so. Since A V_T's X just in case A MAKES X V_I, we can determine whether A MAKES X V_I by

⁸ As evidence for how happily we can coin such verbs, consider: “[it] has a rose to brownish red or reddish brown peridium, yellow context that blues_I when exposed”; and “... a redhead who had blued_T part of her hair.” Both from: “blue, v.1.” *OED Online*. Oxford University Press, March 2020.

determining whether A V_T 's X. And, it turns out, we are all experts at determining whether A V_T 's X. We can flex that expertise by running through some subtle contrast cases:

LIGHTNING: Annie beats Ben and leaves him immobile. A freak thunderstorm rolls in and Ben is hit by a lightning bolt, which stops his heart.

As before, but this time there is no thunderstorm. Instead, the sun sets and Ben dies of exposure overnight.

Does Annie kill Ben?

FLOOD: heavy rain has raised the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house, instead enter Ben's, one door down.

As before, but this time Annie is out of town, is unable to block her doorway and so her house floods. Later, she pumps the waters out of her flooded basement. That water flows down the street and into Ben's house.

Does Annie flood Ben's house?

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Annie. Annie jumps off the track and the trolley instead hits and kills Ben who is tied to the track behind her.

As before, but this time it's not Annie on the tracks, but a barrier. Annie removes the barrier and the trolley instead hits and kills Ben who is tied to the tracks behind it.

Does Annie kill Ben?

DYE: Annie mixes some peroxide into Ben's pomade. Wily Ben realises and so he uses a different bottle instead. Unfortunately, the contents of the second bottle have (independently) fermented into a peroxide-like substance which turns Ben's hair white.

As before, but this time Ben doesn't realise at all. He unwittingly uses the pomade, his hair turns white.

Does Annie turn Ben's hair white?

WIND: the howling wind agitates the bear who jumps up and down on the tree branch until it snaps.

As before, but this time there is no bear. The wind buffets the branch until it breaks and falls to the ground.

Does the wind
break the
branch?

I trust that you answered as follows: Annie doesn't kill Ben when he's struck by the lightning, but does when he dies of exposure; Annie doesn't flood Ben's house when she blocks her doorway, but does when she pumps the water; Annie doesn't kill Ben when she jumps off the track, but does when she moves the barrier; Annie doesn't turn Ben's hair white when he instead uses the second, fermented bottle, but does when he unwittingly uses the bottle she tampered with; and the wind doesn't break the branch when it agitates the bear, but it does when it buffets the branch. Congratulations! You're an expert at determining whether $A V_T$'s X and, in turn, whether A MAKES $X V_I$.

4. We might be experts at determining, for various verbs V , whether $A V_T$'s X and, in turn, whether A MAKES $X V_I$ in a given instance, but this alone doesn't license my claim that each of these instances are tokens of the metaphysical type MAKING. There might be no such (natural) type: Annie might flood the village and Ben might break the window, yet it's possible that those actions share little more than the grammatical form of their description. What should we make of such MAKING-scepticism? Not too much.

As already mentioned, patterns have started emerging across the cases. For example, whatever the verb V , A 's merely causing B to (willingly) $V_T X$ is sufficient for it to be the case that A does not MAKE $X V_I$. Why should that be? An explanation is required.

Just as there are patterns emerging across the cases we've looked at, there are also patterns to be found *within* cases. Recall FLOOD: waters head towards Annie's house, Annie blocks her doorway and the waters instead enter Ben's house; yet Annie does not flood Ben's house. Yet this is nothing special about waters nor flooding_T. Suppose instead that the paint factory explodes and a wave of blue paint heads towards Annie's house, Annie blocks her doorway and the paint instead enters Ben's house; Annie doesn't *blue* Ben's house. Or suppose that a plague of locusts head towards Annie's greenhouse, Annie closes the door and the locusts instead enter Ben's greenhouse; Annie doesn't *destroy* Ben's crop.

Similar patterns are found within the other cases, too. These patterns all require explanation. I submit that the most natural explanation is that MAKING is a unified metaphysical type. It falls to the MAKING-sceptic to provide their own, better explanation; I've not seen one.

5. When metaphysicians have said something about MAKING, it has tended to be that it reduces to causing in one way or other. If they're correct, then their focus on causing is justified. Here I show that each proposed reduction of MAKING to causing fails. The final two proposals have more going for them, so I spend more time with them.⁹

Attempted reduction 1: MAKING is *direct* causing. A natural idea is that while many things might cause X to V_I, only certain things will do so *directly*, where A₁ directly causes X to V_I just in case it causes X to V_I, but doesn't do so by causing some A₂ to cause X to V_I. It then says that A MAKES X V_I just in case A directly causes X to V_I. For instance, suppose that Annie instructs Ben to throw the rock at the window, Ben throws the rock, the window breaks. Friends of this reduction would say that since Annie causes the window to break only by causing Ben to cause the window to break, Ben MAKES the window break, but Annie doesn't. But this reduction immediately fails since Ben too causes the window to break only by causing something else—namely, the rock—to cause the window to break. So it, in fact, falsely returns that Ben doesn't break the window.

Attempted reduction 2: MAKING is *intended* (*foreseen*) causation. This would say that A MAKES X V_I just in case A intentionally (*foreseeably*) causes X to V_I. Since inanimate objects can MAKE X V_I—boulders break windows, lightning burns trees, etc.—and I can (and have) broken things both unintentionally and unforeseeably, this is a non-starter.

Attempted reduction 3: MAKING is *morally loaded* causation. Shelly Kagan (1989) was struck by how we would never accuse a surgeon of killing her patient, even when the surgeon cuts into the patient's heart which causes the heart to stop, which causes the patient to die. As a result, Kagan suggested that whether A kills B is fixed, in part, by moral features of the situation. In particular, he suggested that the surgeon doesn't kill the patient because she acted permissibly. Mightn't we generalise this reduction (or some version of this idea) to MAKING in general? I don't think so. Firstly, whether I MAKE X V_I cuts across every moral distinction: I can break the window either permissibly or impermissibly, either justifiably or

⁹ Much of this section's content has been said before by Thomson (1987) and, less so, by David Lewis (1986).

unjustifiably, either with blame or without it, and so on. Secondly, we already saw in WIND, that wind might break a branch or it might merely cause that same branch to break, even though wind never does anything permissibly nor impermissibly (nor justifiably, etc.). Ultimately, this reduction is no good. (And what of Kagan's surgeon case? I suggest that what we *accuse* people of often differs from what they in fact do. After all, it would be natural for the surgeon herself to say "I feel awful, I killed a patient today.")

It's important not to confuse attempt 3 with a related idea. Namely, that it isn't that A's MAKING X V_I is *fixed* by certain moral facts but instead that it *fixes* certain moral facts. For instance, it could be that in order for to A infringe X's right against being V_I'd, A must MAKE X V_I. As it happens, I think this is correct (see §9), but for now it suffices to say that this is clearly not a reduction: entailment is not reduction.

Attempted reduction 4: MAKING is *causing*, as distinct from being *a cause*.¹⁰ Some authors—e.g. Peter Unger (1977) and Alex Kaiserman (2016)—take seriously the distinction between causing X to V_I and merely being a cause of X V_I-ing, where the mere causes of X V_I-ing are each, in Unger's terms, "a part of that which causes" X to V_I (184). For example, this view would have it that the speeding driver and the torrential rain are each causes of the shopfront breaking (when the car crashes through it) in virtue of their both being parts of what caused the shopfront to break—namely, the driver speeding in the torrential rain. If this is right and what causes X to V_I comes apart from the mere causes of X V_I-ing, then perhaps A MAKES X V_I just in case A causes X to V_I, after all.

Now, I have no idea whether this is the right way of thinking about causation, but I'm sure it isn't the right way of thinking about MAKING. After all, since it says that the speeding driver doesn't cause the shopfront to break, it would also say that the speeding driver doesn't break the shopfront, but clearly the speeding driver breaks the shopfront—he crashes straight through it!

Attempted reduction 5: MAKING is causing that is *insensitive* to other circumstances. This was David Lewis's idea. He said:

It may be that the causation depends on an exceptionally large and miscellaneous bundle of circumstances all being just right. If any little thing had been different, that cause would not have caused that effect. But sometimes causation is comparatively insensitive to small differences in the circumstances. When my strong recommendation causes ... deaths, that is comparatively sensitive causation—there are many differences that would have deflected the chain of events [and so I do not kill]. But if you shoot at

¹⁰ Thanks to the reviewer for this journal for suggesting this as a possible reduction.

your victim point-blank, only some very remarkable difference in circumstances would prevent his death [and so you do kill] (1986: 186).

And we might put the resulting reduction as follows: A MADE X V_1 just in case A (sufficiently) insensitively caused X to V_1 . Setting aside general worries—in particular, just what counts as sufficiently insensitively?—this reduction still returns the wrong verdict in FLOOD. Recall:

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her home, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house instead enter Ben's, one door down.

Annie insensitively causes Ben's house to flood—after all, once she blocks her doorway, it's determined that the waters will enter Ben's house—nonetheless, Annie doesn't flood Ben's house.

So insensitive causing isn't sufficient for MAKING nor, *a fortiori*, necessary and sufficient. We might wonder whether it's merely necessary. It isn't. Suppose that Robin fires an arrow through the crowds at Waterloo Station; it *just* misses every jostling passenger and breaks a window on the far side. Robin causes the window to break, but that he does so is sensitive to the precise speeds and trajectories of all those passengers that the arrow just misses. Nonetheless, Robin clearly breaks the window.

Attempted reduction 6: MAKING is causing *without intervening human acts*. Jennifer Hornsby (1980) says that A MAKES X V_1 just in case A causes X to V_1 and there exists no human that more proximately causes X to V_1 . For instance, when Annie instructs Ben to throw the rock at the window, Ben more proximately causes the window to break and therefore Hornsby says that Annie doesn't MAKE the window break. This reduction returns the wrong verdict in both FLOOD and LIGHTNING. Recall:

LIGHTNING: Annie beats Ben and leaves him slowly dying of his injuries. A thunderstorm rolls in and Ben is hit by a lightning bolt.

Even though there exists no human that more proximately causes Ben to die, Annie does not kill Ben. Similarly, even though there exists no human that more proximately causes Ben's house to flood, Annie doesn't flood Ben's house.

So that there exists no human that more proximately causes X to V_1 is not sufficient for MAKING, nor, *a fortiori*, necessary and sufficient. We might again wonder if it's merely necessary. It isn't: as we saw in

DYE, Annie turns Ben's hair white even when it's Ben himself who applies the pomade (and, in doing so, more proximately causes his hair to turn white).¹¹

6. Each proposed reduction of MAKING to causing fails. Induction alone thus gives us pause to examine the assumption that things do reduce in that manner. So what might motivate that assumption? As far as I know, no one has ever motivated it, but here's how I imagine the argument would go: "that A MAKES X V_I entails that A causes X to V_I, but the converse does not hold. That is, MAKING entails causing, but causing doesn't entail MAKING. So MAKING reduces to causing."¹²

What should we make of this argument? Its sort has been criticised before by Timothy Williamson regarding knowledge's claimed reduction to true belief. I simply repeat his criticism here—namely, that the argument doesn't generalise and, therefore, should be rejected.

Something's being red entails that it's coloured, but not *vice versa*. The argument structure, above, would thus have it that we can reduce things as follows:

$$red = coloured + X.$$

But there is no reason to think that the above equation can be solved in a non-circular way (and its being solved in a circular way—where $X = red$ or $red's\ wavelengths$, for instance—is no evidence that *red* reduces to *coloured*). So that something's being red entails that it's coloured (but not *vice versa*) is no reason to think that being red reduces to being coloured and, in turn, that A MAKES X V_I entails that A causes X to V_I (but not *vice versa*) is no reason to think that MAKING reduces to causing (Williamson 2000: 3).

So we should be suspicious of that argument. But that's only one argument and perhaps there are others that could be given for the same conclusion. So best to change tack: here's an argument that MAKING doesn't reduce to causing.

For the sake of argument, suppose that MAKING does, in fact, reduce to causing. That is, suppose that our concept MAKING reduces to our concept causing, just as, for example, our concept *bachelor* reduces to our concept *unmarried*. Reductions like these come with an epistemic restriction: in the case of *bachelor*,

¹¹ Jonathan Bennett (1988) combines the previous two attempts and suggests that A makes X V_I just in case both A insensitively causes X to V_I and there is no more proximate human cause of X's V-ing. However, this conjunctive reduction still returns the wrong verdict in FLOOD.

¹² Between them, Jerry Fodor (1970) and John Morreall (1976) given four reasons why causing doesn't entail MAKING.

it comes with the restriction that we cannot be more sure that X is a bachelor than we are that X is unmarried.¹³ The same goes for MAKING: if MAKING reduces to causing, then we can't be more sure that A MAKES X V₁ than we are that A causes X to V₁.

And normally, we won't be. If Annie shoots Ben dead then it's clear both that Annie kills Ben and that Annie causes Ben to die. Similarly, if Annie leaves the hose running in the garden, it's clear both that Annie floods the lawn and that Annie causes the lawn to flood. However, that won't always be so. Consider:

POISON: Ben is set to go on a cruise on Monday. The day before, Annie poisons Ben, sending him to hospital instead. The cruise departs without him and sinks on Tuesday with all hands lost. The poison kills Ben the following week.¹⁴

Annie kills Ben. Does she also cause Ben to die? On first glance, it's not at all obvious either way. (Note, the mere fact that she causes him to die *at a given time* is insufficient for it to be the case that she causes him to die. The paramedic's chest compressions might cause the patient to die at 3pm instead of 2pm, but the paramedic doesn't cause the patient to die.¹⁵)

What about on second glance? Isn't this just a case of, so-called, (*early*) *preemption*? Consider:

two assassins plan to kill Dictator. The first assassin will shoot and only if he misses will the second assassin shoot. The first shoots: it's a direct hit and Dictator dies.

The first assassin is said to preempt the second assassin since, by killing Dictator, he prevents the second assassin from doing so. Nonetheless, it's obvious that the first assassin causes Dictator to die. Since Annie's poisoning also prevents the cruise from causing Ben to die, why shouldn't we say of Annie what we say of the first assassin—namely, that she obviously causes Ben to die?

For starters, in poisoning Ben, Annie causes Ben to miss the cruise and, in turn, to live a week longer than he would otherwise have done. That fact alone sits ill with the claim that she causes him to die, and there is no corresponding fact *vis-a-vis* the first assassin. More importantly—and to repeat where I started—it just isn't obvious that Annie causes Ben to die, yet it is obvious that the first assassin causes Dictator to die.

¹³ There might be strange exceptions. For instance, I might tell my young nephew that George is a bachelor. My nephew might thereby come to know that George is a bachelor even though my nephew has no idea what it is to be a bachelor and, in turn, does not know that George is unmarried. I won't be looking at cases like this.

¹⁴ See Hart & Honoré (1959: 220 fn 1).

¹⁵ See Carolina Sartorio (2006).

Of course, we could convince ourselves that Annie does cause Ben to die; we could move from the clear verdict about the first assassin, via some theory of preemption or of causation, to that conclusion—and perhaps that is the right conclusion. But, regardless, that is a conclusion driven by theory and it doesn't undercut the fact that it's (much) clearer that Annie kills Ben than it is that she causes him to die. And that can't be if MAKING reduces to causing.

The same goes for other cases too. Consider:

BUSHFIRE: a bushfire is going to destroy both Annie's house and Ben's house. To protect her house, Annie opens the dam. Its waters quell the fire but also wash away Ben's house.¹⁶

Annie destroys Ben's house. Does she also cause Ben's house to be destroyed? As before, first glances are not definitive. (And, as before, the mere fact that Annie causes the house to be destroyed in a certain, watery way is not sufficient for it to be the case that she causes the house to be destroyed. I can cause the firework to explode quietly—by throwing a blanket atop it—without causing the firework to explode.)

Of course, we could again move from our conclusion about the first assassin, via some theory, to the conclusion that Annie does cause Ben's house to be destroyed. But, again, that reasoning doesn't undercut the fact it's (much) clearer that Annie destroys the house than it is that she causes it to be destroyed. And, again, that can't be if MAKING reduces to causing.

You might think I'm cheating. I have, after all, deliberately picked cases with complex causal structures and then used that complexity to elicit these unclear causal judgements from you (“so of course our causal judgements will be unclear,” goes the complaint). But the point isn't that those judgements are unclear, but that they are *less* clear than the corresponding MAKING judgements. From that perspective, this complaint makes no sense—after all, if MAKING *does* reduce to causing then that complexity should infect our MAKING judgements (at least) as much as our causing judgements. Yet, it doesn't.

Ultimately, we should be very suspicious of the assumption that MAKING reduces to causing: no wonder the attempts of the previous section failed.

7. Before concluding that MAKING is an independent, interesting notion—just like causing—I consider three possible objections to that conclusion.

¹⁶ From Hart & Honoré (1959: 219).

Objection 1: “if MAKING truly were an independent notion then it would, like causing, have a name. But it doesn’t: you just made one up.” Suppose, as this objection assumes, that the words we have somehow reflect reality (more realistically: the *concept* words we have somehow reflect reality). And it is true that “causing,” “cause,” and “causer” are all words in our language, while “MAKING,” “MAKE,” and “MAKER” are not.

But when we move from the generic words to specific ones, the opposite is true. There is no word for someone who causes people to die, but someone who MAKES people die is a *killer*; and there is no word for the action that people who cause people to die do, yet killers *kill*_T. Similarly, there is no word for the action of causing X to flood, but those that MAKE X flood, *flood*_T X, and we might call them *flooders*. Indeed, our language is littered with words for different instances of MAKING and for those who MAKE: *break*_T, *smelt*_T, *bake*_T, *cook*_T, *build*_T, *demolish*_T, *thief*, *builder*, *baker*, *smelter*, etc.

So while only causing has generic words, only MAKING has specific ones—thousands of them. At worst, that’s a wash for MAKING.

Objection 2: “MAKING is too messy and disorderly. What could possibly explain why Annie doesn’t flood Ben’s house in FLOOD but does in its variation? And what could possibly explain why Annie doesn’t kill Ben in LIGHTNING but does in its variation? And so on.” When we look back at all the cases, above, we might conclude that there is no order as to when A does and doesn’t MAKE X V_I. That smacks of defeatism at this early stage, but perhaps it isn’t. But even if it isn’t, disorder hardly leaves MAKING worse off than causing. Ted Sider describes things nicely:

Causation is a particularly unsavoury fundamental posit—at least if the posit is intended to closely match our ordinary concept of causation. It takes only a glance at the recent literature on causation to appreciate how arbitrary and baroque our ordinary concept of causation is (2011: 15-16).

And since this disorder has not dampened our interest in causing, nor should it dampen our interest in MAKING.

Objection 3: “whether A MAKES X V_I is affected by the mental states of agents. We saw this with DYE and its variations: Annie turns Ben’s hair white when he unknowingly uses the peroxide, but not when he does so knowingly. On the other hand, whether A causes X to V_I is a purely physical matter.”

This might all be true, but it isn't obvious how it's objectionable. Perhaps as follows: that MAKING is affected by the agents' mental states renders MAKING unsuited for playing the roles in theories that might otherwise be played by causing. I suggest one sure-fire way of determining whether MAKING *is* suited for playing a certain role is by looking at whether it does, in fact, play that role (see §§9,11); and if it does, then this objection falls flat.

8. We know what MAKING is, we know it when we see it, we have no reason to think it reduces to causing, and there is no in-principle objection to its being an independent, interesting notion. But that might all be besides the point if theorists are not interested in what MAKES what.

What is it for them to be interested? Roughly, if theorists in a given domain draw a distinction that aligns with, for the relevant actors A, whether A MAKES X V₁, then those theorists are interested in what MAKES what. If, for instance, what distinguishes those cases of intentionally bringing about the death of another that *are* homicide, from those that are not homicide, is whether the defendant killed the victim, then lawyers are interested in what MAKES what.

And if lawyers or ethicists or oncologists or ornithologists or whoever *are* interested in what MAKES what, then it falls to metaphysicians to pay attention to MAKING: the more theorists, the more that metaphysicians should pay attention (just how that attention should be paid, I turn to later). That is the conclusion of the preceding seven sections.

But are theorists interested in what MAKES what? As I said, this is an open, empirical matter that can only be settled one theory at a time. Obviously, there are too many theories to address them all here; I shall focus my attention on ethics and the law.

9. There are many ethical theories and their corresponding theorists are interested in different things. The *consequentialist*, for instance, thinks that what fixes whether or not a given action is permissible is whether its consequences were at least as good as any of the other actions available to the agent: if they were, then the action was permissible. Accordingly, consequentialists are interested in consequences.

Deontologists are also interested in consequences: that the child drowned, and wouldn't have done so otherwise, explains why it was impermissible for Annie to ignore the drowning child. But, in their quest

to capture our common-sense morality, deontologists hold that there are additional constraints on how we are permitted to act, ones that go beyond what is a consequence of what.

Here is Philippa Foot's (2002) classic illustration of such a constraint:

RESCUE-I: the tide is rising and five people are trapped in one cove, while a single person, Ben, is trapped in another. Annie can only make it to one of the two coves in time to rescue the people trapped inside. Annie rescues the five; Ben drowns.

RESCUE-II: the tide is rising and five people are trapped in a cove. To reach the five in time, Annie would have to run down Ben who is trapped in the access road. Annie rescues the five; Ben is run down.

In both cases, Ben's death and the survival of the five are consequences of Annie's action, but only in RESCUE-II does Annie act impermissibly. And Foot's explanation of this is as follows: only in RESCUE-II does Annie kill Ben and it's impermissible to kill one in order to save the lives of five others.

Of course, that principle is just a specific instance of a more general one, and so Foot would also endorse:

[IMP] other things equal, it's impermissible for A to $V_T B$ (/B's X) in order to prevent five others (/other Xs) from V_I -ing.

(E.g. other things equal, it's impermissible for Annie to flood Ben's house in order to prevent five other houses from flooding.)¹⁷

If Foot's right then MAKING cuts an important ethical distinction and, as I've put it, deontologists are interested in what MAKES what. Is she right? Her cases don't settle things since they don't differentiate between killing and, amongst other things, causing to die. But we can test [IMP] by looking at other cases and related principles—ones that better isolate MAKING.

For instance, we know that A's causing B to cause X to V_I is insufficient for A's MAKING X V_I . With that in mind:

¹⁷ I have been assuming that A MAKES B die_I just in case A kills_T B. I think that's right, but the editors of *Philosophers' Imprint* pointed out that this conflicts with a common usage of *kill*. Here is their example: if Godfather orders Henchman to kill Victim, then Godfather doesn't MAKE Victim die, but merely causes Victim to die (by causing Henchman to kill him), yet we might well say "Godfather killed Victim." What's going on here? I think what we mean is that Godfather *had* Victim killed. As evidence, suppose that Prosecutor asks for a list of all those killed by Godfather and suppose that Victim is included on that list; Prosecutor might well complain "no, I want a list of just the ones he literally killed"—and that complaint wouldn't make sense if Godfather literally killed Victim. That said, it's unclear to me just what the rules are for this (more figurative) use of *kill*. I suggest that it has something to do with crediting appropriate responsibility, since Godfather is, in some important sense, responsible for Victim's death. (Interestingly, the same goes for many verbs: we might say "Steve Jobs has built a million computers," since he is, in some important sense, responsible for their being built, even though only few Macs may rightly sit beneath a museum sign that reads "Built by Steve Jobs.") Anyway, "A MAKES B die_I just in case A kills_T B" should be understood as "A MAKES B die_I just in case A *literally* kills_T B."

RESCUE-III: the tide is rising and five people are trapped in a cove. Since it's dark, she will have to use her car's headlights in order to reach the five, the lights of which will enable Villain to pick an innocent target and shoot them dead. Annie rescues the five, Villain sees Ben by the lights of Annie's car and shoots him.

It's permissible for Annie to save the five and [IMP] agrees. On the other hand, the causal analogue of [IMP]—that it's impermissible to cause one to die in order to save five others—would wrongly predict that Annie acts impermissibly.

Just as certain common-sense permissibility verdicts are captured by the principle that it's impermissible to kill one to save five others, so too are other common-sense verdicts captured by the principle that it's impermissible to kill another in order to save oneself. Again, that principle is just a specific instance of a more general one:

[IMP2] other things equal, it's impermissible for A to V_T B (/B's X) in order to prevent herself (/her X) from V_I -ing.

As with [IMP], if [IMP2] is correct, then MAKING cuts an important ethical distinction. (Indeed, since the constraints on self-defence presumably parallel those on defence-of-others, [IMP] and [IMP2] come as a package.)

As it happens, we can test [IMP2] by the lights of cases we've already seen. Recall:

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Annie. Annie jumps off the tracks and the trolley instead hits and kills Ben who is tied to the track behind her.

FLOOD: heavy rain has raised the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house, instead enter Ben's, one door down.

It is permissible for Annie to jump off the tracks and to block her doorway and, since she neither kills Ben nor floods his house, these verdicts are rightly predicted by [IMP2].

We can further test [IMP2] by looking at cases inspired by the variations of JUMP and FLOOD, from earlier:

a runaway trolley is heading towards Annie and Ben. Ben is currently protected by a barrier. Unfortunately, the barrier is only big enough to protect a single individual. Annie moves the barrier in front of herself. The trolley kills Ben, while Annie is unharmed.

the floodwaters flood Annie's house while she's away. When she returns, Annie pumps the water out of her basement. That water flows down the street and into Ben's house.

Annie acts impermissibly in both these cases. [IMP2] also gets these right since Annie kills Ben and floods Ben's house.

I think it's no coincidence that common-sense permissibility aligns so neatly with MAKING. I think [IMP] and [IMP2] are correct.

But I've been ignoring the obvious objection. Thomson (1976, 1985) gave the following case:

BYSTANDER: a runaway trolley is heading towards five workers. The only way Smith, a bystander, can save them is by switching the tracks and diverting the trolley down a spur. Smith does so. B is tied to the spur and is killed by the trolley.

And Thomson said that it's permissible for Smith to divert the trolley yet, in doing so, Smith kills B. If Thomson is correct, then BYSTANDER is a counterexample to [IMP]. Everyone took Thomson to be correct, so I must respond.¹⁸ As it happens, we've already seen the makings of that response.

What's so interesting—so vexing—about BYSTANDER's structure is that the trolley that kills B was already set to kill the five and Smith's role was merely to divert that trolley from the five and onto B. Note how we have already seen something very similar to this structure: in FLOOD, the floodwaters were already set to flood Annie's house and Annie's role was merely to divert those waters away from her house and onto Ben's house. Since Annie doesn't flood Ben's house (doesn't MAKE Ben's house flood) in that case, why would it be that Smith does kill B (does MAKE B die) in BYSTANDER?

The natural answer points to the structural difference that remains between the two cases—namely, that in BYSTANDER it's Smith, a third party, who diverts the trolley, while Annie herself diverts the floodwaters. But we can happily do away with this difference:

Heavy rain has raised the level of the river to dangerous levels. Jones sees the waters heading towards a gated hamlet of five houses. The gate is broken, but Jones happens to be carrying some sandbags. Jones blocks the gate with the sandbags and the waters instead enter Ben's house, just down the street.

¹⁸ As it happens, one of the few people who doesn't take Thomson to be correct is Thomson herself. Her current view is that it is *not* permissible for Smith to divert the trolley (2008). Her argument for that conclusion requires that Smith "*makes* B pay the cost of death" which, to my mind, requires that Smith kills B.

Jones doesn't flood Ben's house. And if Jones doesn't flood Ben's house, then surely Smith doesn't kill B.¹⁹

There's another route, this time inspired by JUMP, to the same conclusion.²⁰ Consider:

MALFUNCTIONING: a self-driving trolley has malfunctioned such that it now travels to and runs down the nearest person on connected track. It is currently heading towards Bloggs. Bloggs jumps off the track, so the trolley changes course down a spur towards B who is tied to the track (B now being the nearest person to the trolley). The trolley kills B.

Bloggs doesn't kill B. Is there an important difference between MALFUNCTIONING and BYSTANDER? I don't think so and I suggest that only someone intent on there being one could find it. And, if there isn't a difference, then we must conclude that Smith doesn't kill B in BYSTANDER, after all. (An interesting question is just why everyone has accepted that Smith does kill in BYSTANDER.²¹ I have no answer, but I suggest it turns on the sort of pragmatic issue raised by Kagan's surgeon case.)

Once we understand killing as MAKING die—akin to MAKING flood, etc.—we realise that BYSTANDER is no counterexample to [IMP], after all. Ultimately, Foot was right about [IMP] and [IMP2] and MAKING cuts an important ethical distinction: ethicists *are* interested in MAKING.

10. A brief aside. If I'm right that MAKING cuts an important ethical distinction, you might wonder why the ethicists themselves don't have their own theories of MAKING. As it happens, theories have been proposed. Here are the three most popular and representative ones; I give a certain amount of detail since I will return to them later. (Following their proponents, I present them as accounts of *killing* specifically. I leave it to the reader to generalise them, should they desire.)

The most straightforward theory of killing has its roots in Alan Donagan's work (1979). It says that A kills B just in case B would not have died had A *done nothing at all*. Borrowing two cases from Jonathan Bennett (1998) to illustrate, suppose that in PUSH, A pushes B off a cliff to his death. If A had done

¹⁹ There are always further differences to point to, but each time the question is whether they *matter*. An example of one that surely doesn't matter is that Smith diverts the trolley, but Jones shields the hamlet. As Thomson herself says, "what difference could it be thought to make that a person deflects the trolley by putting a shield around the five instead of by throwing a switch? (2015: 119)" I suggest that the same will go for all further differences between the cases.

²⁰ See "The Bystander Doesn't Kill" for more on this route.

²¹ Interestingly, one person who never accepted that Smith kills in BYSTANDER was Foot, herself. She thought that whether A kills B is fixed by whether A starts the fatal sequence of events that leads to B's death and, in BYSTANDER, "we have the diverting of a fatal sequence and not the starting of a new one" (2002: 85).

nothing, then B would not have died and so A kills B. On the other hand, suppose that in STAYBACK, B is tumbling towards the cliff edge and A could easily catch B, but refuses; B tumbles to his death. B would still have died had A done nothing, so the theory rightly says that A doesn't kill B. (Of course, the theory stands or falls with its definition of *doing nothing at all*—does ignoring B count? does napping? sneezing?—but we can set that aside.)

Warren Quinn (1989) grounds his theory of killing in the distinction between positive and negative rights. To do so, he distinguishes between cases where B's death is a consequence of something A does (of A's *positive agency*) and cases where it's a consequence of something that A does *not* do, but might have done (of A's *negative agency*). He then says that A kills B just in case B's death is a consequence of A's action and that action is an instance of positive agency. A's pushing B in PUSH is an instance of positive agency (it's something A *does*) and thus Quinn says that A kills B. On the other hand, A's refraining from saving B in STAYBACK is an instance of negative agency (it's something A does *not* do) and so Quinn says that A does not kill B here. (We might ask of positive agency what we asked of *doing nothing at all*—does ignoring B count? does napping?—but we can again set such questions aside.)

Lastly, Bennett (1998) says that when B's death is a consequence of A's action ϕ , A kills B just in case the majority of the ways A could have acted, at the time A ϕ 'd, would not have resulted in B's death. In PUSH, the majority of the ways A could have acted (dancing a jig, taking a nap, etc.) would not have resulted in B's death and thus Bennett rightly says that A kills B. On the other hand, when A fails to catch B in STAYBACK, the majority of the ways A could have acted (dancing another jig, taking another nap, etc.) would still have resulted in B's death and thus Bennett rightly says that A does not kill B. (We can also set aside the question of just how these majorities are to be determined.)

The problem for these theories is that they all suffer from multiple, clear counterexamples. For example, these theories all give the wrong verdict in

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Annie. Annie jumps off the tracks and the trolley instead hits and kills Ben who is tied to the track behind her.

Quickly: had Annie not done anything then the trolley would have hit her instead of hitting Ben; similarly, Annie's jumping is an instance of positive agency; and lastly, so long as we suppose that Annie had to jump in a somewhat specific way to avoid the trolley, the majority of ways Annie could have acted would

not have resulted in the trolley missing her and killing Ben instead. So Donagan, Quinn and Bennett each incorrectly say that Annie kills Ben.

For similar reasons, the generalised versions of their theories each also give the wrong verdicts in LIGHTNING, FLOOD and DYE. Ultimately, these theories are no good.

11. Legal causation fixes whether A is legally responsible for a given outcome (at least, in Britain and the USA). This legal responsibility is a central element of the law, both criminal and civil: for instance, to be guilty of murder, it's necessary that A is legally responsible for the victim's death; to be found liable in tort, it's necessary that A is legally responsible for the damage; and so on.

As a primer to the concept of legal causation, recall the case from earlier where A poisons B's gin, B notices, heads to the shop to buy a replacement bottle and is killed by a car while crossing the road. Although A both intended B's death and B would not have died had A not acted as she did, A is not guilty of murder. Why? Because A is not a legal cause—is not legally responsible—for B's death.

What is this legal causation? The received view is that it's a technical legal concept, but one that aligns closely with causing. For instance, H.L.A. Hart and Tony Honoré say:

“Over a great area of the law [the courts] have, in using causal language, sought to apply a group of causal notions embedded in common sense” (1959: 123).

And Michael Moore says:

“The central idea that organises this book is that causation as a prerequisite for legal [responsibility] is intimately related to causation” (2009: xii).

“...what criminal law and the law of torts mean by ‘cause’ is what we ordinarily mean by ‘cause’ as we explain the world” (*ibid*: 5).

The received view is wrong. It is wrong because legal causation aligns with causing only to the extent that causing aligns with MAKING since legal causation is, in fact, MAKING. That is, A legally causes X to V₁ just in case A makes X V₁: A legally causes Ben to die (/the window to break/the house to be destroyed) just in case A kills Ben (/breaks the window/destroys the house).

Demonstrating this is simple enough: we need only examine a number of legal cases and see whether, each time, the court's judgement *vis-a-vis* whether the defendant legally caused the outcome—caused X to V_I—aligns with whether the defendant MADE X V_I.

And, as it happens, we have already examined many such cases: LIGHTNING, FLOOD, JUMP, DYE, BUSHFIRE and POISON (and, often, their variations) are all tidier version of legal cases or distinctions drawn therein.²² And, without exception, Annie MADE X V_I just in case her defendant counterpart was found by the courts to have legally caused X to V_I.

For instance, recall:

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house instead enter Ben's, one door down.

as before, but this time Annie is out of town, is unable to block her doorway and her house floods. Later, she pumps the waters out of her flooded basement. That water flows down the street and into Ben's house.

Annie does not flood Ben's house in the first case and the court found that her counterpart does not legally cause the corresponding property to flood. Conversely, Annie does flood Ben's house in the second case, and the court found that her counterpart did legally cause the corresponding property to flood.²³ The same goes for the other cases.²⁴

Ultimately, legal causation just is MAKING.²⁵ Lawyers are interested in what MAKES what.

²² LIGHTNING: *Alphacell Ltd v. Woodward* [1972] AC 824; *Doss v. Town of Big Stone Gap* [1926], 134, S.E. 563 (Va.); *State v Scates* [1858] 50 N.C. 409. FLOOD: see fn. 23 below. JUMP: (roughly) *Scott v. Shepherd* [1773], 96 Eng. Rep. 525. DYE: *Dalby* [1982] WLR 425, *Roberts* [1971] 65 Cr App R 95. BUSHFIRE & POISON: both discussed in Hart & Honoré (1959). For many more cases, see *Blackstone's Criminal Practice 2006* and Hart & Honoré (1959). Each time, the judgement aligns with MAKING.

²³ The reasoning: "if an extraordinary flood is seen to be coming upon the land the owner of such land may fence off and protect his land from it, and so turn it away, without being responsible for the consequences, although his neighbour may be injured by it" (*Whalley v Lancashire and Yorkshire Railway Co* (1884) 13 Q.B.D. 131). See also: *Nield v London and North Western Railway Company* (1874-75) L.R. 10 Ex. 4; and for locusts instead of floodwaters, *Greyvensteyn v Hattingh* [1911] A.C. 355

²⁴ Perhaps you worry that I have just picked the cases that fit my claim. As it happens, things went the other way: I started with those legal cases that the legal textbooks use to illustrate the intricacies of legal causation—largely, intricacies regarding how it comes apart from causing. I then realised that legal causation comes apart from causing in just those cases that MAKING also comes apart from causation. That realisation was, in part, the inspiration for this paper.

²⁵ There are certain *policy-driven* exceptions. For example, if A negligently starts a fire in New York State, which spreads from one building to many others, A is only legally responsible for the damages done to the first building. As Hart & Honoré say, "such rules ... have nothing whatever to do with causation ... [Instead] they represent a particular *policy* which a particular legal system has adopted" (1959: 84). (See also *Environment Agency v. Empress Car Co. (Abertillery) Ltd* [1999].)

For more on all this see "Legal Causation."

12. Ethicists and lawyers are both interested in what MAKES what. It remains an open, empirical question whether theorists in other domains are similarly interested in MAKING. You might be sceptical as to just how open that question truly is for other domains. For example, you might think that oncologists are surely not interested in MAKING since they are surely interested in what *causes* cancer. I think that such scepticism is rash: it takes only a glance at the commonly listed “causes” of cancer—smoking, exposure to radiation and asbestos, certain viruses, etc—to see that those “causes” are all things that *mutate* cells, that MAKE cells mutate. That alone is sufficient for the question *vis-a-vis* oncology to remain open. I think the same goes for the other domains, too.

That said, it doesn't much matter for my conclusion in this paper. My conclusion is that metaphysicians should start paying attention to MAKING and the fact that ethicists and lawyers are interested in what MAKES what is sufficient reason for metaphysicians to do so.

Perhaps, for now at least, you prefer to leave things to the ethicists and the lawyers. But I don't think that will work. Since MAKING is a general, metaphysical matter, attempting to theorise about it through a domain-specific lens is doomed to failure. The attempts we saw in §10 were examples of this, none clearer than Quinn's theory: he proposed that whether A MAKES X V_I is, in part, fixed by rights, but something as general as MAKING could never be fixed by something so domain-specific. Instead, we should be theorising about it from a general, domain-neutral standpoint.

And so metaphysicians should pay attention to MAKING. What would that look like? In short, by theorising about MAKING in just the same way they've been theorising about causing: asking whether omissions can MAKE or whether MAKING is transitive, etc.; and, ultimately, seeking a theory of just when it is that A MAKES X V_I .²⁶

References

Anscombe, G. E. M. 1971. *Causality and Determination: An Inaugural Lecture*. Cambridge: Cambridge University Press.

Bennett, Jonathan. 1988. *Events and Their Names*. London: Hackett Publishing Company.

— — —. 1998. *The Act Itself*. Oxford University Press.

²⁶ There might well be instances of MAKING that aren't instances of A MAKING X V_I . For example when Annie throws the rock at the window, she MAKES the window break_I: doesn't she also MAKE the rock break_T the window? There's reason to think so. After all, when Annie instructs Ben to throw the rock she not only causes the window to break_I, but she does so *by* causing the rock to break_T the window. And symmetry would have the same go for MAKING—namely, that Annie MAKES the window break_I by MAKING the rock break_T the window. Perhaps that's right. These are the sorts of issues that a theory of MAKING would seek to resolve.

- Donagan, Alan. 1979. *The Theory of Morality*. University of Chicago Press.
- Fodor, Jerry. 1970. 'Three Reasons for Not Deriving "Kill" from "Cause to Die"'. *Linguistic Inquiry* 1 (4): 429–38.
- Foot, Philippa. 2002. *Moral Dilemmas: And Other Topics in Moral Philosophy*. Oxford University Press.
- Hart, H. L. A., and Tony Honoré. 1959. *Causation in the Law*. Oxford University Press.
- Hornsby, Jennifer. 1980. *Actions*. Routledge and Kegan Paul.
- Kagan, Shelly. 1989. *The Limits of Morality*. Oxford University Press.
- Kaiserman, Alex. 2016. 'Causal Contribution'. *Proceedings of the Aristotelian Society* 116 (3): 387–94.
- Lewis, David. 1986. 'Postscripts to "Causation"'. In *Philosophical Papers Vol. II*, edited by David Lewis. Oxford University Press.
- Moore, Michael S. 2009. *Causation and Responsibility: An Essay in Law, Morals, and Metaphysics*. Oxford: Oxford University Press.
- Morreall, John. 1976. 'The Nonsynonymy of Kill and Cause to Die'. *Linguistic Inquiry* 7 (3): 516–18.
- Murphy, Peter, and Eric Stockdale, eds. 2005. *Blackstone's Criminal Practice 2006*. London: Oxford University Press.
- Quinn, Warren S. 1989. 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing'. *The Philosophical Review* 98 (3): 287–312.
- Sartorio, Carolina. 2006. 'On Causing Something to Happen in a Certain Way Without Causing It to Happen'. *Philosophical Studies* 129 (1): 119–136.
- . 2010. 'Causation and Ethics'. In *Oxford Handbook of Causation*, edited by Helen Beebe, Christopher Hitchcock, and Peter Menzies. Oxford University Press.
- Sider, Theodore. 2011. *Writing the Book of the World*. Oxford: Oxford University Press.
- Thomson, Judith Jarvis. 1976. 'KILLING, LETTING DIE, AND THE TROLLEY PROBLEM'. *The Monist* 59 (2): 204–17.
- . 1977. *Acts and Other Events*. London: Cornell University Press.
- . 1985. 'The Trolley Problem'. *Yale Law Journal* 94: 1395–1415.
- . 1987. 'Verbs of Action'. *Synthese* 72 (1): 103–22.
- . 2008. 'Turning the Trolley'. *Philosophy & Public Affairs* 36 (4): 359–74.
- . 2015. 'Kamm on the Trolley Problems'. In *The Trolley Problem Mysteries*, by Frances Kamm. The Berkeley Tanner Lectures. Oxford, New York: Oxford University Press.
- Unger, Peter. 1977. 'The Uniqueness in Causation'. *American Philosophical Quarterly* 14 (3): 177–88.
- Weatherson, Brian. 2007. 'Causation and Causatives'.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

The Bystander Doesn't Kill

1. When person B wouldn't have died had person A not acted as she did—when B's death is a *consequence* of A's actions—in some cases A *kills* B and in other cases she doesn't. Philippa Foot (2002a, 2002b) thought this difference mattered in that she thought that

IMPERMISSIBLE TO KILL: other things equal, it's impermissible to kill B in order to save five others.

But IMPERMISSIBLE TO KILL was soon rejected. Judith Thomson (1976, 1985) gave the following case:

BYSTANDER: a runaway trolley is heading towards five workers. The only way Smith, a bystander, can save them is by switching the tracks and diverting the trolley down a spur. Smith does so. B is tied to the spur and is killed by the trolley.

And Thomson said that it's permissible for Smith to divert the trolley yet, in doing so, Smith kills B.¹

If Thomson is correct, then BYSTANDER is a counterexample to IMPERMISSIBLE TO KILL. People took her to be correct, they rejected Foot's principle and have since sought an alternative: *much* ink has been spilt.

None of that is new. But what's less well known is that Foot denied that Smith kills B in BYSTANDER. She says that A kills B just in case A initiates the fatal sequence that results in B's death, and while this account likely won't stand much scrutiny, Foot's core idea is intuitive enough:

...we think of particular effects as the result of particular sequences, as when a certain fatal sequence leads to someone's death. This idea is implied in coroners' verdicts telling us what someone died of ... [W]e can pick out the particular sequence and ask who initiated it. If the subject died by poisoning and it was I who put poison into his drink, then I [killed him]; likewise if I shot him and he died of a bullet wound (2002b: 85).

Returning to BYSTANDER, Foot says that Smith kills B just in case Smith initiates the fatal sequence that results in B's death, yet the trolley was already fatally heading towards the five before Smith diverted it.

¹ Thomson's current view is that it *isn't* permissible for Smith to divert the trolley (2008). Her argument for that conclusion requires that Smith "*makes* B pay the cost of death" which, to my mind, requires that Smith kills B.

As Foot sees it, “we have here the diverting of a fatal sequence and not the starting of a new one” and so, by her lights, Smith doesn’t kill B (2002b: 85).²

If Foot is correct, then BYSTANDER is no objection to IMPERMISSIBLE TO KILL. But the problem for Foot is that every other account of killing disagrees with her and unequivocally says that Smith kills B in BYSTANDER. As evidence for this, here are three popular and representative accounts.

The simplest account of killing has its roots in Alan Donagan’s work (1979). It says that A kills B just in case B wouldn’t have died had A *done nothing at all*. To illustrate, suppose that in REMOVE, a runaway trolley is heading towards a barrier, which B is trapped behind; A removes the barrier and the trolley hits B. If A had done nothing, then B wouldn’t have died and so A kills B. On the other hand, suppose that in FAIL-TO-PLACE a runaway trolley is heading towards B and A could easily place a barrier in front of him, but refuses; the trolley hits B. B would still have died had A done nothing, so the account rightly says that A doesn’t kill B.

Of course, the account stands or falls with its definition of *doing nothing at all* (does ignoring B count? does napping? sneezing?), but for our purposes it doesn’t matter. Friends of this account take it as plain that B wouldn’t have died had Smith done nothing (since the trolley would instead have killed the five workers) and so the account says that Smith kills B in BYSTANDER.

Warren Quinn (1989) grounds his account of killing in the distinction between positive and negative rights. To do so, he distinguishes between cases where B’s death is a consequence of something A does (of A’s *positive agency*) and cases where it’s a consequence of something that A *doesn’t* do, but might have done (of A’s *negative agency*). He then says that A kills B just in case B’s death is a consequence of A’s ϕ -ing and ϕ -ing is an instance of positive agency. A’s removing the barrier in REMOVE is an instance of positive agency (it’s something A *does*) and thus Quinn says that A kills B. On the other hand, A’s refraining from saving B in FAIL-TO-PLACE is an instance of negative agency and so Quinn says that A doesn’t kill B here.

² Fiona Woollard (2008, 2015) develops an account of killing, inspired by Foot. But her account rejects what I take to be the best bit of Foot’s—that the bystander doesn’t kill.

We might ask of positive agency what we asked of *doing nothing at all* (does ignoring B count? does napping?), but we can again set such questions aside. Quinn takes Smith's diverting of the trolley to be a clear instance of positive agency and so Quinn says that Smith kills B in BYSTANDER.

Lastly, Jonathan Bennett (1998) says that when B's death is a consequence of A's ϕ -ing, A kills B just in case the majority of the ways A could have acted, at the time A ϕ 'd, wouldn't have resulted in B's death. In REMOVE, the majority of the ways A could have acted (dancing a jig, taking a nap, etc.) wouldn't have resulted in B's death and thus Bennett rightly says that A kills B. On the other hand, when A refuses to protect B in FAIL-TO-PLACE, the majority of the ways A could have acted (dancing another jig, etc.) would still have resulted in B's death and thus Bennett rightly says that A doesn't kill B.

So when we consider BYSTANDER by Bennett's lights, what matters is whether the majority of ways Smith could have acted would have resulted in B's death. And they wouldn't: only by diverting the trolley could Smith have brought about B's death and so Bennett says that, in doing so, Smith kills B.

So these accounts of killing agree that Smith kills B in BYSTANDER. But these accounts are all incorrect for they all give the wrong verdict in the following case, from Christopher Boorse & Roy Sorenson (1988):

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Jones, who can avoid the trolley only by jumping onto a ledge. Jones jumps onto the ledge and the trolley instead hits and kills B who is tied to the track behind her.

Boorse & Sorenson said many interesting things about JUMP and cases like it, yet one thing they omitted to point out—and which I take as a datum—is that Jones doesn't kill B here.

That the accounts above each get JUMP wrong is clear. Quickly: had Jones not done anything then the trolley would have hit her instead of hitting B; similarly, Jones's jumping is an instance of positive agency; and lastly, since Jones could only avoid the trolley by jumping onto that ledge, the majority of ways Jones could have acted wouldn't have resulted in the trolley missing her and killing B instead. So Donagan, Quinn and Bennett each incorrectly say that Jones kills B.

And that these accounts get JUMP wrong is noteworthy because of the structural similarity between JUMP and BYSTANDER. Namely, in both cases, what kills B would have killed someone else had the agent not acted as she did.

This is where we are: Thomson gave BYSTANDER as a counterexample to Foot's IMPERMISSIBLE TO KILL, but whether it's truly a counterexample turns on whether Smith kills B in BYSTANDER; Foot says she doesn't, all the accounts of killing say otherwise. But we know those accounts are wrong and, importantly, that they're wrong about a case akin to BYSTANDER—namely, JUMP. I argue that Foot was right, all along.

2. To do so, I start with a case similar to JUMP, but one where the agent plays a more obvious causal role in the victim's death:

MALFUNCTIONING: a self-driving trolley has malfunctioned such that it now travels to and runs down the nearest person on connected track. It is currently heading towards Bloggs. Bloggs jumps off the track, so the trolley changes course down a spur towards B who is tied to the track (B now being the nearest person to the trolley). The trolley kills B.

I say that the agent plays a more obvious causal role in the victim's death here than in JUMP, because of the difference in what the agents do to the respective trolleys. In JUMP, Jones doesn't affect the trajectory of the trolley, but in MALFUNCTIONING, Bloggs causes the trolley to change course away from her and into B. Put plainly, Bloggs *diverts* the trolley away from her and into B. Regardless, I take it as another datum that Bloggs doesn't kill B here.

It's noteworthy that in MALFUNCTIONING it seems as if B would still have died even if Bloggs hadn't been present in the scenario. That is, while B's death *is* a consequence of Bloggs's jumping off the track, it isn't a consequence of Bloggs's simply being present: if Bloggs hadn't been there at all, then B would have been the closest person to the trolley all along and the trolley would still have killed B. Might this account for why Bloggs doesn't kill B in these cases?

No. After all, we might fill in the case some more and suppose that the trolley is initially released by Villain because Villain hates Bloggs. Under these circumstances, Villain wouldn't have released the trolley that kills B had Bloggs not been there; and thus B's death is a consequence of Bloggs's being present. Nonetheless, Bloggs still doesn't kill B.

(I am reminded here of the superstition of the unlucky third light, from the Crimean War. It warns that three soldiers should never light their cigarettes together: the first soldier's cigarette draws the enemy sniper's attention, the second's cigarette fixes the sniper's aim, and the third soldier to light is shot dead. I've no idea whether there were actually any unlucky third lights, but suppose there were. The third soldier wouldn't have died had the first two soldiers not been there, lighting their cigarettes and attracting the sniper's fire. Regardless, the first two soldiers plainly do not kill the third; just as the hated Bloggs doesn't kill B, even though B wouldn't have died had she not been there.)

I said that in MALFUNCTIONING, Bloggs diverts the self-driving trolley. We might make that fact plainer:

DIVERSION: a runaway trolley is heading towards Bloggs, who is tied to the track. Bloggs switches the track and diverts the trolley down a spur. B is tied to the spur and is killed by the trolley.

Given that Bloggs doesn't kill in MALFUNCTIONING, does she kill in DIVERSION? I don't think so. In both cases, Bloggs diverts the trolley away from her and into B. In the first case, she does so by jumping off the track, in the second by switching the track. Yet I cannot see how that difference could possibly alter whether or not the respective agents kill the respective victims. (To hammer the similarity home, you might imagine that instead of jumping off the track in MALFUNCTIONING, Bloggs manoeuvres herself off the track by pulling a lever that controls a platform beneath her.)

To soften the surprise, we might briefly switch from *killing* to another causative verb, e.g. *flooding*. Suppose the river has burst its banks; A places sandbags in her doorway to protect her property; the water arrives and, unable to enter A's property, it instead enters and floods B's property, one door down. A

plainly doesn't *flood* B's house, even though she diverts the water away from her house and into his. I see no reason why what goes for flooding in this case shouldn't also go for killing in DIVERSION.³

So I think we should conclude that Bloggs doesn't kill B in DIVERSION. (Objection: "if Bloggs doesn't kill B in DIVERSION, doesn't that sit ill with the following case, where Martins surely kills B: *a trolley is safely on course for the station; Martins diverts the trolley down a spur where B is tied to the track; the trolley kills B.*" I set this aside until the final section.)

Now, suppose that things had gone a little differently:

BYSTANDER: a runaway trolley is heading towards Bloggs and four other workers, who are all tied to the track. They can only be saved by diverting the trolley down a spur, but this time the lever that switches the track is out of their reach. Smith, a bystander, switches the track and causes the trolley to divert down the spur. B is tied to the spur and is killed by the trolley.

BYSTANDER doesn't share the same structure as the two cases above. In both MALFUNCTIONING and DIVERSION, B's death is a consequence of Bloggs's actions and it's Bloggs who is initially under threat from the trolley. Conversely, in BYSTANDER, it's the bystander Smith who acts.

Regardless, does Smith kill B? No. Given that Bloggs doesn't kill B in DIVERSION, Smith doesn't kill B in BYSTANDER. How could the mere fact that the person who diverts the trolley into B is or is not under threat from that trolley affect whether or not she kills B by diverting the trolley? (Of course, it might (might!) bear upon the permissibility of Smith's diverting the trolley vs Bloggs's diverting the trolley. But since killing is independent of permissibility—and it must be if Foot's IMPERMISSIBLE TO KILL isn't to be circular nonsense—that it might bear upon permissibility is a separate matter.)

But suppose we were to insist that Smith's not being under threat makes all the difference. We would have to insist the very same goes for our previous cases too:

as in JUMP, but this time Jones doesn't see the trolley heading her way. Smith tackles Jones off the track and the trolley instead kills B, who is tied to the track behind Jones.

But Smith doesn't kill B here. (The bodyguard that tackles her VIP to save them from the assassin's bullet doesn't kill whomever is standing behind.)

³ See, of course, "MAKING Metaphysics" for more on this move.

Ultimately, given that Bloggs doesn't kill B in DIVERSION, I think we must conclude that Smith doesn't kill B in BYSTANDER.

3. But what of the objection I left open, above? It comprises two distinct worries:

- (i) can it be the case that Bloggs doesn't kill B in DIVERSION, but Martins does kill B when she diverts the trolley from the empty track onto B (henceforth EMPTY); more generally, can it be the case that whether A kills B depends on where the trolley was heading *before* A acted?
- (ii) (if so) what could explain *why* whether A kills B depends on such matters?

I take them in turn.

We know what Foot would say about (i). She would say that what matters is whether A initiates the fatal sequence of events that results in B's death. And she would point out that whether a sequence of events is fatal before A diverts it of course depends on where that sequence of events was heading: if it was heading towards a death, then it was already a fatal sequence which A merely diverts towards B (and so A doesn't kill B); if it wasn't heading towards a death, then A diverts a sequence of events that *becomes* fatal only in that diversion (and so A does kill B). And, she'd further point out, that DIVERSION is an instance of the former, while EMPTY is an instance of the latter.

Nonetheless, it would smack of convenience (for Foot and for me) if this contrast only showed up in DIVERSION vs EMPTY. Thankfully it doesn't and we've already seen it elsewhere. Recall:

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Jones, who can avoid the trolley only by jumping onto a ledge. Jones jumps onto the ledge and the trolley instead hits and kills B who is tied to the track behind her.

Jones doesn't kill B here. And notice how, *vis-a-vis* (i), DIVERSION is to EMPTY exactly what JUMP is to

REMOVE: a runaway trolley is heading towards a barrier, which B is trapped behind. A removes the barrier and the trolley hits B.

After all, what differentiates JUMP from REMOVE is also where the trolley was heading *before* the agent acted: in JUMP it was heading towards a (fatal) collision with Jones, while in REMOVE it was heading

towards a (harmless) collision with a barrier.⁴ Yet it's plain that this does make a difference since it's plain —just as it was plain in §1— that Jones doesn't kill B in JUMP, even though A does kill B in REMOVE.

The point is this: we know, *independently of everything said in §2*, that where the trolley was heading *before* A acts can make a difference to whether A kills B.

Which brings us to (ii): *why* does it make a difference? I have no answer. I suppose Foot thought she had an answer, or the beginnings of one, but her language of “diversions” and “fatal sequences” is too metaphorical—too metaphysically dark—to illuminate much.

Note, however, that we needn't draw the distinction between these cases with Foot's language. We can draw it counterfactually:

in DIVERSION and JUMP, the trolley that kills B *would* have killed someone had the agent not acted as they did,

in EMPTY and REMOVE, the trolley that kills B *wouldn't* have killed someone had the agent not as they did.⁵

And we can probably draw it in other ways, too. But regardless of how it's drawn, we should still wonder why it makes the difference that it does.

But that's just all to say that we need a better theory of killing. We already knew that the existing theories weren't up to scratch, since they all wrongly insist that Jones kills B in JUMP, but if IMPERMISSIBLE TO KILL is again a candidate moral principle, then there is renewed demand for a better understanding of just when A kills B. And it should be a candidate again since the bystander doesn't kill.

References

Bennett, J. (1998). *The Act Itself*. Oxford, New York: Oxford University Press.

Boorse, C., & Sorensen, R. A. (1988). Ducking Harm. *Journal of Philosophy*, 85(3), 115–134.

Byrne, T. (forthcoming). MAKING Metaphysics. *Philosophers Imprint*.

⁴ Remember: the fact that it's Jones moving herself off the track is merely an artefact of the case; it could just as easily be Jones's bodyguard tackling her off the track.

⁵ To be clear, I'm not suggesting that these counterfactuals distinguish every killing from every non-killing (since they certainly don't), but only that they distinguish between the cases mentioned. Nor am I suggesting that these counterfactuals align (/are equivalent) with Foot's fatal sequences; Foot never fleshed out her metaphors, so it's impossible to know what her account is equivalent to.

- Donagan, A. (1979). *The Theory of Morality*. University of Chicago Press.
- Foot, P. (2002a). *Virtues and Vices: And other essays in moral philosophy*.
- Foot, P. (2002b). *Moral Dilemmas: And Other Topics in Moral Philosophy*.
- Quinn, W. S. (1989). Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing. *The Philosophical Review*, 98(3), 287–312.
- Thomson, J. J. (1976). KILLING, LETTING DIE, AND THE TROLLEY PROBLEM. *The Monist*, 59(2), 204–217.
- Thomson, J. J. (1985). The Trolley Problem. *The Yale Law Journal* 94 (1985): 1395–415
- Thomson, J. J. (2008). Turning the Trolley. *Philosophy & Public Affairs*, 36(4), 359–374.
- Woollard, F. (2008). Doing and Allowing, Threats and Sequences. *Pacific Philosophical Quarterly*, 89(2), 261–277.
- Woollard, F. (2015). *Doing and Allowing Harm*. Oxford, New York: Oxford University Press.

Legal Causation

1. Compare:

Wanting to break Ben's leg, Annie sabotages the lintel of his front door. When Ben returns home and opens the door, the lintel collapses upon him, breaking his leg.

As above, but this time Ben notices the sabotage and instead of opening the door, he heads to buy a new lintel. On the way, he is hit by a car, breaking his leg.

In both cases, Ben's leg wouldn't have broken but for Annie's actions, but only in the first case will the courts (rightly) find Annie guilty of breaking Ben's leg. Lawyers say that this is because only in the first case is Annie a *legal cause* of Ben's broken leg.¹

But what is legal causation? There are two broad schools of thought. The first, *legal formalism*, holds that whether someone is a legal cause of some harm is a matter of amoral, descriptive fact. Just as, for example, whether Ben's leg wouldn't have broken but for Annie sabotaging the lintel is a matter of amoral, descriptive fact.

More specifically, the received view (amongst the legal formalists) is that while legal causation is a technical legal concept, it's one that aligns closely with our everyday, concept *common-sense causation* (henceforth, just "causation").² For instance, H.L.A. Hart and Tony Honoré say:

"Over a great area of the law [the courts] have, in using causal language, sought to apply a group of causal notions embedded in common sense."³

Why is the received view only that legal causation "aligns closely with" causation, as opposed to that it just is causation and that someone is a legal cause of a harm if and only if they cause that harm? Well, because there are many clear counterexamples to that stronger thesis. Here are three such counterexamples (each of which will prove useful later on):

WINDOW: Annie is throwing rocks at the wall of Ben's house. Seeing this, a passerby is inspired to throw his own rock. The passerby's rock smashes one of Ben's windows.

¹ Often also called *proximate cause*.

² Unfortunately, various concepts go by the name "causation." The law talks of both legal causation (/proximate causation) and factual causation (where X is a factual cause of Y if and only if Y wouldn't have occurred but for X). There is also common-sense causation, as used by the person on the street when they say that, e.g., "the firework caused the fire"—the study of this common-sense causation is a mainstay of contemporary metaphysics. To further add to the confusion, metaphysicians think of factual causation simply as one possible account of common-sense causation, as opposed to anything distinct. In this paper I will not talk of factual causation. I will talk of legal causation and I will always refer to it as "legal causation" (and cognates). Since there is no natural way of saying "X common-sense causes Y," I will reserve all unqualified uses of "causation" to refer to common-sense causation.

³ H. L. A. HART & TONY HONORÉ, CAUSATION IN THE LAW 123 (1959)

LIGHTNING: Annie beats Ben and leaves him immobile. A freak thunderstorm rolls in and Ben is hit by a lightning bolt, which stops his heart; he dies.⁴

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house instead enter Ben's, one door down.⁵

These are counterexamples since, each time, Annie plainly causes the harm suffered by Ben, yet in none of the cases is Annie a legal cause of that harm. For example, in the first case Annie plainly causes Ben's window to break (by inspiring the passerby to break it), yet the courts are in no doubt that Annie is not a legal cause of that broken window. The same goes for the other cases too and that is why the received view isn't that legal causation just is causation.

Since legal causation aligns closely, but not perfectly with causation, the standard formalist approach to legal causation (the one found in both legal textbooks and the rhetoric of court judgments) first asks whether the defendant was a cause of the harm in question and then provides various principles ("tests") for determining whether the defendant was additionally a legal cause of that harm.

One such principle asks if someone else willingly did something that was a temporally more proximate cause of the harm in question. If someone did, then that action would qualify as a so-called *novus actus interveniens* and "break the chain of causation" between the defendant's action and that harm. This principle neatly returns that Annie isn't a legal cause of the broken window in WINDOW since, in that case, the passerby's throwing of the rock is such a *novus actus interveniens*.

However, that principle is no help with cases like LIGHTNING, which lack any such third-party action. Accordingly, the standard formalist approach provides another principle, this time asking whether there was some intervening "act of God" that similarly amounted to a *novus actus interveniens*, "breaking the chain of causation"; if so, then the the defendant is not a legal cause of the harm. This principle neatly returns that Annie isn't a legal cause of Ben's death in LIGHTNING, since the lightning bolt is such an "act of God."

⁴ When appropriate, I reference the judgment(s) that corresponds to the case vignette; and when possible, I provide one from England & Wales and one from the US. In this instance, see *Alphacell Ltd v Woodward* [1972] AC 824 and *Hames v. State*, 808 S.W.2d 41, [1991] Tenn.

⁵ See *Nield v London and North Western Railway Company* (1874-75) L.R. 10 Ex. 4 and *Heins Implement v. Hwy. Transp. Com'n*, 859 S.W.2d 681 (Mo. 1993).

Many other principles exist for many other sorts of cases.⁶ Nonetheless, there are still cases that defy any useful generalisation. FLOOD is one such case: the courts agree that Annie isn't a legal cause of Ben's house flooding, yet there is no general principle that accounts for why that is so.⁷

The piecemeal nature of this formalist approach motivates a certain scepticism about legal formalism. Those sceptics—the *legal realists*—hold that, despite appearances, whether someone is a legal cause of some harm isn't a factual matter whatsoever. Instead, they hold that it is a moral matter. In turn, they hold that the explanation of why Annie isn't a legal cause of the harms in WINDOW, LIGHTNING, etc. is simple: namely, Annie isn't morally responsible for those harms (or, perhaps, that public policy is such that she shouldn't be held responsible—or whatever it might be).

I don't know whether the standard formalist approach warrants that scepticism—perhaps it does, perhaps it doesn't. Either way, I want to propose a different formalist account of legal causation. This account is formalist in that it holds that legal causation is a matter of amoral, descriptive fact. However, unlike the standard formalist approach, it holds that legal causation has nothing to do with common-sense causation whatsoever. Moreover, it is unified and principled.

Here is a simple primer for my account. Recall the pair of cases that I opened with:

Wanting to break Ben's leg, Annie sabotages the lintel of his front door. When Ben returns home and opens the door, the lintel collapses upon him, breaking his leg.

As above, but this time Ben notices the sabotage and instead of opening the door, he heads to buy a new lintel. On the way, he is hit by a car, breaking his leg.

Only in the first case is Annie a legal cause of Ben's broken leg. Why is that? I say it's because, only in the first case, does Annie *break* Ben's leg (you break those legs that you collapse lintels onto, but not those that someone else hits with their car). More generally, it says that the defendant is a legal cause of someone's broken leg if and only if the defendant breaks that leg. Similarly, it says that Annie isn't a legal cause of Ben's death in LIGHTNING because Annie doesn't *kill* Ben. More generally, it says that a defendant is a legal cause of someone's death if and only if the defendant kills that person.

⁶ See, for example, DAVID ORMEROD QC & DAVID PERRY QC eds., *BLACKSTONE'S CRIMINAL PRACTICE 2019* §A1 (2018) and W. PAGE KEETON, DAN B. DOOBS, ROBERT E. KEETON & DAVID G. OWEN, *PROSSER AND KEETON ON THE LAW OF TORTS* Ch. 7 (5th ed. 1984).

⁷ To be clear, I do not claim that the courts offer no principle, but only that it is not one that can generalise beyond the watery specifics of the case. E.g. "if an extraordinary flood is seen to be coming upon the land the owner of such land may fence off and protect his land from it, and so turn it away, without being responsible for the consequences, although his neighbour may be injured by it" *Whalley v Lancashire and Yorkshire Railway Co* (1884) 13 Q.B.D. 131. See also *Heins Implement v. Hwy. Transp. Com'n*, 859 S.W.2d 681 (Mo. 1993).

The job for the next few sections is to generalise that account and to define it—that is an exercise in metaphysics; afterwards I argue for it. To be clear, my conclusions in this paper are strictly metaphysical, concerning the nature of legal causation; I make no claims about whether and to what extent those conclusions should be adopted by the courts as a matter of practice. They are also limited: I do not, for instance, offer a rubric that will, for every possible case, immediately settle whether the defendant is a legal cause of a given harm.

2. In this section I introduce a metaphysical relation, akin to but distinct from common-sense causation. I begin with some background.

When, for example, a window breaks, it might be the case that

Annie caused the window to break,

and it might be the case that

Annie broke the window.

Often the two will go hand-in-hand—for instance, when Annie throws a rock at the window—but they needn't and we've already seen one such case where they don't:

WINDOW: Annie is throwing rocks at the wall of Ben's house. Seeing this, a passerby is inspired to throw his own rock. The passerby's rock smashes one of Ben's windows.

Annie causes the window to break (by inspiring the passerby to break it), even though she doesn't break it herself.

Similarly, when the house floods, it might be that Annie caused the house to flood and it might be that Annie flooded the house. These can also come apart and, again, we've already seen one case where they do:

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house instead enter Ben's, one door down.

Here, Annie doesn't flood Ben's house, even though she causes his house to flood by diverting the waters his way.

And, it's important to note, this is nothing special about Annie nor people, since the same holds more generally. When the chocolate melts, it might be that the fire caused the chocolate to melt and it might be that the fire melted the chocolate. These can come apart, too: if the fire cuts power to the air-conditioning, then the fire doesn't melt the chocolate, even though it causes the chocolate to melt (by causing the

summer air to melt it). However, having noted this generality, I will now set it aside and restrict my attention to only those cases involving people, since only people are candidates for legally causing something.

There's a pattern here. *Break*, *flood* and *melt* are all verbs that can be used both *intransitively* and *transitively*. A verb *V* is used intransitively in a given sentence when it doesn't allow a grammatical object (e.g. "the window broke"), while it is transitive when it does (whether explicitly, e.g. "Annie breaks the window," or implicitly, e.g. "stop stealing [cars]!"). If we mark these intransitive and transitive uses with the subscripts "I" and "T," then we can capture the pattern as follows. For thing *X* and verb *V*, when

X V_I's (e.g. the window breaks),

it might be the case, for person *A*, that

(i) *A* causes *X* to V_I (e.g. Annie causes the window to break)

and it might be the case that

(ii) *A* V_T's *X* (e.g. Annie breaks the window).

Now, it has long been understood that each instance of (i) shares something important with every other such instance; namely, that they are all instances of the same metaphysical relation—namely, common-sense causation. Accordingly, it has long been understood that we can theorise about what it is for, e.g., Annie to cause the window to break by theorising about causation in general (or, indeed, by theorising about what it is for, e.g., Annie to cause the house to flood). Such theorising is a mainstay of contemporary metaphysics.

Elsewhere, I've argued that the same goes for (ii) and that each instance of (ii) shares something important with every other such instance; namely, that they too are all instances of the same metaphysical relation—a relation that, while akin to causation, is distinct from it.⁸ My reasoning was simple (and, indeed, was exactly as you'd expect for a claim of this sort): stark similarities exist across all instances of (ii).

These similarities become apparent once we consider minimal variations of some of the cases, above. For instance, recall

WINDOW: Annie is throwing rocks at the wall of Ben's house. Seeing this, a passerby is inspired to throw his own rock. The passerby's rock smashes one of Ben's windows.

Even though Annie inspired the passerby to break Ben's window, Annie herself did not break that window. And it turns out that this is but a single instance of a general principle; namely, that if *A* causes *X*

⁸ See, of course, "MAKING Metaphysics."

to V_I merely by causing B to (willingly) V_T X, then that is sufficient for it to be the case that A does not V_T X. After all, if Annie had instead inspired the passerby to, e.g., repair Ben's window, then nor would Annie have repaired that window; or if she had instead inspired the passerby to open the window, then Annie herself wouldn't have opened it; and so on.

And recall FLOOD: waters head towards Annie's house, Annie blocks her doorway and the waters instead enter Ben's house; yet Annie does not flood Ben's house. This is nothing special about floodwaters nor floods. Suppose instead that the paint factory explodes and a wave of red paint heads towards Annie's house, Annie blocks her doorway and the paint instead enters Ben's house; Annie doesn't *red* Ben's house, even though she causes it to *red*. Or suppose that a plague of locusts heads towards Annie's greenhouse, Annie closes the door and the locusts instead enter Ben's greenhouse; Annie doesn't *destroy* Ben's crop, even though she causes it to be destroyed. (Generalising this similarity is a messy affair and since it isn't relevant here, I won't get into it.)

Other cases reveal further similarities and all these similarities require explanation. I concluded that the best such explanation is that every instance of (ii) is an instance of the same metaphysical relation. For lack of a better name, I christened that relation *MAKING*, where

A MAKES X V_I if and only if A V_T 's X.

(For example, A MAKES the window break if and only if A breaks the window; A MAKES the house flood if and only if A floods the house; and so on.)⁹

3. In that same paper, I also pointed out that while we might not be familiar with *MAKING*, as such, we are very familiar with its various instances and, in turn, we are experts at determining when A does and doesn't V_T X. We can flex that expertise by running through some subtle contrast cases.

Does Annie kill Ben?

Annie beats Ben and leaves him immobile. A freak thunderstorm rolls in and Ben is hit by a lightning bolt, which stops his heart; he dies.

⁹ Certain instances of this schema are ungrammatical. For instance, consider the verb *smile* which lacks a transitive form ("Annie smiled_T Ben" is ungrammatical): A MAKES X smile if and only if A smiles X. That this instance of the schema is ungrammatical tells us nothing about the world, but only reveals the caprice of our language. After all, there is no reason why *smile* couldn't have a transitive form and there is no reason why we couldn't say, for instance, "could you go and smile that child, he'll ruin the photo otherwise." Just as there's no reason why, for instance, you can *red* something, but there is no corresponding word for orange ("he orange_T the picture" is bad English). The same issue arises for those verbs that lack an intransitive form, such as *harm* (e.g. "Ben harms_I"). This again is an artefact of our language. For starters, *harm* used to be used intransitively (e.g. "The men is fresh, too, and won't harm_I for a bit of exercise"). In both cases, the correct response to the issue is to simply coin the missing verb forms. However, to avoid clunky language, I will restrict myself to verbs that do have both transitive and intransitive forms: *break*, *flood*, *kill*, *melt*, etc. See JUDITH JARVIS THOMSON, *ACTS AND OTHER EVENTS* 129 (1977) and "MAKING Metaphysics" §2.

As before, but this time there is no thunderstorm. Instead, the sun sets and Ben dies of exposure overnight.

Does Annie flood Ben's house?

Heavy rain has raised the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house, instead enter Ben's, one door down.

As before, but this time Annie is out of town, is unable to block her doorway and so her house floods. Later, she pumps the waters out of her flooded basement. That water flows down the street and into Ben's house.

Does Annie kill Ben?

A runaway trolley with enough momentum to kill a single individual is heading towards Annie. Annie jumps off the track and the trolley instead hits and kills Ben who is tied to the track behind her.

As before, but this time it's not Annie on the tracks, but a barrier. Annie removes the barrier and the trolley instead hits and kills Ben who is tied to the tracks behind it.

Does Annie break Ben's arm?

Annie sabotages the lintel of Ben's front door. When he returns home and opens the door, the lintel collapses upon him, breaking his leg.

As before, but this time Ben notices the sabotage and instead of opening the door, he heads to buy a new lintel. On the way, he is hit by a car, breaking his leg.

I trust that you answered as follows: Annie doesn't, as we've already seen, kill Ben when he's struck by lightning, but she does when he dies of exposure; Annie doesn't, as we've also already seen, flood Ben's house when she diverts the waters his way, but she does when she pumps the waters out of her basement; Annie doesn't kill Ben when she jumps off the track, but she does kill him when she removes the barrier; and, to return to where this paper began, Annie breaks Ben's leg when the lintel collapses upon him, but not when he's hit by the car. Congratulations: you're an expert at determining when A does and doesn't V_T X.

And that you're an expert is no surprise. Each of these different instances of MAKING—floodings, killings, breakings, etc.—are as familiar to us as, e.g., chairs, parties and the colour red; and just as we instinctively know whether something is a chair (or a party, or coloured red), so too do we instinctively know whether some action is, e.g., a killing. (To be clear, I do not claim that for *every* action we immediately know whether it's an instance of MAKING. I say more about those unclear cases in §7, but for

now it suffices to point out that nor do we immediately know for every thing whether it's a chair (consider beanbags) or whether it's red (consider salmon pink), etc.; and just as our lack of omniscience there doesn't undercut our general expertise *vis-a-vis* chairs, nor does our lack of omniscience here undercut our general expertise *vis-a-vis* MAKING.)

I am emphasising this expertise because I will, throughout the rest of this paper, be asserting of this-or-that case that it is (/isn't) an instance of MAKING—that Annie does kill Ben here, that she doesn't flood Ben's house there, and so on—and I don't want the grounds of those assertions to be mysterious. And they aren't mysterious: my grounds for asserting that a given action is an instance of MAKING are the same as my grounds for asserting that a given thing is a chair, or red, or whatever—namely, we know it when we see it.

4. With MAKING introduced, I can now state my account of legal causation:

A is a legal cause of X V_I -ing if and only if A MAKES X V_I .

And since A MAKES X V_I if and only if A V_T 's X, that amounts to the following:

A is a legal cause of X V_I -ing if and only if A V_T 's X.

For example, Annie is a legal cause of the window breaking if and only if Annie breaks the window; Annie is a legal cause of Ben dying if and only if Annie kills Ben; Annie is a legal cause of the house flooding if and only if Annie floods the house; and so on.

Earlier I claimed that this account of legal causation is amoral in that, by its lights, whether A is a legal cause of some harm isn't itself at all fixed by any moral features of the situation. That claim amounts to the claim that whether A V_T 's X is not at all fixed by any moral features of the situation.

Why think that claim is true? For two reasons. Firstly, whether A V_T 's X cuts across every possible moral distinction (and indeed every combination of them): A might, e.g., break the window either permissibly or impermissibly; A might flood the house either justifiably or unjustifiably; A might destroy the car either with blame or without it; and so on. And given that whether A V_T 's X cuts across every possible moral moral distinction, what moral features of the case could be left to fix whether A does or doesn't V_T X?

And secondly (and as mentioned above) it's not only people that can MAKE Xs V_I : fires can melt things, boulders can break things, and meteors can exterminate things, yet fires, boulders and meteors are each ineligible for moral evaluation. (After all, it makes no sense to say, e.g., that the fire melted the

chocolate permissibly or impermissibly, or that it was wrong for the meteor to exterminate the dinosaurs —just as it makes no sense to say that the fire melted the chocolate gleefully, etc.) And since fires, boulders, etc. are ineligible for moral evaluation, yet fires and boulders can MAKE Xs V_I (and can also, as we saw, cause X to V_I without MAKING X V_I), it cannot be moral features of the situation that are fixing whether they do indeed MAKE X V_I. Ultimately, whether A V_T's X is an amoral matter.

I also claimed earlier that my account of legal causation has nothing to do with (common-sense) causation (and, thus, that MAKING has nothing to do with causation). What does that claim amount to?

One way that MAKING might have something to do with causation is that MAKING might just *be* causation. However, we've already seen that that isn't so since whether A MAKES X V_I can come apart from whether A causes X to V_I (recall, for instance, when Annie causes the window to break by causing Ben to break it, yet doesn't break it herself).

More interesting is the possibility that A MAKES X V_I if and only if A causes X to V_I *in a particular way* (just as, for example, someone roasts a potato if and only if he cooks it in a particular way). The standard formalist approach to legal causation is like that: it says that A is a legal cause of X's V-ing if and only if A causes X to V_I in a particular way—i.e. without there existing a more proximate third-party that also caused the harm, and without there being an intervening “act-of-God,” and so on. So if MAKING is also causing in-a-particular-way, then my account of legal causation would just be some variation of the standard formalist approach.

However, I think it's clear that MAKING isn't just causing in-a-particular-way. I think that because every plausible specification of just what that particular-way is, admits of clear counterexamples. Elsewhere, I show that that is the case for an exhaustive range of specifications, but here I will consider only the most plausible ones.¹⁰

Proposed specification 1: MAKING is direct causation. A natural idea—plausibly one that can be traced back to Francis Bacon—is that while many things might cause X to V_I, only certain things will do so directly, where A₁ directly causes X to V_I if and only if it causes X to V_I, but doesn't do so by causing some A₂ to cause X to V_I. It then says that A V_T's X if and only if A directly causes X to V_I. For instance, recall when Annie inspires Ben to throw the rock at the window, Ben throws the rock, the window breaks. Friends of this specification would say that since Annie causes the window to break only by causing Ben to cause the window to break, Ben breaks the window, but Annie doesn't. But this specification

¹⁰ See “MAKING Metaphysics.”

immediately fails since Ben too causes the window to break only by causing something else—namely, the rock—to cause the window to break. So it, in fact, falsely returns that Ben doesn't break the window.

Proposed specification 2: MAKING is *intended* causation. This would say that A V_T 's X if and only if A intentionally causes X to V_I . Since inanimate objects can MAKE Xs V_I —boulders break windows, lightning burns trees, etc.—despite not doing so intentionally, this specification is a non-starter.

Proposed specification 3: MAKING is *foreseeable* causation. When A performs some action and, in doing so, causes X to V_I , sometimes it is foreseeable that A's action would cause X to V_I and sometimes it isn't. This specification thinks that this foreseeability makes all the difference and holds that A V_T 's X if and only if A foreseeably causes X to V_I . Regardless of whether we understand this foreseeability subjectively (that is, foreseeable to A) or objectively (foreseeable to the reasonable person), this proposal cannot be right. Recall

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house instead enter Ben's, one door down.

We can stipulate that, in blocking her doorway, Annie foreseeably (both subjectively and objectively) caused Ben's house to flood, but that doesn't alter the fact that Annie didn't flood Ben's house.

That's a case where A foreseeably causes X to V_I , yet doesn't V_T X. We might wonder whether there are also cases where A V_T 's X, but doesn't foreseeably cause X to V_I . There are. It's said that champagne was invented by mistake when French monks bottled some wine before it had finished fermenting, causing it to finish fermenting in the pressurised bottle and, in turn, to carbonate. If that's true, then the monk who bottled that wine caused it to carbonate without it being foreseeable to anyone that it would do so. Even so, the monk clearly carbonated the wine.

Proposed specification 4: MAKING is causation *without intervening human acts*. Jennifer Hornsby says that A V_T 's X if and only if A causes X to V_I and there exists no human that more proximately causes X to V_I .¹¹ For instance, when Annie instructs Ben to throw the rock at the window, Ben more proximately causes the window to break and therefore Hornsby says that Annie doesn't break the window break. This specification returns the wrong verdict in both FLOOD and LIGHTNING. Recall:

LIGHTNING: Annie beats Ben and leaves him immobile. A freak thunderstorm rolls in and Ben is hit by a lightning bolt, which stops his heart; he dies.

¹¹ JENNIFER HORNSBY, *ACTIONS* (1980).

Even though there exists no human that more proximately causes Ben to die, Annie doesn't kill Ben. Similarly, even though there exists no human that more proximately causes Ben's house to flood, Annie doesn't flood Ben's house.

Those are cases where A causes X to V_I without any intervening human acts, yet A doesn't V_T X. As before, we might wonder whether there are also cases where A V_T 's X despite there being intervening human acts. There are: if I lay a bear trap outside your door, which breaks your leg when you unwittingly step on it, then I break your leg even though you walked into it (and, in doing so, more proximately caused your leg to break).

5. In the previous two sections, I introduced MAKING—where A MAKES X V_I if and only if A V_T 's X—and I defended my claims that MAKING is amoral and that it has nothing to do with causation. I also presented my account of legal causation:

A is a legal cause of X V_I -ing if and only if A MAKES X V_I .

And since A MAKES X V_I if and only if A V_T 's X, that amounts to the following:

A is a legal cause of X V_I -ing if and only if A V_T 's X.

Unfortunately, there is no simple criterion for judging whether this account is correct. The natural idea—that it is correct just in case its verdicts align with the courts' judgments *vis-a-vis* legal causation—won't do since some of those court judgments are certainly incorrect.¹² After all, those judgments are judgments of fact (at least, they are for the legal formalist), yet the courts are fallible and so we can be sure they'll have made mistakes—we just don't know where. Clearly, it's not a mark against my account that it doesn't align with an incorrect court judgment of legal causation.¹³

Instead my approach will be more indirect—inductive rather than deductive. In this section, I will detail various elements of legal causation that my account aligns with. To do so, I will run through a number of cases (or, more precisely, case vignettes of the sort used throughout this paper), each of which represents a central, settled and uncontroversial element of legal causation. To my mind, these elements

¹² To regiment things, the courts will have 'judgments' of legal causation, while my account will have 'verdicts.'

¹³ There are also *policy-driven* judgments that would cause problems for this criterion. The classic example is New York State's "first building rule." That rule states that if A starts a fire, which spreads from one building to many others, destroying them all, then A is only a legal cause of the destruction of the first building. Since my account makes no such ordinal distinction and simply holds that A is a legal cause of the destruction of any building that A destroys, any judgment of legal causation that falls under the first building rule would not align with whether A destroyed the building; and, in turn, would fail the criterion, above. But as Hart & Honore say, such judgments "have nothing whatever to do with causation ... [Instead] they represent a particular *policy* which a particular legal system has adopted." H. L. A. HART & TONY HONORE, CAUSATION IN THE LAW 84 (1959) For England & Wales, see *Environment Agency v Empress Car Co. (Abertillery) Ltd* [1999] 2 AC 22.

constitute the central tenets of legal causation, but reasonable people can disagree over whether, and to what extent, that is so. Regardless, this section will detail elements of legal causation that my account aligns with.

On the other hand, in the next section I will detail some court judgments with which my account conflicts. It is a straightforward implication of my account that those judgments are incorrect. Ultimately, my view is that the alignment exhibited in this section renders those conflicts palatable, at the very least (I say more about this in that section).

I start with those cases that lack any complexity *vis-a-vis* legal causation—the “easy” cases—before turning to more complex ones. If A shoots B, which results in B’s death (in the usual way being shot results in one’s death) then A is a legal cause of B’s death and, of course, A killed B. If A drops a brick through B’s windshield then A is a legal cause of B’s broken windshield and, of course, A broke B’s windshield. On the other hand, if A merely stands back while someone else shoots B in the heart, then A didn’t kill B, nor is A a legal cause of B’s death. And if A stands back while someone else drops a brick through B’s windshield, then A didn’t break B’s windshield, nor is A a legal cause of the broken windshield. Unsurprisingly, these cases, and other easy ones like them, pose no problem for my account.

But what about “hard” cases that do involve causal subtleties? For example, those cases where the standard formalist approach to legal causation becomes piecemeal? We saw three in the introduction:

WINDOW: Annie is throwing rocks at the wall of Ben’s house. Seeing this, a passerby is inspired to throw his own rock. The passerby’s rock smashes one of Ben’s windows.

LIGHTNING: Annie beats Ben and leaves him immobile. A freak thunderstorm rolls in and Ben is hit by a lightning bolt, which stops his heart; he dies.

FLOOD: heavy rain has raised the level of the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie’s house instead enter Ben’s, one door down.

In each of these cases, Annie is not a legal cause of the harm in question despite being a cause of that harm and so the standard formalist approach must introduce further principles to capture those judgments. On the other hand, my account happily aligns with those judgments: in the first case, Annie doesn’t break Ben’s window and so my account rightly returns that Annie isn’t a legal cause of that broken window; in the second case, Annie doesn’t kill Ben and so my account rightly returns that Annie isn’t a legal cause of his death; and in the third case, Annie doesn’t flood Ben’s house and so my account rightly returns that Annie isn’t a legal cause of his flooded house. (We might also consider a variation of FLOOD alluded to earlier: suppose that a plague of locusts heads towards Annie’s greenhouse, Annie closes the door and so

the locusts instead enter Ben's greenhouse and destroy his crop. Annie isn't a legal cause of that destruction and my account agrees since Annie doesn't destroy Ben's crop.¹⁴)

Here are four more hard cases, each possessing some confounding causal factor.

Annie mortally wounds Ben and leaves him for dead. Before Ben dies of those wounds, a passerby with a grudge shoots Ben dead.¹⁵

Even though Annie mortally wounded Ben, Annie is not a legal cause of his death. My account agrees since, even though Annie mortally wounded Ben, Annie didn't kill Ben. That's a case where Annie isn't a legal cause of the death despite having inflicted mortal wounds and we might contrast it with the following case where the wounds inflicted are not themselves fatal:

Annie seriously wounds Ben. While those wounds by themselves are not fatal, they prevent doctors from operating on Ben's duodenal ulcer. Ben dies when the ulcer bursts.¹⁶

Here Annie is a legal cause of Ben's death. My account agrees since Annie killed Ben.

Another case:

Annie stabs Ben. At the hospital, an overworked doctor fails to notice that the knife has pierced one of Ben's lungs and, as a result, Ben is not given appropriate treatment. Ben dies.¹⁷

Even though the doctor's failure was a substantial cause of Ben's death, Annie is a legal cause of Ben's death. My account agrees since Annie killed Ben—despite the doctor's error. That case is often contrasted with the following:

Annie stabs Ben. Ben is treated at the hospital, where his wound largely heals. A doctor mistakenly prescribes Ben antibiotics, to which he has already had one allergic reaction. Ben dies of anaphylaxis.¹⁸

Despite the similarities in these two cases—a stabbing combined with a medical error, resulting in a death—Annie is not a legal cause of Ben's death here. My account agrees since this time Annie didn't kill Ben.

The previous four cases all concern death—a reflection of the relevant case law. Notice however, how the same contrast hold if we switch out death for, e.g., paralysis:

¹⁴ See *Greyvenstejn v Hattingh* [1911] A.C. 355.

¹⁵ See *State v Scates* (1858), 50 NC 409; *People v. Elder*, 100 Mich. 515, Mich. Supreme Court (1894).

¹⁶ See *R v McKechnie* (1992) 94 Cr App R 51.

¹⁷ See *R v Smith* [1959] 2 QB 35 and *Thompson v. Fox*, 326 Pa. 209, 192 A. 107 (Pa. 1937).

¹⁸ See *R v Jordan* (1956) 40 Cr App R 152 and *Brown v. N.Y. State Training School*, 285 N.Y. 37, 32 N.E.2d 783 (N.Y. 1941).

Annie stabs Ben. At the hospital, an overworked doctor fails to notice that the knife has pierced one of Ben's nerves and, as a result, Ben is not given appropriate treatment. Ben is paralysed.

Annie stabs Ben. Ben is treated at the hospital, where his wound largely heals. A doctor mistakenly prescribes Ben antibiotics, to which he has already had one allergic reaction. The anaphylaxis temporarily prevents blood from reaching his brain; Ben is paralysed.

Annie is a legal cause of Ben's paralysis in the first case, but not the second; and my account agrees since Annie paralysed Ben in the first case, but not in the second. (The subtlety of these contrasts makes stark an issue that has been lingering in the background: just what is it for A to kill B, or for A to paralyse B, or to break his arm, etc? I briefly turn to these questions in §7.)

As I mentioned in the introduction, one general principle of legal causation asks whether there exists an action (willingly) performed by some other party that "breaks the chain of causation" between the defendant's action and the harm in question—a *novus actus interveniens*. We saw an application of this principle in WINDOW, where Annie is not a legal cause of the broken window since the passerby's throwing the rock is a *novus actus interveniens*. Now, the parenthetical "willing" is important. Consider, for example:

Annie leaves a bear trap outside Ben's front door. Ben leaves for a work and, unwittingly, steps in the trap, breaking his leg.

As above, but this time Ben sees the bear trap. Nonetheless, and with nothing to lose, Ben decides to step into it, breaking his leg.¹⁹

In both cases, Ben is a more proximate cause of his broken leg (after all, he is the one who steps into the bear trap), yet only in the second case is he a willing cause of his broken leg. Accordingly, the principle returns that while Annie is a legal cause of Ben's broken leg in the first case, she isn't in the second. My account agrees since only in the first case does Annie break Ben's leg.

Actually, the principle is more nuanced, still. It's not sufficient that the third-party be a willing cause, they must also be an *informed* cause—they must appreciate what it is they're doing. Consider:

Annie secretly doses Ben's martini with a fatal dose of arsenic. Ben drinks the martini and dies.

As above, but this time Annie informs Ben of what she had done. Nonetheless, Ben still decides to drink the martini; he dies.

¹⁹ See *R v Latif* [1996] and *Rodgers v. Yellow Cab Co.*, 395 Pa. 412, 147 A.2d 611 [Pa. 1959].

In both cases, Ben is a more proximate cause of his death (after all, he is the one who pours the drink into his mouth); moreover, and unlike in the previous pair of cases, Ben willingly causes his death in the sense that he willingly pours the drink into his mouth. However, only in the second case, does Ben informedly cause his death since only in that case does Ben appreciate what it is he's doing—namely, drinking poison. Accordingly, the amended principle returns that while Annie is a legal cause of Ben's death (/his poisoning) in the first case, she is not a legal cause of his death in the second case. My account agrees since only in the first case does Annie kill (/poison) Ben.

6. In this section I detail three cases in which the verdict returned by my account conflicts with the judgment of the court. Each time, I briefly explain where, by my account's lights, the court's reasoning went wrong.

Conflict 1: *O'Neill v. Port Jervis*.^{20, 21} In response to building work, the city of Port Jervis erected a barrier that jutted out from the construction site, crossed the pavement and continued several feet into the road. A pedestrian, O'Neill, was prevented from continuing along the pavement by the barrier and, in order to continue on her way, O'Neill walked around the barrier and into the road. She was struck by an incoming vehicle and killed. On appeal, the courts held these facts did not constitute an in-principle barrier to the city of Port Jervis being a legal cause of O'Neill's death.

Since those same facts entail that the city of Port Jarvis didn't kill O'Neill (which it plainly didn't), those facts are sufficient for my account to return that the city was not a legal cause of O'Neill's death. Accordingly, the verdict returned by my account conflicts with the judgment of the court.

The court's reasoning is instructive: "...[if] the city personified had been there and forced O'Neill ... into the wheels of traffic, no doubt would exist in the minds of any one as to [legal causation]." The court then asked if it should be any different in the actual case where instead of being forced by the city personified, O'Neill was instead forced into the road by the "necessity" of doing so in order to continue on her way. They answered in the negative and judged accordingly. My account deftly explains where their reasoning went wrong since in the hypothetical case, the city personified would have killed O'Neill and that makes for all the difference in the world between that case and the actual one *vis-a-vis* legal causation.

²⁰ *O'Neill v. Port Jervis*, 253 N.Y. 423

²¹ Why the switch from case vignettes to actual cases? Because in the previous section it was instructive to screen away specific details of a particular case in order to emphasise the generalisable elements involved. Here, though, it is those specifics that drive the conflicts.

Conflict 2. *Hain v Jamison*.²² As a result of the farm-owners' negligence, one of their calves escaped its enclosure and wandered onto a nearby road, at night. While driving home, Hain spotted the calf, parked, and tried to move the calf off the road to safety. As she was doing so, another vehicle came around the bend and, unable to stop in time, collided with Hain, killing her. On appeal, the court ruled that the facts of the case did not constitute an in-principle barrier to the farm-owners being a legal cause of Hain's death.

Since those same facts entail that the farm-owners didn't kill Hain (which they plainly didn't), those facts are sufficient for my account to return that the farm-owners were not a legal cause of Hain's death.

It's again instructive to look at how the court arrived at its judgment. All parties agreed that if the escaped calf had wandered into the road and collided with Hain's car, killing Hain, then the farm-owners would have been a legal cause of Hain's death. The issue in contention was whether that case differed from the actual case *vis-a-vis* legal causation. The court framed that issue in terms of foreseeability: all parties agreed that if Hain had collided with the calf, then the manner of her death would have been foreseeable, but was the actual manner of her death similarly foreseeable (or, more precisely, could a jury reasonably conclude that it was)? The court concluded that it was similarly foreseeable and judged accordingly. By my account's lights, the court's mistake lay in framing the issue in that way. What explains why the farm-owners would have been a legal cause of Hain's death in that hypothetical case is not that Hain's death would have been foreseeable, but that the farm-owner's would have killed Hain—just as, e.g., I kill you if my tiger escapes and eats you. And that difference makes all the difference between the hypothetical case and the actual one *vis-a-vis* legal causation.)

Conflict 3: *Wallace*.²³ Wallace threw acid over van Dongen, inflicting horrific injuries that left van Dongen maimed, paralysed from the neck down and in terrible, incurable pain. Fifteen months later, and unable to take his own life, van Dongen travelled to Belgium, where he was euthanised in accordance with Belgium law. On appeal, the courts held that those facts of the case—in particular, how van Dongen had decided to travel to Belgium to be euthanised—did not constitute an in-principle barrier to Wallace being a legal cause of van Dongen's death (and, in turn, guilty of his murder).

Since those same facts entail that Wallace didn't kill van Dongen (which she plainly didn't), those facts are sufficient for my account to return that Wallace was not a legal cause of van Dongen's death.

²² *Hain v Jamison*, 28 N.Y.3d 524

²³ *R v Wallace* [2018] EWCA Crim 690.

The court's reasoning is instructive once again—although, the details given leave the ultimate crux of that reasoning a little unclear. The central issue (seemed to be) whether van Dongen's decision to end his life was voluntary in the sense required to render the resulting euthanasia a *novus actus interveniens*. On the one hand, the court considered *Kennedy*,²⁴ in which the defendant provided the victim the heroin with which the victim later overdosed. In that case, the court held that the victim's decision to inject was voluntary and, in turn, a *novus actus interveniens*. On the other hand, the court considered *Blaue*,²⁵ in which a stabbing victim refused a blood transfusion that would otherwise have saved her life (she was a Jehovah's Witness). There, the court held that the victim's religious beliefs rendered her decision to refuse the transfusion involuntary—since, in some sense, she had no choice but to refuse—and, in turn, not a *novus actus interveniens*. Returning to *Wallace*, given the “truly terrible situation” van Dongen was in, and since death was the only possible escape from that situation, the court held that this case fell in with *Blaue* as opposed to with *Kennedy*. They concluded: “there was nothing that could decently be described as voluntary in the decision by [van Dongen] to end his life” (76) and, therefore, his ending his life didn't amount to a *novus actus interveniens*. By my account's lights, and as was the case with *Hain*, the court's mistake lies in how they framed the issue: what explains why the defendant wasn't a legal cause in *Kennedy* is not that the victim voluntarily injected the heroin, but instead that the defendant didn't kill the victim; what explains why the defendant was a legal cause of the victim's death in *Blaue* is not that the victim had no choice but to refuse the transfusion, but that the defendant did kill the victim.²⁶ The extent to which the actions in question were voluntary is a red-herring. It's also a red-herring in *Wallace*: Wallace didn't kill van Dongen and that alone fixes that she is not a legal cause of his death.

In each of these cases, the verdict of my account conflicts with the judgment of the court—and while these are obviously not the only such cases, I believe they're representative. It's a straightforward implication of my account of legal causation that those judgments are incorrect. My view is that they are incorrect and, moreover, that we now know exactly where those judgments went wrong.

²⁴ *R v Kennedy* (No. 2) [2008] 1 AC 269.

²⁵ *R v Blaue* [1975] 3 All ER 446.

²⁶ In conversation, some people have questioned whether the defendant killed in *Blaue*, since the victim could easily (“easily”?) have averted her fate. I am reminded here of the scene in *Austin Powers* where Powers drives towards an anonymous henchmen in an extremely slow road roller. The henchman has minutes to move out of the way (and could easily have done so), but doesn't and is eventually squashed. The fact that the victim could easily have averted his fate has no bearing upon the obvious fact that Powers killed him.

7. A is a legal cause of X V_I-ing if and only if A V_T's X. What does that amount to, practically speaking? In short, it amounts to the following: whensoever we need to determine whether A is a legal cause of, e.g., the broken window, we only need ask whether or not A broke the window: if she did, then she is a legal cause; if she didn't, then she isn't. Thankfully, that question is almost always easy to answer: we know MAKING we when see it. (As further evidence for that ease, notice how there was no difficulty, in any of the cases considered above, in determining whether A V_T'd X: no difficulty in determining whether Annie flooded Ben's house in FLOOD; that the farm-owners didn't kill Hain; that Wallace didn't kill van Dongen, and so on.)

But what about those cases in which it isn't easy to determine whether or not A V_T'd X—what happens then? For example, consider *Scott v Shepherd* (“the famous squib case”):

[Shepherd] threw a lighted squib made of gunpowder into a market where a large concourse of people were assembled. The squib fell on the stall of Yates. To prevent injury to himself and Yates's goods, Willis took up the squib and threw it across the market where it fell on the stall of Ryal, who, to save his own goods, threw it across another part of the market, where it exploded and [blinded Scott].²⁷

By my account's lights, Shepherd is a legal cause of Scott's blindness if and only if Shepherd blinded Scott, yet whether Shepherd blinded Scott isn't clear: on the one hand, he is the one who lit and threw the squib, but on the other, he didn't throw it anywhere near Scott. And, without knowing whether Shepherd blinded Scott, my account is silent on whether Shepherd is a legal cause of Scott's blindness.

However, the fact that we are not currently in a position to determine whether, e.g., Shepherd blinded Scott doesn't entail that we will never be. Instead, (I suggest that) it's merely a consequence of how we currently lack a theory of MAKING and that, in turn, is merely a consequence of the fact that MAKING has been overlooked by contemporary metaphysics.

What would it be to have such a theory? One type of theory about some concept X is an analysis of X—i.e. a theory of the form “something is X if and only if it is Y and Z.” However, searches for analyses of this form are almost always futile (perhaps the only uncontroversial example is that something is a bachelor if and only if it is an unmarried man—yet even that raises uncomfortable questions about the Pope). I suspect that the project of analysing MAKING is similarly futile.

Instead, what we need is to start learning about MAKING, in general: whether that be by drawing connections between various instances of MAKING, or by discovering that every instance of MAKING is

²⁷ *Scott v. Shepherd* 96 Eng. Rep. 525 (K.B. 1773).

also an instance of some other phenomenon (or, conversely, that being an instance of some further phenomenon is incompatible with MAKING), and so on.

And even though metaphysicians haven't paid attention to MAKING, it's not a project that needs to start from scratch. After all, if legal causation is MAKING, then lawyers have been paying oblique attention to MAKING for centuries. For starters, they have documented endless cases—data points—to work from; they have also drawn all sorts of distinctions between those cases and it's no coincidence that some of those distinctions neatly align with MAKING (recall, for instance, the discussion of *novus actus interveniens* in §5). I have high hopes for this project.

I end this section by returning to *Scott v Shepherd* and (speculatively) illustrating how the theorising would go in that instance.²⁸ The general approach here is to start with cases where it's clear whether A MADE X V_I and to use those verdicts to help us in the unclear cases—moving from the known to the unknown. In that vein, consider this variation of a familiar case:

Villain blows the dam sending a torrent of water towards the open doorway of Annie's house. Annie blocks her doorway with sandbags and so the floodwaters instead enter Ben's house, one door down.

We already seen that Annie doesn't flood Ben's house, but what about Villain? He certainly does: he unleashed the torrent of water and that is sufficient for him to have flooded any house that the water ended up in—despite Annie's involvement.

Now consider a second variation:

As above, but this time Ben also blocks his doorway. The result is that the water is diverted from Annie's house, to Ben's and, ultimately, into Charlie's house, one further door down.

I see no reason to think that this variation should be any different to the above *vis-a-vis* whether Villain flooded the house, despite the additional “diverter.”

We can now ask whether there is a difference between the previous case and *Scott v Shepherd*. There are certainly many similarities: in both cases one person unleashes some threat (Villain, the water; Shepherd, the squib) towards one person, which is then diverted twice by a third- and fourth-party, before harming some victim. Given those structural similarities, there's good reason to think that what goes for Villain and Charlie *vis-a-vis* flooding should also go for Shepherd and Scott *vis-a-vis* blinding and, therefore, that Shepherd does blind Scot (and, ultimately, is a legal cause of Scott's blindness).

²⁸ I have already undertaken this sort of theorising in my discussion of *Blaue*, fn. 26.

But, like I said, that's speculative and it glosses over certain factors that would need to be considered were this theorising not merely illustrative.²⁹ Even so, it demonstrates how to resolve the tricky cases—those cases where it isn't immediately obvious whether A V_T'd X.

8. Just above, I suggested that the search for an analysis of MAKING is futile; it's worth mentioning a straightforward implication of that futility. There will exist (at least, in principle) pairs of cases for which the first is, e.g., a killing, the second is not a killing, and yet there is no generalisable principle that can be pointed to that explains why that is so. I say "in principle" because it's difficult to even contrive such a pair, but for illustration consider:

Annie slowly drives at Ben in a road roller. Ben could easily move out of the way, but doesn't. Annie runs Ben down. Annie kills Ben.

Annie pushes Ben, face first, into the swimming pool. Ben is an able swimmer and could easily rotate himself, but he doesn't. He remains face down and drowns. Annie does not kill Ben.

As I say, it's not clear that this is such a pair, but suppose it is and that there really is no generalisable principle that can be pointed to that explains why the first is a killing, but the second isn't. It would follow, therefore, that Annie is a legal cause of Ben's death in the first case, but isn't in the second case. In turn, it would follow that the only thing that could be pointed to to explain those verdicts is that Annie killed Ben in the first case, but didn't in the second—no further generalisation could be given. The existence of such cases may or may not be tolerable to the courts as a matter of practice, but I have no view on that and so, having raised the issue, I set it aside.

9. Consider the toy account of legal causation that I quickly set aside in the introduction (albeit, presented in now-familiar terminology):

A is a legal cause of X V_I-ing if and only if A causes X to V_I.

(E.g. A is a legal cause of the window breaking if and only if A causes the window to break.) Of course, that account is false, but if it weren't, no one could deny it was principled and unified; nor could they use it as an objection to legal formalism. Why is that? Because it starts with the general metaphysical relation causation and it says that legal causation aligns exactly with that relation—without codicil or amendment.

²⁹ One such factor is that, as described, Annie and Ben each had no choice but to divert the waters towards someone else's house, while perhaps Willis and Ryal could, if they'd been clearer-headed, have thrown the squib into a gap in the crowd. But, then again, would it have made a difference if there was some particular way that Annie *et al.* could have arranged their sandbags that would have instead diverted the waters into a ditch?

As principled and unified as that account would be, it still wouldn't immediately settle, for every possible case, whether the defendant is a legal cause of the harm in question. That's because facts about causation are themselves sometimes unclear. (Consider, for example, when A and B each simultaneously shoot Victim in the head; did A cause Victim's death?) Nonetheless, the account makes plain exactly where the answer would lie in such cases: not in legal theory, nor in moral theory, but in metaphysics.

The same goes for my account of legal causation. It says that A is a legal cause of X V_I-ing if and only if A MAKES X V_I. In doing so, it starts with a general metaphysical relation, the one I have called MAKING, and it says that legal causation aligns exactly with that relation—without codicil or amendment.

As we've seen, this account also won't immediately settle, for every possible case, whether the defendant is a legal cause of some harm. That's because facts about MAKING are also themselves sometimes unclear. Nonetheless, the account tells us exactly where to look in order to settle the matter: not in legal theory, nor in moral theory, but in metaphysics.

On Offsetting

1. There's a cafe in Raglan, New Zealand, that makes fantastic pancakes. Emily lives in London, England, and she hankers for those pancakes. As good fortune would have it, she's a private jet owning billionaire and so she flies to Raglan through the night, eats a short-stack for breakfast and then flies home. In doing so, Emily emits around 160 tonnes of carbon dioxide into the atmosphere (for reference, that's ten times the annual carbon footprint of the average American).¹

Those affected by climate change will have two distinct complaints against Emily in virtue of her trip to Raglan. The first is general: Emily spent vast sums on that trip and that money could have been put to much better use (promoting green policies, funding green initiatives, etc.). This is a general complaint because the same goes for every billionaire, even those who spend their fortunes without emitting much CO₂ at all. It's the complaint of "you could have done more to help," and it isn't my topic.²

The second complaint is more pointed: in making that trip, Emily emitted 160 tonnes of CO₂ and, as a result, Emily contributed to climate change and to the suffering of those affected by it: those affected by the bigger floods, the drier droughts and the hotter heatwaves that climate change will bring. (Exactly how Emily contributes to their suffering will be a central issue of this paper—I set it aside, for now.) I'll call this the *directed complaint*: it's the complaint that those affected by climate change have against Emily in virtue of the fact that, in emitting those 160 tonnes of CO₂, she contributed to climate change. The directed complaint is my topic.

Not everyone will agree that those affected by climate change have a directed complaint against Emily. They'll say that while Emily might have contributed to the suffering of those affected by climate change in the future, her contribution to the suffering of any one individual will be so marginal that it will be imperceptible to that individual. And then they'll say that if no individual can even perceive Emily's contribution to his/her suffering, then no individual can have a directed complaint against Emily. And then

¹ There are many types of greenhouse gas: carbon dioxide, nitrous oxide, methane, etc. CO₂ receives the most attention because it accounts for about 80% of our greenhouse gas emissions. Since "CO₂" also lends itself to easier expression than "greenhouse gas," I similarly focus on CO₂. However my conclusions go for all our greenhouse gas emissions.

Regarding carbon footprints, see <https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks>.

² Instead see, e.g., Peter Singer (1972).

they'll conclude that if no individual has a directed complaint against Emily, then there can be no aggregated directed complaint, either.

This cuts across various issues—of ethical aggregation,³ of imperceptible harms⁴—but I'm going to sidestep them all here since the debate I'm interested in lies downstream of this one. Accordingly, I'm going to assume that those affected by climate change do have a directed complaint against Emily. That is, I'm going to assume that those affected by climate change have a complaint against Emily in virtue of that fact that she contributed to climate change that weekend.

2. John Broome (2012) says that if Emily offsets the CO₂ emissions from her trip, then those affected by climate change will have no directed complaint against her. As I understand the term, the action A1 performed by emitter *E* (and all emitters in this paper will, like Emily, have names beginning with E) *offsets* the CO₂ of some other action of hers A2 just in case the total amount of CO₂ in the atmosphere is no greater than it would have been had E performed neither A1 nor A2.

There are various ways Emily might offset her CO₂ emissions for a given activity (more on those ways later), but one way is by capturing the same amount of CO₂ from the atmosphere that performing the activity emitted into the atmosphere. As it happens, one of Emily's factories can—at the press of a button—be reconfigured to capture CO₂ and so when Emily returns from Raglan, she presses that button and captures 160 tonnes of CO₂ from the atmosphere.⁵ In doing so, Emily offsets the emissions from her trip since the total amount of CO₂ in the atmosphere (+160 tonnes and then -160 tonnes) will be no higher than it would have been had Emily neither taken her trip nor captured that CO₂. And once Emily's emissions are offset, Broome's claim is that those affected by climate change will have no directed complaint against her in virtue of her taking that trip to Raglan.

³ See, e.g. Thomas Scanlon (1998).

⁴ See, e.g., Shelly Kagan (2011).

⁵ This is also called *direct air capture* since the CO₂ is captured directly from the atmosphere. Not to be confused with *point of source capture*, which, e.g. uses filters to capture the CO₂ from smokestacks before it enters the atmosphere—which I won't be talking about in this paper.

Now, Broome's claim is a general one and he thinks that what goes for Emily goes for all of us: we are all contributing to climate change, that contribution gives those affected by climate change a directed complaint against us, yet, if we offset, then they will have no such complaint. But when it comes to assessing Broome's claim, Emily's trip to Raglan is a usefully simple case, for two reasons. Firstly, her trip was extreme in that she emitted a particularly large quantity of CO₂ for particularly frivolous reasons and so if offsetting gets her off the hook then we can assume that it will also get us off the hook when we, e.g., heat our homes.

Secondly, in the case of Emily's trip to Raglan, Emily *herself* emitted the 160 tonnes of CO₂ (in that, there was no more proximate human emitter), but that will rarely be the case when it comes to our own contributions to climate change: I don't emit any CO₂ when I run the kettle, the power plants do; I don't emit any CO₂ when I take a flight, the airline does, etc. (Indeed, for most of my contribution to climate change, I don't even cause the CO₂ to be emitted since the electricity generated by the power plant, or the flights flown by the airline, aren't sensitive to my demand—at least not in any simple way.⁶ Moreover, much of my carbon footprint is out of my control: the police will patrol and emit CO₂ regardless of what I do and even whether I exist.)

Just as Emily herself emitted those 160 tonnes of CO₂, so too did she capture those 160 tonnes herself, but this will likely never be the case for any offsetting I do: I might, e.g., pay someone else to capture CO₂ on my behalf, but I won't ever capture any CO₂ myself. That Emily herself emits and captures the CO₂ in question simplifies matters, so I focus on her case for now—I'll return to our everyday, more indirect cases later.

So Broome says that since Emily offset the emissions from her trip to Raglan, those affected by climate change have no directed complaint against her. If Broome is right then, since Emily loves pancakes and hates injustice, this is great news for Emily. (And if Broome's general claim is right, then it's great news for all us, as well.)

⁶ See Kagan (*ibid.*).

3. In fact, it's *suspiciously* great. Robert Goodin (1994) compares it to the Catholic practice of selling indulgences, where medieval sinners paid the Church in order to escape the consequences of their sins. Elizabeth Cripps (2016) says it's akin to disrupting your neighbour with a loud party, having first paid someone else to cancel their own noisy event. The website www.cheatneutral.com mocks the practice with the following £2.50 service: "When you cheat on your partner you add to the heartbreak, pain and jealousy in the atmosphere. Cheat Neutral offsets your cheating by funding someone else to be faithful and not cheat. This neutralises the pain and unhappy emotion and leaves you with a clear conscience." Environmentalist George Monbiot (2006) calls it "pernicious and destructive nonsense."

Here's one way of getting at what they all find so objectionable about offsetting, from Caspar Hare (2013). Consider:

BARRELS (CAPTURED): emitter Esha's factory is on the west side of the river, her waste processing plant is on the east side and, upstream, is another factory that regularly disposes of its waste into the river. Esha's factory throws a barrel of its toxic waste into the river. Later, Esha's waste processing plant removes a barrel from the river that the upstream factory had disposed of. (Why does Esha do this? Perhaps the cost of transporting her barrel across the river is prohibitive.) Esha's barrel pollutes victim Veena's farm downstream, ruining 1/3 of Veena's crop. (Note: if both barrels had polluted Veena's farm, then 2/3 of her crop would have been ruined.)

Since the total amount of toxic waste in the river is no higher than it would have been had Esha neither thrown her waste into the river nor captured the other factory's barrel, Esha has offset her waste.

The objection then proceeds as follows:

(P1) even though Esha offset her waste, Veena (obviously!) has a complaint against Esha in virtue of the fact that Esha threw her barrel into the river,

(P2) what goes for the offsetting of waste thrown into the river goes for the offsetting of CO₂ emitted into the atmosphere,

(C) even when Emily offsets her CO₂ emissions, those affected by climate change (obviously!) have a complaint against her in virtue of the fact that Emily emitted CO₂ into the atmosphere,

where that complaint is the directed complaint. (C) is the denial of Broome's claim.

Much of this paper will be about (P2), but for now let's focus on (P1); it's true, but it's important to be clear about exactly why it's true.

To that end, consider a simpler case, one without any offsetting:

BARRELS*: emitter Esha* throws a barrel of toxic waste into the river. The barrel pollutes victim Veena*'s farm downstream, ruining Veena*'s crop.

Veena* clearly has a directed complaint against Esha*, but what is the nature of that complaint? I think Veena* actually has two independent grounds for such a complaint against Esha*:

- (i) Esha* *ruined* Veena*'s crop,
- (ii) Esha* *increased the risk* that Veena*'s crop would be ruined.⁷

In the next two sections, I'll explain what (i) and (ii) amount to—what *ruining* something and what *increasing the risk* of something each amount to—and I will explain why they are each sufficient for it to be the case that Veena* has a directed complaint against Esha*. But until then we can take (i) and (ii) at face value since, for now, it suffices to show that they are independent in that either one could hold without the other. (A further question is how these complaints interact when they both hold—as they do in BARRELS*. Perhaps (i) trumps (ii), perhaps they aggregate somewhat. However, it doesn't bear on anything in this paper, and so I set the question aside.)

To see that you can have (ii) without (i), consider:

Just above Veena*'s farm, the river flows through ten sluice gates that dam the river. The first gate directs the flow towards Veena's farm and anything that passes through it will head in that direction, while the other nine gates direct the flow elsewhere. A highly toxic (and cohesive) algae bloom is floating downstream. Before it reaches the

⁷ It might, in this case, be more natural to talk of counterfactual dependence instead of risk—i.e. Veena*'s crop wouldn't have been ruined if Esha* hadn't thrown the barrel into the river. However we'll soon be considering cases that can only be captured by risk talk and so it makes sense to put things that way from the off.

gates, Ferryman sails his ferry in front of the last five gates, blocking them. The river now flows through only the first five gates and, as it happens, the algae bloom passes through the first gate, where it poisons Veena*'s farm and ruins her crop.

While Ferryman didn't ruin Veena*'s crop it is the case that he increased the risk that her crop would be ruined. Accordingly, while Veena* cannot complain that

(i) Ferryman ruined Veena*'s crop,

she can complain that

(ii) Ferryman increased the risk that Veena*'s crop would be ruined.

And so you can have (ii) without (i).

And to see that you can have (i) without (ii), consider any familiar causal preemption case. If, for example, Pyro sets fire to Veena*'s crop just before Esha*'s barrel of toxic waste arrives, then while Veena* can complain that

(i) Pyro ruined Veena*'s crop,

(after all, he burnt her crop), she can't complain that

(ii) Pyro increased the risk that Veena*'s crop would be ruined,

since the crop was certain to be ruined, regardless of whatever Pyro might have done.

So (i) and (ii) are independent in that either one could hold without the other. With that in mind, we can return to (P1) and BARRELS (CAPTURED). Recall that if Esha had had nothing to do with the river that day (had she neither thrown her barrel into the river nor captured the upstream barrel from the river) then 1/3 of Veena's crop would still have been ruined, since it would have been ruined by the upstream barrel. Because of this, Veena cannot in fact complain of Esha that

(ii) Esha increased the risk that 1/3 of Veena's crop would be ruined.

After all, 1/3 of Veena's crop would have been ruined even if Esha had had nothing to do with the river.

So that Esha offset her pollution doesn't count for nothing, but it also doesn't count for very much. It doesn't count for very much since even though Veena cannot complain that

(ii) Esha increased the risk that 1/3 of Veena's crop would be ruined,

she can nonetheless complain that

(i) Esha ruined 1/3 of Veena's crop, since, despite everything else, it was still Esha's barrel of waste that did the damage. So that is why (P1) is true: even though Esha offset her pollution (and even though that meant that she didn't increase the risk that 1/3 of Veena's crop would be ruined), that doesn't change the fact that she nonetheless ruined 1/3 of Veena's crop; and that she ruined 1/3 of Veena's crop is sufficient for it to be the case that Veena has a directed complaint against her.

(P2) says that the same goes for Emily's offsetting of her CO₂ emissions. The rest of this paper argues that that isn't so. Here's the plan. Over the next two sections, I explain what facts like these amount to:

- (i) Esha* ruined Veena*'s crop,
- (ii) Esha* increased the risk that Veena*'s crop would be ruined.

And, in turn, why they are each sufficient to give Veena* a directed complaint against Esha*. In §§6-7, I argue that the only possible grounds of the directed complaint that those affected by climate change have against Emily (in virtue of her taking her trip to Raglan) is that, in taking that trip, Emily increased the risk that they would be affected by climate change. In §§8-9, I argue that since Emily offset the emissions from that trip, it isn't the case that she increased the risk that they would be so affected. In §§10-11, I tie up some loose ends and, ultimately, I conclude that Broome was right, after all: if we offset our emissions, then those affected by climate change will have no directed complaint against us. (Although, it's worth mentioning, that my defence of Broome's claim is quite different to Broome's and, as far as I can tell, for everything Broome himself argues, the objection above is quite appropriate.)

4. Recall:

BARRELS*: emitter Esha* throws a barrel of toxic waste into the river. The barrel pollutes victim Veena*'s farm downstream, ruining Veena*'s crop.

Veena* plainly has a directed complaint against Esha*. Just now, I said that the following are both sufficient to ground such a complaint:

- (i) Esha* *ruined* Veena*'s crop,
- (ii) Esha* increased the *risk* that Veena*'s crop would be ruined.

But what do (i) and (ii) each amount to? I take them in turn.

Elsewhere, I introduced what I called *MAKING* (see “MAKING Metaphysics”). I said that each instance of, e.g., Annie *breaking* the window, Annie *flooding* the house, Annie *melting* the chocolate, etc. is an instance of *MAKING*; just as Annie causing the window to break and Annie causing the house to flood are each instances of causing. More generally, I said that when

the window breaks,

Annie breaks the window just in case Annie *MAKES* the window break. And that when

the house floods,

Annie floods the house just in case Annie *MAKES* the house flood. More generally still, I said that (for agent *A*, verb *V* and thing *X*) when

X Vs (e.g. the window breaks),

that

A Vs X (e.g. breaks the window) just in case *A MAKES X V*.⁸

I also made the following two introductory points. Firstly, I said that *MAKING* is distinct from causing since *A* might cause *X* to *V* without *V*-ing *X* (*/MAKING X V*). For example, if Annie inspires Bob to throw a rock at the window, then while Annie caused the window to break (by causing Bob to break it), Annie herself didn't break the window (didn't *MAKE* it break). (In fact, I also argued for something stronger: that *MAKING* doesn't even reduce to causing. But this claim isn't relevant here.) Relatedly, I said that *MAKING* is distinct from counterfactual dependence: if, for example, Bob threw the rock through the window then he broke it, regardless of whether Annie would have broken it instead if he'd missed.

Secondly, I said that while we might be unfamiliar with the general concept *MAKING*, we are very familiar with its specific instances. Indeed we are able to draw subtle distinctions as to whether *A Vs X*.

For instance, we instinctively know that Annie doesn't flood Ben's house here:

⁸ To be more precise, I should distinguish between the transitive uses of a verb *V*, which require an object—e.g. *A* (the subject) broke the window (the object)—and those intransitive uses, that don't (e.g., the window (the subject) broke). If we mark these transitive and intransitive uses with the subscripts *T* and *I* then this should instead read that when

X V_I's,

that

A V_T's X (e.g. breaks_T the window) just in case *A MAKES X V_I* (e.g. *MAKES* the window break_I).

Since these details don't bear upon anything there, I've omitted them.

FLOOD: heavy rain has raised the river to dangerous levels. To protect her house, Annie blocks the doorway with sandbags. The river soon floods and the floodwaters, unable to enter Annie's house, instead enter Ben's, one door down.

Even though she does flood it here:

As before, but this time Annie is out of town, is unable to block her doorway and so her house floods. Later, she pumps the water out of her flooded basement. That water flows down the street and into Ben's house.

Similarly, we know that Annie doesn't kill Ben here:

JUMP: a runaway trolley with enough momentum to kill a single individual is heading towards Annie. Annie jumps off the track and the trolley instead hits and kills Ben who is tied to the track behind her.

Even though she does kill him here:

As before, but this time it isn't Annie on the tracks, but a barrier. Annie removes the barrier and the trolley instead hits and kills Ben who is tied to the tracks behind it.

Lastly, we know that the wind doesn't break the branch here:

WIND: the howling wind agitates the bear who jumps up and down on the tree branch until it snaps.

Even though it does break it here:

As before, but this time there is no bear. The wind buffets the branch until it breaks and falls to the ground.

And other examples abound. Ultimately, we are all experts at determining whether A Vs X and, in turn, whether A MAKES X V.

I then argued for a number of metaphysical and moral claims about MAKING, but I didn't introduce MAKING here to rely upon those claims—they're controversial and stronger than anything needed here. Instead, I introduced MAKING because it provides a neat way of framing two useful principles. Relatedly, I emphasised our intuitive expertise with specific instances of MAKING to assuage any worry that we might be ignorant as to whether these principles apply to any given situation.

With all that in mind, here is the first principle:

if A Vs B (/B's X) and if that is bad for B, then, other things equal, B has a directed complaint against A.

For example if Annie floods Ben's house then, other things equal, Ben has a directed complaint against Annie; or if Annie breaks Ben's window then, other things equal, Ben has a directed complaint against Annie. Lastly, if Esha (/Esha*) ruins Veena's (/Veena*'s) crop, then, other things equal, Veena (/Veena*) has a directed complaint against Esha (/Esha*). This isn't a controversial moral principle. (What things might be unequal? Perhaps Ben consented to Annie's flooding his house, perhaps Annie had a right to flood Ben's house, etc. Since none of the cases discussed here will engage that clause, I set it aside.)

The second principle is the causal analogue of the first:

if A causes B to V (/B's X to V) and if that is bad for B, then, other things equal, B has a directed complaint against A.

Unlike the first principle, this one is NOT true and we've already seen counterexamples to it. Recall that in FLOOD, that while Annie doesn't flood Ben's house there, she does nonetheless cause his house to flood (by blocking her doorway which causes the water to instead flood his house). Even so, Ben clearly doesn't have a directed complaint against Annie. And we saw another counterexample too: when Annie inspires Bob to throw the rock at Ben's window and, in doing so, causes Ben's window to break (by causing Bob to break it), Ben doesn't have a directed complaint against Annie—his complaint is instead against Bob. That this second principle is false is uncontroversial.

5. I now turn to

(ii) Esha* increased the risk that Veena*'s crop would be ruined.

What does that amount to? In particular, what does *increasing the risk* of something amount to?

Here is a paradigm case of increasing the risk of something:

ROULETTE: Gunslinger is determined to play Russian roulette with Vasha. Gunslinger has four bullets and a revolver with eight chambers. Before Gunslinger proceeds, Elena swaps Gunslinger's revolver for one with only six chambers. (Why does Elena do this? Let's suppose that the eight-chambered revolver is an heirloom that she wants to

reclaim.) Gunslinger loads the four bullets into the six-chambered revolver, vigorously spins the cylinder and, when it stops spinning, he aims at Vasha and pulls the trigger: Vasha is shot dead.⁹

In switching the revolvers,

Elena increased the risk that Vasha would be killed.

Importantly, this risk isn't merely so-called *epistemic* risk. It isn't simply a result of our limited evidence about the situation; in particular, it's not a result of our ignorance about whether Vasha would have been killed had Elena not switched the revolvers.

This is because there is nothing *to* know that we are ignorant of since there is no fact of the matter of whether Vasha would have been killed had Elena not swapped the revolvers. Suppose that Elena hadn't, in fact, swapped the revolvers: what would have happened? Well, Gunslinger would instead have loaded those four bullets into his eight-chambered revolver and he would have vigorously spun that gun's cylinder, before aiming at Vasha and pulling the trigger—that much we know. But whether Vasha would have been killed depends on whether there would have been a bullet in front of the firing pin; and that depends on *exactly* how Gunslinger would have spun that revolver:

if he had spun it with 8.54N of force, then (we can suppose) a bullet would have been in front of the firing pin,

if he had spun it with 8.55N of force, then a bullet wouldn't have been in front of the firing pin,

if he had spun it with 8.52N of force, then a bullet wouldn't have been in front of the firing pin,

and so on.

But there just isn't any fact of the matter of exactly how Gunslinger would have spun the eight-chambered revolver if Elena hadn't swapped the revolvers. In the language of possible worlds (and following the standard approach to evaluating conditionals such as these), the possible world in which Elena doesn't swap the revolvers and Gunslinger spins his eight-chambered revolver with 8.54N of force

⁹ Much of this section is heavily indebted to Hare (2011) and this example, to which I'll return throughout, is akin to his Wheel of Fortune case.

is no closer to the actual world than the possible world in which Elena doesn't swap the revolvers and Gunslinger spins with 8.55N of force.¹⁰

And so there is no fact of the matter as to whether Vasha would have been killed had Elena not swapped the revolvers. Instead it is *indeterminate* whether Vasha would have been killed had Elena not swapped the revolvers. *A fortiori*, there is nothing (relevant) to know that we don't know when we say that

Elena increased the risk that Vasha would be killed.

Ultimately, this isn't epistemic risk. Instead, it's *objective* risk. And by switching the revolvers, Elena increased that risk that Vasha would be killed from 4/8 to 4/6.¹¹

That it isn't epistemic risk is important because the mere fact that

A increased the (epistemic) risk that B would be killed,

(in that A did something that made it more likely, given someone's evidence, that B would be killed) isn't sufficient to give B a directed complaint against A (even when there are no countervailing moral facts). To see this, suppose that A places—what she sincerely believes to be—a deadly spell on B. In doing so, A made it more likely, given her evidence, that B would be killed, but clearly B has no directed complaint against A. (Not to mention, that whether B has a directed complaint against A is an objective matter, but epistemic risk is agent relative: some passerby C might have believed A's spell to be a protection spell and, therefore, given C's evidence, A made it less likely that B would be killed.)

Running through another example of increasing the (objective) risk will help, since it better parallels what's to come. Recall:

¹⁰ After all, what could make it that the world in which Elena swaps the revolvers and Gunslinger spins with, say, 8.54N of force is closer? To think there is a fact of the matter here is to endorse a view akin to epistemicism about vagueness (that is, the view that insists that there is a precise number of grains that is sufficient for a heap, albeit a number we are in-principle barred from knowing): there is a precise force with which Gunslinger would have spun the cylinder, albeit a force that we are in-principle barred from knowing. This isn't the place to engage with that view.

"But wouldn't Gunslinger have spun the eight-chambered with just the same force that he did in fact spin the six-chambered revolver?" Not likely. For starters, the revolvers would have different weights and that would affect Gunslinger's blood flow which, in turn, would affect the force of the spin.

¹¹ Some people are suspicious of objective risk (/chance); this paper is not the place to defend a concept so central to how we naturally understand the world. But it won't hurt to quote David Lewis on the topic: "Along with subjective credence we should believe also in objective chance. The practice and the analysis of science require both concepts. Neither can replace the other. Among the propositions that deserve our credence we find, for instance, the proposition that (as a matter of contingent fact about our world) any tritium atom that now exists has a certain chance of decaying within a year" (1980).

Just above Veena*'s farm, the river flows through ten sluice gates that dam the river. The first gate directs the flow towards Veena's farm and anything that passes through it will head in that direction, while the other nine gates direct the flow elsewhere. A highly toxic (and cohesive) algae bloom is floating downstream. Before it reaches the gates, Ferryman sails his ferry in front of the last five gates, blocking them. The river now flows through only the first five gates and, as it happens, the algae bloom passes through the first gate, where it poisons Veena*'s farm and ruins her crop.

I said that

(ii) Ferryman increased the risk that Veena*'s crop would be ruined.

This risk is also objective risk.

To see this, suppose that Ferryman hadn't, in fact, blocked those gates: what would have happened? Well, perhaps we know that he would have docked the ferry instead. And perhaps we know that he would have done it by heading roughly west-by-northwest at roughly half throttle. But whether the algae would have passed through the first gate is fixed by how the current is flowing just above the gate. And that flow, I can inform you, is *chaotic*: tiny differences rapidly multiply, while bigger differences—e.g. the location of a ferry—multiply with even greater consequence. (You might think of the movement of the algae here as akin to the movement of a given bead through a Galton Board—you know, those devices with beads, pegs and slots, used to demonstrate probability distributions. Just as tiny differences in the initial conditions of the beads will dramatically affect where each of the beads ends up, so too will tiny differences in the river's flow dramatically affect the flow of the water.)

This all means that whether the algae would have passed through the first gate had Ferryman not blocked the gates depends on exactly how Ferryman would have headed to the dock:

if he had taken heading 331.4° engaging 50.1% of the engine's potential at 12:00:00, then (we can suppose) the algae would have passed through the first gate,

if he had taken heading 331.5° engaging 50.4% of the engine's potential at 12:00:02, then the algae would have passed through the ninth gate,

if he had taken heading 331.9° engaging 49.6% of the engine's potential at 11:59:59, then the algae would have passed through the third gate,

and so on.¹²

But there just isn't any fact of the matter of exactly how Ferryman would have headed to the dock, just as there wasn't any fact of the matter of exactly how Gunslinger would have spun the eight-chambered revolver. And so there is no fact of the matter as to whether Veena*'s crop would have been ruined by the algae had Ferryman not blocked those gates.

Instead, all we can say is that the risk that the algae would have passed through the first sluice gate had Ferryman not blocked the last five gates would have been lower than it in fact was. (Just as, e.g., the risk that a given bead would enter the left-most slot on a Galton Board is higher when some other slot is blocked off, than when it isn't.)

Let's recap. We started with this simple case:

BARRELS*: emitter Esha* throws a barrel of toxic waste into the river. The barrel pollutes victim Veena*'s farm downstream, ruining Veena*'s crop.

And I said of it that

(ii) Esha* increased the risk that Veena*'s crop would be ruined.

We now know what that means: if Esha* hadn't thrown her barrel into the river then the objective risk that Veena*'s crop would have been destroyed would have been lower (at 0—or close to it, accidents do happen) than it in fact was (at 1—or close to it).

I also said that (ii) is sufficient to give Veena* a directed complaint against Esha*. That follows from the following principle:

if A increases the risk that B will V (/B's X will V) and B does V (/B's X does V) and if that is bad for B, then, other things equal, B has a directed complaint against A.¹³

But should we accept this principle? Well, I take it as a datum that when, e.g., Elena swapped the revolvers in ROULETTE, that Elena didn't merely do a terrible thing, but that she did a terrible thing *to*

¹² Perhaps these are fanciful claims—who knows. But we can at least imagine that the river just about these sluices gates is as chaotic as these conditionals claim, and that's all we need for this explanatory example.

¹³ The “and B (/B's X) does V” clause restricts the principle to only those cases in which, e.g., B is killed. Not everyone will agree that this restriction is necessary; they'll say that even when B doesn't V (e.g. B doesn't die), that B has a directed complaint against A since so-called “pure risk” imposition is sufficient for a such a complaint. That's controversial and I don't want to take a stand on it. However, if that's right (and if everything I say in §§6-7 is right), then even those people who aren't in fact affected by climate change will have a directed complaint against us emitters; and if that's right, then the matter at hand is doubly important. For discussion of pure risk see, e.g., Judith Thomson (1986) and John Oberdiek (2012).

Vasha; and I take it as another datum that Vasha has a directed complaint against Elena as a result. And what could ground that complaint if not the fact that Elena increased the risk that Vasha would be killed? And if that fact *does* ground that complaint then it must do so off the back of this principle.

But perhaps you have different data, aren't compelled by rhetorical questions and don't think we should accept this principle—either way, I don't want argue about it. Instead, over the next two sections, I'll show that the only possible grounds of the directed complaint that those affected by climate change (might) have against us emitters is that we increased the risk that they would be affected by climate change. As a result, if the fact that we increased that risk isn't sufficient to ground such a complaint (as the deniers of the principle would have it), then those affected by climate change mustn't have a directed complaint against us, after all. But that they have such a complaint was the starting assumption of this paper; not to mention, that if they don't have such a complaint, then Broome's claim (that if we offset then they will have no directed complaint against us) is trivially true. That's why I don't want to argue about this principle. Instead, let's turn to the possible grounds of that complaint.

6. Those affected by climate change will more proximately, be affected by some extreme weather event: a flood, a drought, a heatwave, a forest fire, etc.¹⁴ If we don't offset our carbon emissions (and, perhaps, even if we do?) just how are our emissions related to those affected in that way? To focus the mind, consider:

Emily's twin, Evelyn, also flies her private jet to Raglan and back for some pancakes—emitting 160 tonnes of CO₂. However, unlike Emily, Evelyn doesn't offset her emissions. 100 years later, victim Victoria's crop is ruined when a large storm floods her farm.

I have assumed that Victoria has a directed complaint against Evelyn as a result of Evelyn emitting those 160 tonnes of CO₂. What we want to know is just how Evelyn's emissions are related to the destruction of

¹⁴ “What about the immediate health effects that come from inhaling that carbon dioxide?” That isn't a result of climate change, but merely shares a common cause with climate change.

“What about, e.g., those affected by disappearing glaciers; what's the corresponding extreme weather event?” They are disappearing because they are not being replenished with additional snow, but the lack of snowfall (because of a lack of snow storms) is structurally akin to an excess of rainfall.

Victoria's crop and, in particular, what is it about that relationship that gives rise to Victoria's directed complaint against Evelyn. As I said above, I think the answer is that

Evelyn increased the risk that Victoria's crop would be ruined, and I will say more about that shortly, but let's first consider other possible answers.

(a) *If Evelyn hadn't taken her trip to Raglan, then Victoria's crop wouldn't have been ruined.* This counterfactual might be understood in different ways, depending on how fine-grained we understand its sides to be. For instance, we might understand them to be fine-grained:

(a1) If Evelyn hadn't taken her trip to Raglan *just as she did*, then Victoria's crop wouldn't have been ruined *just as it was*—by that exact flood at that exact time, etc.

Now, (a1) is true because the weather—just like the river above the sluice gates—is chaotic: it's extremely sensitive to past conditions, since tiny changes in those conditions rapidly multiply. Indeed, they multiply so rapidly that while forecasters can pretty accurately predict tomorrow's weather, their forecasts for more than ten days hence are no better than guesses. Given the weather's sensitivity to past conditions, it's a statistical certainty that if Evelyn hadn't taken the trip to Raglan *just as she did* in 2020—down to the millisecond, the milligram and the millidegree—then Victoria's crop wouldn't have been ruined *just as it was* in 2120.

Unfortunately, (a1) counts for little since the same goes for pretty much everything that happened prior to 2120. For example, it's also a statistical certainty that if Evelyn hadn't swatted that butterfly just as she did in 2020, then Victoria's crop wouldn't have been ruined just as it was. Yet whatever complaint Victoria has against Evelyn in virtue of her taking her trip, it's certainly not a complaint that she also has against Evelyn in virtue of her, e.g., swatting a butterfly or, indeed, doing anything else whatsoever.

Given how sensitive the weather is to tiny changes, perhaps a coarser counterfactual will help:

(a2) If Evelyn hadn't taken her trip to Raglan, then Victoria's crop wouldn't have been ruined by an extreme weather event in 2120.

Unfortunately, while (a1) was true but indiscriminating (in that it didn't discriminate between Evelyn's trip to Raglan and Evelyn's swatting of that butterfly), (a2) is both indiscriminating and indeterminate. It

is indeterminate for the very same reasons that it was indeterminate whether Veena*'s crop would have been ruined had Ferryman not blocked the sluice gates. Namely, given the weather's sensitivity to past conditions, whether Victoria's crop would have been destroyed in 2120 had Evelyn not taken her trip to Raglan, depends on *exactly* what Evelyn would have done instead, yet there is no fact of that matter. For instance, it might be that if Evelyn hadn't taken her trip, then she would have spent the weekend gardening instead, but the weather in the vicinity of Victoria's farm in 2120 is dependent upon exactly how she would have gardened instead:

if Evelyn had gardened in exactly *this* way (buffeting *these* CO₂ molecules with her trowel at *this* time and in *this* way etc.), then (we can suppose) there wouldn't have been an extreme weather event in the vicinity of Victoria's farm in 2120,

if Evelyn had gardened in exactly *that* way, then a tornado would have destroyed Victoria's farm in 2120,

if Evelyn had gardened in exactly *that other* way, then there would have been ideal farming weather in the vicinity of Victoria's farm in 2120,

and so on.

Yet there is no fact as to exactly how Evelyn would have gardened, and so (a2) is indeterminate.

Not only is (a2) indeterminate, but it's also indiscriminating—just as (a1) is. That's because the following is, for example, similarly indeterminate:

if Evelyn hadn't swatted that butterfly in 2020, then Victoria's crop wouldn't have been ruined by an extreme weather event in 2120,

and for the very same reasons (that is, whether Victoria's crop would have been destroyed depends on exactly what Evelyn would have done had she not swatted that butterfly, yet there is no fact of the matter).

Ultimately, determinate counterfactuals like these cannot be what grounds Victoria's directed complaint against Evelyn. (It's worth pointing out, too, that even if some counterfactual like this did single out Evelyn's trip, it would still be far from clear that it could ground Victoria's directed complaint:

Ben might not have drowned if Annie hadn't ignored his cries for help, for example, but even so, Ben has no directed complaint against Annie—Annie didn't contribute to his suffering.)

(b) *In taking her trip to Raglan, Evelyn caused Victoria's crop to be ruined (/was a cause of Victoria's crop ruining).* It's not at all clear whether Evelyn does cause Victoria's crop to be ruined. Sometimes common-sense is a good guide to whether *c* causes *e*, but I'm sceptical that common-sense has anything much to say about this sort of case. So to convince ourselves one-way-or-other, we'd have to instead see what the various competing accounts of causation said about the case, before deciding which of those accounts to believe—no small task.¹⁵

But, thankfully, it's not one we need undertake here. Since even if did convince ourselves that Evelyn causes Victoria's crop to be ruined, we've already seen that that fact is insufficient for Victoria to have a directed complaint against Evelyn. It's insufficient since, as we saw in §4, the following principle is false:

if A causes B to V (/B's X to V) and if that is bad for B, then, other things equal, B has a directed complaint against A.

(And, recall, the fact that it's false followed uncontroversially from cases like the one in which Annie causes Ben's window to break by inspiring Bob to throw a rock through it, yet Ben has no directed complaint against Annie.)

Of course, having first convinced ourselves that Evelyn does cause Victoria's crop to be ruined, we could undertake a second, even bigger task and endeavour to discover just when A's causing B to V (/B's X to V) *is* sufficient to give B a directed complaint against A. And, having done that, we could then convince ourselves that Evelyn not only caused Victoria's crop to be ruined, but that she did so in that particular way (whatever it might be) that is sufficient for Victoria to have a directed complaint against her. But I see no reason to think these steps will each work out.¹⁶ So to pursue (b) further is, I think, to clutch at straws.

¹⁵ We'd have to decide because they don't all agree. For example, David Lewis (1986) would say that Evelyn doesn't cause (/isn't a cause of) the destruction of Victoria's crop, while Johann Frick (ms) would say that she does.

¹⁶ Indeed, I think we've good reason to think they won't work out since, as I've argued elsewhere, it's MAKING, not causation, that draws the line in question—but I won't argue that here.

(c) *Evelyn ruined Victoria's crop*. By throwing her barrel into the river, Esha ruined Veena's crop and, in virtue of her doing so, Veena has a directed complaint against her. So if, by taking her trip to Raglan, Evelyn ruined Victoria's crop, then Victoria would similarly have a complaint of just against Evelyn. But Evelyn clearly doesn't ruin Victoria's crop (any more than she *floods* the crop, or *kills* the bees that live amongst the crop, etc.).

7. We've have been trying to figure out just how Evelyn's trip to Raglan is related to the destruction to Victoria's crop and, in turn, to pinpoint what it is about that relationship that grounds Victoria's directed complaint against Evelyn. The previous section's attempts were all dead-ends. In this section, I return to the fact that, by taking her trip to Raglan,

Evelyn increased the risk that Victoria's crop would be ruined.

The mechanism behind it is well understood, and breaks into two steps. Firstly, the CO₂ emitted by Evelyn warms the Earth: solar radiation enters the Earth's atmosphere and is absorbed by the land and sea; the land and sea then radiate energy this energy back up towards space, but CO₂ in the atmosphere traps a proportion of this radiation, warming the Earth; the more CO₂ in the atmosphere (i.e. the more that is emitted into the atmosphere) the greater this proportion will be and, in turn, the more the Earth will warm.

Secondly, the warmer the Earth, the greater the risk that there will be an extreme weather event in the vicinity of Victoria's farm and, in turn, that Victoria's crop will be destroyed. At this point, the rest of the mechanism differs for each of the different types of extreme weather event—forest fires, droughts, etc.—but I won't consider each type here. Instead I'll focus on floods, since they've been my working example.

The warmer Earth results in more water evaporating into vapour; that vapour will eventually condense and fall back to Earth as rain; and the more vapour there is, the greater the risk that enough of it will collect and fall in a single location, resulting in a flood. The warmer Earth also results in more land ice melting which raises the sea level; the higher the sea level, the less protective flood barriers become; and

the less protective flood barriers become, the greater the risk that any given storm will be strong enough to breach them and flood the area. And so on. Similar things go for other extreme weather events, too.¹⁷

So Evelyn's trip to Raglan warmed the Earth which, in turn, increased the risk that an extreme weather event would ruin Victoria's crop. (We might again model this with the Galton Board—albeit, this time, one with a few red beads amongst thousands of black ones. If we think of Victoria's crop being destroyed by an extreme weather event as akin to at least one of the red beads ending up in a specific slot, then Evelyn's trip to Raglan is akin to adding another red bead into the Galton Board: it increases the risk of that destruction (/of a red bead ending up in that slot).)

Since we have been unable to pinpoint any other possible grounds of Victoria's directed complaint against Evelyn—and since we have assumed that Victoria does have that complaint—we must conclude that its grounds are that

Evelyn increased the risk that Victoria's crop would be ruined.

What about when someone doesn't emit any CO₂ and only captures CO₂ from the atmosphere—how are they related to the destruction of Victoria's crop? Suppose that

Emily's other sister, Edwina doesn't emit any CO₂ that weekend in 2020. Instead, she reconfigures one of her factories and captures 160 tonnes of CO₂ from the atmosphere.

100 years later, victim Victoria's crop is ruined when a large storm floods her farm.

How is Edwina related to the destruction of Victoria's crop in 2120? Well, for symmetric reasons, by capturing that CO₂

Edwina reduced the risk that Victoria's crop would be ruined.

After all, there was less CO₂ in the atmosphere and so less radiated heat was trapped and so there was less risk that an extreme weather event would occur in the vicinity of Victoria's farm.

What remains is determining how emitting and capturing interact.

8. Recall

¹⁷ To learn more about the mechanisms driving climate change, the Intergovernmental Panel on Climate Change is a good place to start: <https://www.ipcc.ch/>.

ROULETTE: Gunslinger is determined to play Russian roulette with Vasha. Gunslinger has four bullets and a revolver with eight chambers. Before Gunslinger proceeds, Elena swaps Gunslinger's revolver for one with only six chambers. (Why does Elena do this? Let's suppose that the eight-chambered revolver is an heirloom that she wants to reclaim.) Gunslinger loads the four bullets into the six-chambered revolver, vigorously spins the cylinder and, when it stops spinning, he aims at Vasha and pulls the trigger: Vasha is shot dead.

I said that, by switching the revolvers,

Elena increased the risk that Vasha would be killed,

just as, by taking her trip to Raglan, Evelyn increased the risk that Victoria's crop would be ruined.

Now, suppose that after Elena swaps the revolvers but before Gunslinger acts, some passerby had stolen one of Gunslinger's four bullets. Since Gunslinger would now have had to play Russian roulette with only three bullets instead of four, it would be the case that

the passerby reduced the risk that Vasha would be killed,

just as, by charitably capturing those 160 tonnes of CO₂, Edwina reduced the risk that Victoria's crop would be destroyed.

But what if Elena both swaps the revolvers *and* steals a bullet:

ROULETTE (CAPTURED): Gunslinger is determined to play Russian roulette with victim Vasha*. Gunslinger has four bullets and a revolver with eight chambers. Emitter Elena* swaps his eight-chambered revolver for one with six chambers. She then takes one of Gunslinger's bullets. (Why does Elena* do this? Let's suppose that Elena* didn't want to doom Vasha* and so she takes the bullet to make amends—at least by her lights.) Gunslinger loads the three bullets into the six-chambered revolver, vigorously spins the cylinder and, when it stops spinning, he aims at Vasha* and pulls the trigger: Vasha* is shot dead.

Did Elena* increase the risk that Vasha* would be killed?

The short answer is “no, she didn't.” And the short explanation as to why she didn't is that the risk that Vasha* would be killed (3/6) is exactly as it would have been had Elena* had nothing to do with the

situation (4/8)—and no one else intervened, and there weren't any confounding factors, etc. However, it'll pay to look at things more carefully.

Firstly, it's important to distinguish this question from a nearby one. The nearby question I have in mind asks whether there exists some action that Elena* performed such that that *action* increased the risk that Vasha* would be killed—setting aside that it was Elena* who performed it and, in turn, whatever else Elena* might or mightn't have done before or after. (As a heuristic, you could suppose that the action in question was instead performed by some third-party and ask “did that third-party increase the risk that Vasha* would be killed?”) The mere fact that Elena* swapped the revolvers is sufficient for this nearby question to be answered in the affirmative—regardless of whether Elena* later stole a bullet, many bullets or, even, all the bullets. This nearby question is NOT relevant here; for starters, it's far too broad to fix whether Vasha* has a directed complaint against Elena* (e.g. consider the case where she later steals all the bullets!).

Instead, what we want to know is whether Elena* the *agent* increased the risk that Vasha* would be killed—taking into account that she both swapped the revolvers and stole the bullet.

Returning to that question, notice how in ROULETTE (CAPTURED) Elena* first swapped the revolvers and then stole the bullet. Now, that she performed those actions in that order is an artefact of the case, yet, in discussion, some have suggested that that ordering matters *vis-a-vis* whether Elena* increased the risk that Vasha* would be killed—we should wonder whether that's right. Those discussions also pointed to another artefact of the case that might be playing a bigger role than my description of the case allows for—namely, the length of the delay between Elena*'s two actions. In particular, some suggested that it makes a difference whether the two actions were performed within, say, a few minutes of each other (a short delay) or, say, on subsequent days (a long delay).

With all that in mind, here are four ways ROULETTE (CAPTURED) might have gone:

- (1) Elena* swaps the revolvers ... short delay ... Elena* steals the bullet,
- (2) Elena* swaps the revolvers ... long delay ... Elena* steals the bullet,
- (3) Elena* steals the bullet ... short delay ... Elena* swaps the revolvers,

(4) Elena* steals the bullet ... long delay ... Elena* swaps the revolvers.

Of each of them we should ask whether Elena* increased the risk that Vasha* would be killed.

With regards to (1) and (2), I think it's plain that Elena* didn't increase that risk. After all, there is no action that Elena* performed—neither swapping the revolvers nor stealing the bullet—such that, if she hadn't performed it, then the risk that Vasha* would have been killed would have been lower than it in fact was: if Elena* hadn't swapped the revolvers, then she wouldn't have later stolen the bullet, and so the risk would have been just as it was; and if Elena* hadn't later stolen the bullet, then the risk would have been higher than it in fact was— $4/6$ instead of $3/6$. Given that, I see no room for thinking that Elena* increased the risk that Vasha* would be killed in (1) nor (2).

I also think that (3) is a clear case. Suppose that Gunslinger had left his ordnance on the table and that Elena* swapped the revolvers (/stolen the bullet) just moments before she stole the bullet (/swapped the revolvers). It is surely confused to think that it would have made all the difference in the world—*vis-a-vis* whether Elena* increased the risk that Vasha* would be killed—if Elena* had swapped before stealing (as in (1)) or had stolen before swapping (as in (3)). Since it's plain that Elena* didn't increase the risk that Vasha* would be killed in (1), we are forced to conclude the same of (3).

That leaves

(4) Elena* steals the bullet ... long delay ... Elena* swaps the revolvers,

which, in discussion, was the case that most inclined people to think that Elena* increased the risk that Vasha* would be killed. It did so because this time there is some action that Elena* performed—namely, swapping the revolvers—such that if she hadn't performed it, then the risk that Vasha* would have been killed would have been lower ($3/8$ instead of $3/6$). Moreover, given the long delay between Elena* stealing the bullet and then later swapping the revolvers, it's (perhaps?) natural to think that she swapped the revolvers independently of how she first stole the bullet and, therefore, it's (perhaps?) natural to think that her swapping of the revolvers should be evaluated without consideration of how she first stole the bullet. And once we make it that far, the inevitable conclusion is that, in swapping the revolvers, Elena* increased the risk that Vasha* would be killed.

Is *any* of that right though? For starters, I'm sceptical of this "independence," and I'm sceptical that it — whatever it might be — could entail that Elena*'s swapping of the revolver should be evaluated without consideration of how she first stole the bullet. Most of all though, I see no reason to accept the claim that the long delay between Elena*'s two actions entails, or even indicates, anything interesting about the second action's relationship to the first. That I perform some action A2 a long time after I perform A1 is compatible with me performing A2 because I performed A1 (I returned the book Friday because I returned it Monday); it's compatible with the my performance of A1 being necessary for my performance of A2 (I can't return a book without first having borrowed it); it's compatible with my having made a single decision to perform both A1 and, later, A2 (I planned on Monday to borrow the book and to return it in four days); and so on. Importantly, once we return to offsetting, it's implausible to think that the delay between one's emitting and capturing (/one's capturing and emitting) is indicative of anything other than one's timeliness (/one's lack thereof).

There's also a broader reason not to worry about the ordering—at least not in the context of Emily's offsetting. This is because, in reality, Emily won't emit and capture once, but repeatedly throughout her lifetime: she might emit 160 tonnes of CO₂ on the 1st, capture 160 tonnes of CO₂ on the 5th and emit another 160 tonnes of CO₂ on the 10th, and so on. The following is a bad question: in capturing those 160 tonnes of CO₂ on the 5th, did Emily offset her emissions from the 1st or from the 10th? It's a bad question because there is surely no fact of the matter—what *could* settle it? And if there is no fact of that matter, there can be no fact of the matter as to whether Emily captured before or after she emitted. And if there is no fact to that order, then that order cannot make a difference.

Ultimately, however we cut it, Elena* didn't increase the risk that Vasha* would be killed in ROULETTE (CAPTURE).

Even once we accept that, we might still wonder *why* that's the case. That is, we might wonder why the fact that Elena* later stole one of Gunslinger's bullets makes it the case that she (Elena*) didn't increase the risk that Vasha* would be killed; particularly when contrasted with how, in BARRELS (CAPTURED), the

fact that Esha later captured the upstream barrel had no bearing on the fact that she (Esha) nonetheless destroyed Veena's crop.

The answer, I believe, is that it's because risk is *fungible*: its units are interchangeable in every way. Electronic money is also fungible and it behaves similarly. Suppose, for example, that Philanthropist first transfers \$1000 into PETA's account and, later, hacks into PETA's account and transfers \$1000 elsewhere. Only a confused fur-trader could complain of Philanthropist that she enriched PETA. That's because electronic money is fungible and so it's nonsense to say of any particular unit of money in PETA's account that it came from some donor or other. Instead, the only facts in the vicinity are quantitative ones: that the balance is higher (/lower) than it would have been were it not for a certain credit (/debit); that, e.g., the balance is \$20 higher than it would have been but for Sally's donation. Yet, given that Philanthropist performed both transfers and, as a result, the balance is the same as it would have been had Philanthropist performed neither transfer, there is no such quantitative fact. *A fortiori* there is nothing for the fur-trader to complain about.

The very same goes for risk. Risk is fungible and it's similarly nonsensical to say in, e.g., ROULETTE that any given unit of risk that Vasha would be killed is "Elena's risk" (what could that even mean?). Instead, all we can say in that particular case is that the risk that Vasha would be killed is higher than it would have been had Elena not acted as she did. However, in ROULETTE (CAPTURED) we can't even say that: given that Elena* both swapped the revolvers and stole the bullet and, as a result, the risk that Vasha* would be killed was the same as it would have been had Elena* performed neither of those actions, there is no such quantitative fact.

At least, that's my best effort at an explanation. Ultimately, though, what's important isn't what explains why Elena* didn't increase the risk that Vasha* would be killed, but only the fact that she didn't. And that fact is plain enough: Elena* didn't increase the risk that Vasha* would be killed in ROULETTE (CAPTURED).

9. It's time to return to Emily:

EMISSIONS (CAPTURED): Emily takes her private jet to Raglan and back for some pancakes—emitting 160 tonnes of CO₂. Upon her return, she captures 160 tonnes of CO₂ from the atmosphere. 100 years later, in 2120, Victoria's crop is ruined when a large storm floods her farm.

We know that if Victoria is to have a directed complaint against Emily in virtue of her taking her trip to Raglan, the grounds of that complaint must be that

Emily increased the risk that Victoria's crop would be ruined.

However, it's now clear that that isn't the case. Just as Elena* didn't increase the risk that Vasha* would be killed in ROULETTE (CAPTURED), nor did Emily increase the risk that Victoria's crop would be ruined.

That means that Victoria has no directed complaint against Emily in virtue of her emitting those 160 tonnes of CO₂. And the reason why she doesn't is because Emily offset those emissions; Broome was right, after all.

The upshot for Emily is that, so long as she continues to offset her emissions, then she can jet around the world with abandon and those affected by climate change will have no directed complaint against her in virtue of her doing so. (Although, to return to where this paper began, they might well complain of Emily that she could have better spent her money—but that's a different matter.)

10. What about the rest of us? I remarked in §2 that while Emily herself both emits and captures the CO₂ in EMISSIONS (CAPTURED), that's not the case for the majority of the emissions that make up our own carbon footprints; it also very likely won't be the case for any offsetting we might do. Let's say that the portion of our carbon footprints that we emit ourselves (e.g. from driving the car, from cooking with gas, etc.) are our *direct* emissions; and let's say that the portion that we don't emit ourselves (e.g. from the electricity we use, from the flights we take, etc.) are our *indirect* emissions. Similarly, let any offsetting that we do ourselves be *direct* offsetting, and any offsetting that others do on our behalf be *indirect* offsetting. Lastly, let's say that E *successfully* offsets her emissions from a given CO₂-emitting activity

just in case that offsetting means that those affected by climate change have no directed complaint against her in virtue of that activity. (On the other hand, Esha's offsetting in BARRELS (CAPTURED), for example, wasn't successful, since Veena still had a directed complaint against her.)

Now, it's unclear to me just what moral responsibility I bear for my indirect emissions vs my direct emissions: perhaps I am less responsible for them, perhaps I am equally responsible—I don't know.¹⁸ Similarly, it's unclear to me just what moral credit I bear for my indirect offsetting vs my direct offsetting. But what is clear enough is that the two answers will mirror each other: if I bear less credit for indirectly offsetting n tonnes of CO₂, then I see no reason to think that I won't bear less responsibility, by the same degree, for indirectly emitting n tonnes of CO₂. That is, I see no reason to think that the two sides won't balance each other out. This is good news because around 80% of carbon footprint is indirectly emitted and, therefore, we can successfully offset that proportion of our carbon footprint indirectly.

What about the remaining 20% of our carbon footprint? Without being clear about the two issues, above, it's not possible to say what our own prospects are for successfully offsetting those emissions. However, since that still leaves 80% of our carbon footprints to work on, I set this issue aside.

I also remarked in §2, that while one way of offsetting the emissions of a given activity is by capturing the equivalent amount of CO₂ from the atmosphere, it's not the only way. Indeed, the dominant form of offsetting today is offsetting by prevention, in which emitter E offsets the emissions for a given activity by paying offsettee O not to emit an equivalent amount of CO₂. For example, E might offset the tonne of CO₂ emitted from her flight to Paris by paying (/funding) O not to power his home by burning coal, but instead with solar power; in doing so, E prevents O from emitting a tonne of CO₂ that he would otherwise have emitted and, in turn, E offsets the CO₂ from her flight.

¹⁸ There are considerations in both directions. Since the direct emitter of my indirect emissions is, surely, partly responsible for those emissions, doesn't their partial responsibility lessen my own? Then again, can it really make a moral difference whether I boil my kettle with gas (and so directly emit) or with electricity (and so indirectly emit)?

Not only is offsetting by prevention the dominant form of offsetting, it's also the cheapest—by some margin: at present prices, it would cost you around \$20 to offset a tonne of CO₂ by prevention, but around \$900 to do so by capture.¹⁹

So why has this paper focused on offsetting by capture? Well, offsetting by prevention is cheap because supply currently outstrips demand. However, while that demand is increasing, there is a hard upper-limit on supply (there are, e.g., only so many people that are currently burning coal that can be paid to instead use solar). And so offsetting by prevention will only become more expensive, before disappearing altogether.²⁰ On the other hand, offsetting by capture is expensive because the technology is new and, in turn, because demand is low. Moreover, there is no limit to possible supply (at least, not while atmospheric CO₂ levels remain a concern). As a result, prices are quickly dropping and are estimated to be closer to \$100 a tonne in just a few years.²¹ This all means that offsetting by capture is sustainable in the longterm, while offsetting by prevention is not—and that is why I focused on it. (And, as we'll see shortly, it's rarely prohibitively expensive, anyway.)

Nonetheless, we might wonder whether E can successfully offset her emissions by offsetting by prevention. If we focus on indirect emissions and offsetting—as I've suggested we should—then the relevant test case would be the following:

Gunslinger is determined to play Russian roulette with victim Vasha*. Gunslinger has *three* bullets and a revolver with eight chambers. Indirect emitter Indiana pays Elena* to swap Gunslinger's eight-chambered revolver for one with six chambers. Indiana then pays offsettee Oleg not to supply Gunslinger with a fourth bullet (which he would have done, but for Indiana's payment to him). Gunslinger loads the three bullets into the six-chambered revolver... Vasha* is shot dead.

If Indiana had neither paid Elena* to swap the revolvers nor paid Oleg to withhold the additional bullet, then the risk that Vasha* would be killed would have been 1/2 (Gunslinger's original three bullets plus

¹⁹ E.g. at this company: <https://www.climeworks.com/>.

²⁰ There's also a potential free-rider problem for those offsetting by prevention today, when the prices are low. See Kai Spiekermann (2014).

²¹ See, e.g., [https://www.cell.com/joule/fulltext/S2542-4351\(18\)30225-3](https://www.cell.com/joule/fulltext/S2542-4351(18)30225-3) and <https://www.sciencemag.org/news/2018/06/cost-plunges-capturing-carbon-dioxide-air>.

Oleg's single bullet in the eight-chambered revolver), and that is the same risk as it, in fact, was (three bullets in the six-chambered revolver). And so, by paying Oleg to withhold the bullet, Indiana has offset her payment to Elena* to swap the revolvers. But has she done so successfully?

I'm inclined to think so. As far as I can tell the only difference between this case and ROULETTE (CAPTURED) is that in this case Indiana's "emissions" and "offsetting" are both indirect. And, as I said above, since they are both indirect, I see no reason to think that this difference won't balance out. And, if it does and if that truly is the only difference, then we should conclude of this case exactly what we concluded of ROULETTE (CAPTURED); and, in turn, we should conclude that we can successfully offset our emissions by prevention. That said, I have no great confidence in that conclusion, but since offsetting by prevention isn't sustainable anyway, I don't think we should worry about it all that much.

11. What are the practical implications of all this? Here's one. Suppose that climate-conscious Helena is planning her summer holiday; she wants to experience those Raglan pancakes that she's heard so much about, but she's concerned about contributing to climate change. By this paper's reckoning, so long as Helena offsets the emissions from her trip, then she can travel guilt-free.

Here's another. Suppose that climate-conscious Maia is choosing between two (realistic) options for a journey from London to Madrid:

1. fly with a budget airline for £50, emitting 150kg CO₂, and (successfully) offset-by-capture those emissions for a further £135,
2. take a train, emitting 50kg of CO₂ for £200.

As far as those affected by climate change are concerned, Maia should choose the first option. This is good news for Maia since the first option is both cheaper and significantly quicker. (Note, that not only is it cheaper, but it's still relatively affordable. So while offsetting by capture is expensive, it won't, for many of us, be prohibitively so. And, as I say, it's quickly becoming cheaper still.)

Ultimately, so long as we (successfully) offset the emissions from our activities, then we can continue doing them guilt-free, since, having offset those emissions, those affected by climate change will have no directed complaint against us in virtue of our having done them.

References

- Broome, J. (2012). *Climate Matters: Ethics in a Warming World*. W. W. Norton.
- Byrne, T. (forthcoming) MAKING Metaphysics. *Philosophers' Imprint*.
- Cripps, E. (2016). On Climate Matters: Offsetting, Population, and Justice. *Midwest Studies In Philosophy*, 40(1), 114–128.
- Frick, J. (Submitted). Chancy Causation and the Problem of Aggregate Events.
- Goodin, R. E. (1994). Selling Environmental Indulgences. *Kyklos*, 47(4), 573–596.
- Hare, C. (2011). Obligation and Regret When There is No Fact of the Matter About What Would Have Happened if You Had not Done What You Did. *Noûs*, 45(1), 190–206.
- Hare, C. (2013). Symposium on the work of John Broome at Boston University.
- Kagan, S. (2011). Do I Make a Difference? *Philosophy & Public Affairs*, 39(2), 105–141.
- Lewis, D. (1981). A Subjectivist's Guide to Objective Chance. In W. L. Harper, R. Stalnaker, & G. Pearce (Eds.), *IFS: Conditionals, Belief, Decision, Chance and Time* (pp. 267–297). Springer Netherlands.
- Lewis, D. (1986). Postscripts to 'Causation'. In D. Lewis (Ed.), *Philosophical Papers Vol. II*. Oxford University Press.
- Monbiot, G. (2006, October 18). Paying for our sins. *The Guardian*. <https://www.theguardian.com/environment/2006/oct/18/green.guardiansocietysupplement>
- Oberdiek, J. (2012). THE MORAL SIGNIFICANCE OF RISKING. *Legal Theory*, 18(3), 339–356. Cambridge Core.
- Scanlon, T. (1998). *What We Owe to Each Other*. Belknap Press of Harvard University Press.
- Singer, P. (1972). Famine, affluence, and morality. *Philosophy & Public Affairs*, 229–243.
- Spiekermann, K. (2014). Buying Low, Flying High: Carbon Offsets and Partial Compliance. *Political Studies*, 62(4), 913–929.
- Thomson, J. J. (1986). Imposing Risks. In *Rights, restitution, and risk: Essays, in moral theory*. Harvard University Press.