CONTENTS

1

Killing

ON THE ETHICAL SIGNIFICANCE OF CAUSATIVES

MUCH HAS BEEN written about the moral significance of the contrast between killing and letting die. The canonical contrast here is between cases where someone's death is the result of some positive action and cases where someone's death is the result of a failure to do something.[1] But there has been rather less written about the distinction between cases where someone's death is counterfactually dependent on some positive action by an agent and cases where an agent *kills* someone, and there has in particular been little extended discussion of the ethical significance of this latter distinction. Section 1 introduces the distinction. Section 2 explores its conceptual underpinnings. Section 3 explores the ethical significance of the distinction. Some concluding remarks follow. An appendix discusses some relevant recent literature in philosophy and linguistics. A second appendix explores the concept of ethical significance.

1. Philippa Foot (1967) did much to bring this kind of contrast to the attention of philosophers. There is a vast literature on the topic of killing and letting die, as well as on the more general distinction between doing and allowing. Particularly influential works in this area include (in addition to Foot 1967) Bennett 1995, Foot 1984, Quinn 1989, and Thomson 1976. We note that Foot distinguished two uses of "allowing." The use she was "concerned with" applies where an agent could "intervene" to stop a sequence of events "somehow already in train" but does not intervene, as when one "could stop a leaking tap but allows the water to go on flowing." The other "is roughly equivalent to *enabling*," as when "someone may remove a plug and *allow* water to flow" (1967, 3). Neither use maps onto the key distinction that animates this paper.

Introducing the Distinction

We are all going to die. But some of us will die sooner than others. And the actions of other agents make a difference to when we die. Let us say that x, in doing something, is a *factor* in y's death at t just in case y would have died later than t had they not done that.[2] (We are using "factor" as a term of art here—it has obviously been chosen because it bears *some* resemblance to ordinary usage, but it is not important to us that it matches ordinary usage.) And to ensure that our current topic does not get too muddled with doing/allowing discussions, let us not be too expansive about what counts as doing something. In particular we shall generally not count omissions, failures to do something, as a kind of doing something.

Someone can kill someone else and not be a factor (in the defined sense) in that person's death. Suppose Jones pushes a button that releases a trapdoor that sends Smith plummeting into a canyon to his death. If Jones had not pushed Smith, Smith would have been killed by a sniper right after the time of the actual trapdoor release. It is a deep canyon: Smith dies later than he would have if Jones had not released the trapdoor. Jones doesn't know that—they just want Smith dead. Obviously, Jones killed Smith, but Jones wasn't, in the sense defined, a factor in Smith's death.[3]

Our focus here will be on failures of entailment in the *other* direction. It is clear that one can be a factor in someone's death without killing them. Suppose Jones takes Smith to the beach and a lunatic sniper, whose presence Jones did not foresee, kills various people on the beach, including Smith. Jones did not kill Smith (the sniper did), and this is true even if Jones coerced Smith into going to the beach. But it is easy to fill in the details so that Jones was a factor in the defined sense—in particular, it is easy to fill in the details so that if Jones hadn't

2. When the choice is not binary, there is a slightly different formulation that might be considered: x, in doing something, is a factor in y's death just so long as there is some alternative action available such that y would have died later had x taken that alternative course of action. This comes apart from the conception in the main text: consider a case where x performs action A, would have performed B if they hadn't done A, but where it would have taken the performance of a third available action C to extend y's life. We slightly prefer the more convoluted formulation, but as it doesn't really matter much to what we have to say, we shall work with the simpler formulation in the text. Note that there is also a natural use of "factor" where it is predicated of an action—an action by x is a factor in y's death just in case had that action not been performed, y would have lived longer. Sometimes this use will be more convenient. As the connection between the two uses is clear enough, we shall freely go back and forth between them.

3. That is one reason why the intended meaning of our term of art does not quite match ordinary uses of "being a factor in someone's death."

taken Smith to the beach, Smith would have died later. For pretty much anything we do, it is easy to imagine ways in which, by doing that thing, we are a factor in someone's death. (Indeed, given butterfly effects, it is plausible that we can be pretty sure, and perhaps even know, for any action we perform, that there is someone who would have lived longer had we not performed that action.[4])

When x is a factor in y's death, y would have lived longer had x not acted the way they did. How much longer? This varies from case to case, but it bears emphasis that the time differential provides no clean guide to whether x killed y or not. If x shoots person y in the head, x kills y. But the narrative can be filled in so that y would have lived only a tiny bit longer had they not been shot (as when one shoots a person who is close to death anyway); and of course it can also be filled in so that y would have lived a great deal longer had they not been shot (as when a person is fatally shot in their youth and would have otherwise lived a long life). The relevant time differential is not going to be the key to the contrast between cases where one is a mere factor in someone's death and cases where one kills someone. What *does* explain that contrast?[5]

Investigating the Distinction

"To kill" is an example of what linguists and grammarians call a causative verb, a kind of verb that implies causality. (Where V is a causative verb, "x V'd y but x had no effect on y," makes no sense.[6,7]) We wish to draw attention to four

4. This is especially plausible if we assume the principle of counterfactual excluded middle (CEM), according to which, for any P and Q, either Q would be the case if P were the case, or Not-Q would be the case if P were the case. For each future person we can ask, with reference to a particular nose scratching, "Would they have lived the very same length of time if Jones had not scratched his nose?" Given butterfly effects and CEM, we can expect many "No" answers and it is obvious that plenty of the "No" answers will concern cases where the agent would have lived less long. For an extended defense of CEM, see Dorr and Hawthorne (ms.).

5. One could also in principle conduct a somewhat parallel investigation into the contrast between saving someone and being an anti-factor in their death (where x is an anti-factor in y's death just in case y would have lived less long had x not acted the way they did). But there may be disanalogies. For example, while one can kill without being a factor, can one save someone's life without being an anti-factor?

6. What about a case in which x kills y, but y would have died at the exact same time, in the exact same manner, had x done otherwise? Even here it is natural to say x's action had an effect on y, even though it may be that if x had acted otherwise, someone else's action would have had the same effect.

7. There are some interesting questions about the legal use of "kill." In many jurisdictions, a person can be found guilty of murder by reason of being an accomplice in a certain felony during which a killing occurred. Suppose you are correctly found guilty of murdering someone. If

plausible ideas concerning how (at least a large class of[8]) causative verbs work, all of which straightforwardly apply to "kill." The first two draw upon prominent themes in Hart and Honoré's classic *Causation in the Law*, an important resource for any study of the interplay of ethics and causation. The focus of those authors, of course, is on applications of causal ideology in a legal setting, but they emphasize that the courts' uses of causal language "have their roots in certain features of a variety of concepts which permeate the daily non-legal discourse of ordinary men" (1985, 1). And while their focus is not on causative verbs per se, their discussion can easily be adapted into some promising heuristics concerning how such verbs are put to work.[9]

Mediating Agents

The first idea is this. If x, by doing F, is a factor in y's death, but y's death also crucially depends on some later intervention by a voluntary agent that exploits x's doing F, then x does not kill y. This applies a general idea about causation in the law that Hart and Honoré put as follows:

> [F]ree, deliberate and informed act or omission of a human being, intended to exploit the situation created by defendant, negatives causal connection. (1985, 136)[10]

(For the purposes of this essay, we can rewrite "causal connection" as "causative connection.")

And here is one of their vivid examples:

> *A* throws a lighted cigarette into the bracken which catches fire. Just as the flames are about to flicker out, *B*, who is not acting in concert with *A*, deliberately pours petrol on them. The fire spreads and burns down the forest. *A*'s action, whether or not he intended the forest fire, was not the cause of the fire: *B*'s was. (1985, 74)

---

you murder, does it follow you kill? And if so does that mean that, after all, you can kill without having an effect on the victim? We doubt that legal practitioners would be comfortable asserting the latter claim. But they may be prone to fudge the issue about exactly where this chain of inference breaks down.

8. Our reason for this qualification will become clear (see the discussion of "hasten" and "accelerate" below).

9. The connection between causatives and the notion of causation operative in the law is a prominent theme in Thomas Byrne (2022).

10. We are aware that the use of "negative" as a verb will sound a bit jarring to many readers. Read as you would "negate."

The phrase "intends to exploit" in their formulation is crucial. As Hart and Honoré note, "The act of a consumer who uses a product without knowing that it is defective will not bar an action against a negligent manufacturer" (1985, 149). Similarly, if x plants a car bomb and y—ignorant of the danger—voluntarily turns the key, resulting in the death of y's family, that does not stop us saying that x killed y's family.

We find *something* like this idea in some more recent work on causatives. For example, Richmond Thomason (2014) distinguishes "efficient agents" from "automatic agents" (where paradigms of the latter include the wind or a fire) and offers the generalization that "an efficient agent is precluded [from being in the relevant causal relation to the outcome of a process] by another efficient agent between it and the body [of a process]" (2014, 70). (Note, though, that for reasons just given, Thomason's generalization is too incautious.)[11]

Hart and Honoré explain what they have in mind by a "free, deliberate and informed act": Intermediate agents that are acting out of "self-preservation" to prevent significant harm coming to them, or who are acting under some legal or social obligation, are not intended to count as engaging in such an act (1985, 144). In this way historians often give some number in the millions when asked how many people Stalin killed.[12] As Thomason would put it, when we use "kill" in this kind of way, various intermediate animate agents are "automatized."

Abnormal Interventions

Here is a rough first pass at the second idea: If x is a factor in y's death by doing F, but the death crucially depends on an abnormal intervention between x's F-ing and y's death, then x does not kill y. Hart and Honoré offer a version of this idea, generalized to causal responsibility in general:

> The basic principle is that normal physical events, even subsequent to the wrongful act, do not relieve a wrongdoer of responsibility but that an abnormal conjunction of events (in this case the wrongful act and the third factor) negatives causal connection, provided that the conjunction is not designed by human agency. (1985, 162–63)

11. There is a significant body of work on intervening agency in the ethics of harming. See, for example, Frowe 2022 and Tadros 2007.

12. See, for example, Stanford Report, "Stalin Killed Millions: A Stanford Historian Answers the Question, Was it Genocide?" September 23, 2010, https://news.stanford.edu/2010/09/23/naimark-stalin-genocide-092310/.

(Again, for the purposes of this essay, we can rewrite "causal connection" as "causative connection.") They emphasize that it is crucial that the abnormality is posterior to the action. And they are right to do so. If one of us pushed the other into a zoo cage with a tiger in, dooming them, it is quite natural to say that the one killed the other, even if, owing to a recent ban on tiger captivity, it is quite abnormal for zoos to still have tigers.[13]

Here are a few pairs inspired by their discussion that illustrate the contrast between cases where abnormality is already in place and those where it is posterior to the action:

Pair 1.

Scenario 1: Jones pushes Smith over. Smith falls onto a spike in the ground at the spot where he lands, with fatal effect.

Scenario 2: Jones pushes Smith over. Right after the push, a bolt of lightning hits a tree that falls on top of Smith while Smith is lying on the ground. Smith is crushed and dies.

Intuitively, Jones kills Smith in scenario 1 but not scenario 2.

Pair 2.

Scenario 1: Jones drops a brick out of a second-story window. Smith, trespassing, happens to be standing beneath the window. The brick lands on Smith, with fatal effect.

13. What if the tiger was typically sleepy and not hungry? The fact is that our ordinary sensibilities tell us that there is nothing strange about a continuation in a world where one is pushed into a tiger cage to get attacked by the tiger. One relevant point here is that many uses of normality have an existential flavor: having four friends at a dinner party may be perfectly normal, even if having three and having five are also perfectly normal (see Carter and Hawthorne [ms.] for further discussion). It is best to tap into this existential use for these purposes—don't ask which continuation normally happens, but instead which continuations would be normal. Also, since we are so used to thinking of tiger attacks as perfectly normal when someone enters a tiger cage, our ordinary habits of thought may cause noise were we to try to switch to an unusual tiger case, as when the tiger is sedated and only wakes thanks to an unusual pistol shot that occurs after the cage is entered. Better instead to switch to cases that are not contaminated by, as it were, our normal normality thoughts. (Thanks to Thomas Byrne for helpful discussion here.) There are certainly residual questions about the relevant sense of abnormality. But it certainly doesn't just amount to "low probability." Imagine a dystopian future where people are paid to attach an explosive device to their pet. Post attachment, a bingo machine selects a ball associated with one device and this automatically triggers one device attached to an unlucky pet. If Jones's dog dies, that is due to a low probability intervention, but the claim that Jones killed their dog sounds pretty good. The discussions in Carter and Hawthorne (ms) and Smith (2010) may be helpful here.

> Scenario 2: Jones drops a brick from a second-story window, and as it is falling, a bolt of lightning hits it, redirecting it into a neighboring garden, where it fatally strikes Smith.

Scenario 1 is a normal continuation scenario. Two is not. We are much more comfortable saying that Jones killed Smith in scenario 1 than in scenario 2.

We glossed our presentation of the second idea as a "rough first pass." Hart and Honoré's caveat "provided that the conjunction is not designed by human agency" points to some need to qualify. Suppose Jones knows that a tree is about to fall because of some abnormal weather conditions and knows where it is going to land. Jones pushes Smith so that Smith lands on the relevant spot just prior to the tree's breaking and falling on that spot. Here the conjunction of falling and then being hit by a tree is "designed by" Jones, so the abnormality of the breaking does not block causativity.[14]

The idea that abnormal interventions block causality has been taken up in some recent literature, including the substantial literature on causal modeling.[15] For example, as a prelude to a formal model of causal strength, Icard, Kominsky, and Knobe offer both the principle of "Abnormal Inflation" that says that if C and A are both factors in an outcome, "people will be more inclined to say" C is the cause when "they regard C as abnormal than when they regard C as normal" (2017, 81) and the principle of "Supersession" that says

14. The example of x pushing y to the ground just as a tree is about to fall on them is discussed at length in Hart and Honoré, chapter 3. There they make the point that foreknowledge of the event makes a difference. One might reasonably wonder whether "abnormal invention between the action and the death" could be replaced by "intervention between the action and the death that could not reasonably be foreseen." Unless the notion of intervention is refined so as to do a lot of heavy lifting, this alternative proposal seems to get the wrong result in cases where the intervention is not reasonably foreseen but is a normal continuation of a situation already in place before the action. Suppose, for example, that one has no reason to believe that someone is in one's house, throws a brick out of the window, and then, right after, someone steps out through the back door and gets hit by a brick, dying as a consequence. The stepping out was not foreseeable, was after the throwing, and resulted in a death. But it seems very natural in this case to suppose that a killing took place. Recall also from fn. 15 that the existential sense of normality need not imply anything like reasonably high probability and for that reason need not imply foreseeability. Suppose a screen opens for a few seconds every minute. The screen is between Smith and Jones. Smith shoots an arrow toward the screen hoping that it will open and that Jones will be hit. The screen does open and Jones dies from an arrow wound. The death required a change that could not reasonably be foreseen—the opening of the screen—but the case is clearly one of killing.

15. See, for example, discussions of normality and deviance in Gallow 2022, Hall 2007 and Halpern 2008.

people are less inclined to treat C as the cause when A is abnormal than when A is normal. (We note in passing, though, that temporal order does not figure in these principles: Arguably, principles that take account of temporal order will be even more predictive of folk practice.)

Elsewhere, Knobe remarks that folk judgments of causation are driven by which counterfactuals they focus on rather than ignore. Moreover, he goes on to suggest (citing Kahneman and Tversky 1982; Roese 1997) that people are inclined to "consider counterfactuals in which events of *unusual* types are replaced by events of *usual* types" (2009, 241).

Let us juxtapose this abnormality-driven idea with David Lewis's suggestion about what is distinctive about killing, namely,

> insensitivity to circumstances . . . if you shoot at your victim point-blank, only some very remarkable difference in circumstances would prevent his death. The same is true if you set . . . a delayed action-bomb, working inexorably towards its lethal outcome. The case of a bomb with a randomizer is comparatively insensitive: the bomb might well have chanced to go off, but it isn't the fine details of the circumstances that would make the difference. (1986a, 186)

The idea (adapted to the ideology of factors) is that when you are a factor but don't kill, the death depends much more on the fine details of circumstances than when you kill.[16] We don't think this does as well as the normality-theoretic idea, at least when it comes to understanding folk practice. Suppose Jones takes Smith to Liechtenstein. A powerful nuclear missile has already been directed to Liechtenstein, big enough that there is no chance of anyone in Liechtenstein surviving when it hits. Smith's subsequent annihilation does not seem to depend very much on the fine details. Compare this with a case where Jones shoots at Smith from a distance while Smith is running. Jones is not highly skilled. If Smith had slowed down just a tiny bit or speeded up just a bit, or if Jones had aimed just a tiny bit to the left or right, Smith would not have died. Comparatively speaking there is more sensitivity to fine details. (And we can ramp up the sensitivity even more in the second case by having the bullet enter a small hole in a slightly defective bulletproof vest.) But in the latter case, Jones kills Smith, whereas in the former Jones does not. Contrast similarly a case where Jones takes Smith into a car that has a car bomb with no randomizer with a case where Jones pushes Smith who lands on a spike. The

16. Lewis is clearly aware that killings may be sensitive to certain fine details, as when the bomb's going off depends on the outcome of rolling dice. It is the comparative sensitivity that is supposed to matter.

death of Smith may depend on fine details of spike location and the trajectory of the fall, while death subsequent to the turning of the key may be comparatively less sensitive to details. At least if we are trying to understand what drives ordinary judgments, the normality-theoretic idea seems superior to that of Lewis.

Bad Actors

A third theme—one that is prominent in work (some coauthored) by Joshua Knobe—is that, ceteris paribus, people are more inclined to consider bad actors the cause rather than good actors even if both are factors.[17] Here is an example much in the spirit of this discussion, adapted to killing:[18]

> A patient needs 50 mls of drug A each day. However, if that patient receives 100 mls over the space of an hour, they will die. One morning, within the space of an hour, the nurse gives that patient a 50 ml dose, and someone who shouldn't be handling hospital drugs also gives the patient a 50 ml dose. (To rid ourselves of distractions due to temporal order, we can even make the injections simultaneous—perhaps the nurse cannot see the other agent because they are separated by a screen or can see the other agent, who is dressed as a nurse and who tells the real nurse they are injecting something else.)

Clearly, we are more inclined to say that the impostor killed rather than the nurse.[19] Of course in this case there is obviously a clear normality contrast as

17. This a central theme of Knobe (2009), where he draws on the principle that "people are inclined to consider counterfactuals in which bad events are replaced by good events" (240), which then has knock-on effects for judgments of causality. See also Hitchcock and Knobe 2009.

18. Knobe and Hitchcock (2009) present cases very similar to the one that follows.

19. This phenomenon is arguably less prominent when the verb has positive valence. Suppose that due to a rare condition a patient needs 100 mls of drug A to save them (when 50 mls is normally enough). A good nurse injects 50 mls. A bad nurse, falsely believing drug A to be lethal also simultaneously injects 50 mls of drug A. The two doses combine to produce the needed 100 mls. Here there is less of a temptation to say that the bad nurse saved the patient. (Nor is it particularly tempting in this case to say that the good nurse saved the patient.) Can we say more generally that the phenomenon is far less prominent when the outcome is good? That is less clear. If we adapt the patient case in the main text so that, unbeknownst to both injectors, the person killed is Dr. Evil and so their death is a good thing, it is not clear that this makes a difference. It is also worth exploring how some of the other main text themes carry over to "save." Suppose Christina throws a ball out of a window and a freak gust of wind carries it next door and knocks unconscious a predator that was about to kill x. Isn't it a least a little more

well—what the nurse is doing is normal, but what the imposter is doing is abnormal. But the case can be tweaked so that the relative normality of what is being done is evened out. Suppose that the hospital is full of sickly enemy soldiers and the administration quite frequently allows government agents to interfere by adding fatal extra doses here and there to supplement drug A medication. Here the intervention by the non-nurse, while bad, may be quite normal. (And if we suppose that it is not normal for this nurse to be administering drug A—perhaps they were called in for the day from a very different unit owing to an unusual shortage of staff—then it may be that the nurse's intervention was more abnormal than that of the non-nurse.) The asymmetry in badness thus need not come with an asymmetry in abnormality, and still the effect of badness on causal and causative judgments persists.[20] Of course there is still a sense in which the bad actors are acting abnormally: Icard et al. (2017, 81) subsume cases of bad actors under abnormality principles by including violations of *prescriptive* norms as well as statistical abnormality within the category of the abnormal.[21] In any case, this phenomenon cannot be subsumed under the last as it is not about normal *continuations*.

It is also worth noting that the asymmetry persists even if we set things up so that the protagonists have the same subjective risk of doing something that has the patient's death as a consequence.[22] Suppose, for example, that the imposter has a million vials, knows one of them has drug A (the rest being saline), but doesn't know which. Hoping to kill the patient, the imposter (who

---

tempting to say that Christina saves x in this case than it is to say that the ball tosser killed x in a case where a freak gust brings about x's demise?

20. For some related points, see Knobe and Fraser (2008), where they critically discuss Julia Driver's (2008) suggestion that it is atypicality and not morality per se that explains the relevant data.

21. It's a good question how the bad/good principle interacts with the other ideas discussed here. For example, are we more willing to say x killed y when x kidnaps y and a tree freakishly falls on the back of the car while x is driving y than when x is taking y to the hospital and a tree freakishly falls on the back of the ambulance, killing y? Hart and Honoré cite plenty of case law in which an abnormal intermediary blocks causation by some agent even though some bad action by that agent is a factor. (See, for example, their discussion of *Toledo & Ohio Central R. Co. v. Kibler & Co.*, 163–64.) This suggests that making the initial action bad has little effect on the capacity of subsequent abnormal interventions to block killing. We shall not pursue this further here.

22. We use "consequence" in a way that is intended to be neutral between killing and not killing. That seems to fit with ordinary usage. If Jones drops a brick out the window which is then transported to Smith's head by a freak weather event, then even though we demur from "Jones killed Smith," we are much happier with "Smith's death was a consequence of Jones's dropping a brick out of the window."

has only time to use one vial before making their escape) by chance injects the vial containing drug A. Meanwhile the nurse knows there is an overdoser on the loose. But it's a big city and the temporal window for fatal double dosing is quite short, so their credence is one in a million that the overdoser will be a factor on this occasion. Here the subjective risks are the same but there is a clear asymmetry in our disposition to make judgments about who killed who.[23]

Accelerated Processes

Let us turn to the fourth idea. The key thought here is that when it comes to judgments of killing, there is a big difference between cases where one is a factor because one accelerates a death-culminating process that has already been initiated, and cases where one *initiates* a process.[24] Consider the following pairs:

A1: Smith has been bitten by a death adder. Jones (perhaps wishing to spare Smith suffering) squeezes Smith's extremities in such a way as to accelerate the action of the venom. As a result Jones dies one minute earlier than if Jones had done nothing.

A2: Smith has been bitten by a death adder. Jones (perhaps out of a desire that Smith does not suffer too much) injects a faster-acting poison into Jones's bloodstream (or, alternatively, shoots Smith through the head). As a result, Smith dies one minute earlier than if Jones had done nothing

B1: Smith is in a pot of water over a flame. Jones, a cannibal, wants dinner quickly. Jones adds a large amount of salt to the pot (which has been placed over the flame by other cannibals).[25] Smith dies one minute earlier than they would have if Jones hadn't added the salt.

B2: Smith is in a pot of water over a flame. Jones, a cannibal, wants dinner quickly. Jones (who is very strong) lifts the pot and moves it to a red-hot griddle some yards away. Smith dies one minute earlier than they would have if Jones had left the pot alone.

23. Of course if the nurse knows that the imposter is administering A then if they administer A themselves they are also at fault. So we don't want to even up subjective risk in *this* kind of way, since that would remove the moral asymmetry.

24. The ideology here is rather similar to Foot's concept of a sequence "thought of as somehow already in train" (1967, 3).

25. We realize that it is something of a folk myth that salt does much to speed up boiling. Readers who care about this should substitute some chemical that really does substantially speed things up.

In A2 and B2, we are more inclined to say that Jones killed Smith than in A1 and B1. (Similarly for various other causatives. For example, we are more inclined to say that Jones *boiled* Smith in B2 than in B1, and, assuming that both the venom and the faster-acting poison work by paralysis, we are more inclined to say that Jones *paralyzed* Smith in A2 than in A1.)[26]

Clearly this kind of contrast is playing a role in discussions about when various sorts of actions during palliative care constitute killing, suicide, and so on. Here is the Michigan Supreme Court:

> There is a difference between choosing a natural death summoned by uninvited illness or calamity, and deliberately seeking to terminate one's life by resorting to death-inducing measures unrelated to the natural process of dying. (cited by Cantor 2006, 411)

Here is the idea. Accelerating the "natural death" by, say, switching off life support or voluntarily lessening one's food and water intake, is a matter of the "natural process of dying" playing itself out. By contrast, measures that kill in a way that does not merely accelerate the natural dying process already underway, but which induces death by different means, are a completely different matter.

Hart and Honoré have a variant of the idea presented here. It is that one merely accelerates death (as opposed to "causes death") when (a) one slightly accelerates death and (b) one does something that "would not be sufficient to kill a person in a normal state of health" (1985, 344). This may in one way suggest an improvement to our own formulation, but in another way it does worse. In our cases, the process accelerated makes only a modest amount of difference to the time of death. But what if it made a difference of days rather than minutes? Here we do seem to be more willing to count the relevant action as a killing, but intuitions are still pretty murky. Suppose a snake venom was untreatable, but unaccelerated would take two agonizing weeks to kill someone. Someone massages their extremities, accelerating the passage of venom, and the victim dies a full week earlier than they would have done. How much

26. We note that there is also an interesting category of cases where one *enables* a process to culminate in a death but where it is not natural to say either that one initiates or that one merely accelerates. For example: (i) A boulder is tumbling down a hill. A car stands between the boulder and the eventual victim. One drives the car away, enabling it to kill the victim. (ii) Every home has a special filter that stops a deadly agent in the water supply reaching the tap—it gets as far as the filter but is then blocked. Someone removes the filter, enabling the deadly agent to reach the resident's glass, with lethal effect. We shall not attempt here to explore the variety of differential reactions in various enabling cases, though the materials presented here ought to be helpful for those interested.

more comfortable are we saying the massager killed the victim? We don't think the judgment is especially clear. Here it is worth noting that various palliative care activities that accelerate death by days rather than minutes are nevertheless *not* treated as killing (though it is unclear how much self-delusion is going on here).[27] Thus, while the length of time by which a life is shortened may make some difference in our willingness to say that x killed y, the judgments are not sharp enough to warrant further epicycles on our initial articulation of the idea. In any case, the ideas articulated here mark generic pressures for and against judgments of killing rather than clean exceptionless generalizations, and should be taken in that spirit.

In another way our articulation is superior. Suppose Smith is in a hospice. Jones wants Smith dead and lets a venomous snake into the bed. This is the kind of snake—perhaps a pygmy rattlesnake—whose bite will be survived when suffered in a good state of health, but is fatal to people who are as weakened as Smith is. Smith, as expected, dies from the bite. Even if it was true that Smith would have in any case died a very short while later from the illness that put them in the hospice, Jones killed Smith. But Hart and Honoré's formulation suggests a different (and incorrect) judgment, since a bite from that snake "would not be sufficient to kill a person in a normal state of health."

We note in passing that when someone's action is a factor but doesn't kill, we are—as Hart and Honoré's discussion suggests—often reasonably comfortable describing the case as one where one agent "accelerates" or "hastens" the death of another. (And we note that the expression "hastening death" is quite pervasive in palliative care literature.)[28] Could our main topic have indeed been presented as the contrast between killing and merely hastening death (or as the contrast between killing and merely accelerating death)? We chose our official formulation of the central topic in part because "factor" is clearly defined and in part to sidestep one tricky issue. Verbs like "hasten" and "accelerate" are typically treated by grammarians as causatives. Yet while the ideas about killing presented above apply to all sorts of paradigmatic causatives, they do not seem to apply so smoothly to "hastening" and "accelerating." As we have just noted, the cases where someone's action is a factor in another death and they don't kill them, we are often more comfortable saying they accelerate death and that they hasten death. For example, in case A1 above, we are very comfortable saying the person hastens the death by

27. Thus Cantor writes: "It is self deception if people think they are not killing anyone when they deliberately choose a regimen of treatment which they know will result in the patient's death when there is an alternative which will keep the patient alive" (2006, 411).

28. See, for example, Cavanaugh 1996, Cantor 2006, and Billings 2011.

massaging the extremities. But the connection between being a factor and ordinary uses of "hasten" is still, arguably, imperfect. For example, many informants felt that one would be taking semantic liberties if one said "x hastened y's death" in a case where x takes y to the beach and a sniper shoots y.[29] It is thus important to realize that the ordinary use of "hasten" is less demanding than paradigmatic causatives but is arguably more demanding than "being a factor." Because of these complications we urge readers not to assume that "being a factor" can be glossed in terms of "hasten" or "accelerate."

There is a related question. Could our core topic have been posed as the difference between being a factor in someone's death and causing someone to die? The discussions cited by Hart and Honoré and by Knobe would seem to suggest that the kinds of consideration that make us fall short of calling a factor a killing would also make us reluctant to say that the relevant agent caused someone to die by performing the relevant action. But the issues are delicate: Take the case where Jones's brick is redirected by lightning. Consider: (i) Jones killed Smith. (ii) Jones was the cause of Smith's death. (iii) Jones caused Smith's death. (iv) Jones was a cause of Smith's death. The fourth sounds quite a bit better than the first to our ear and the third a bit better than the first (though we may have been corrupted by philosophical fashion). A famous example from Katz has generally convinced linguists that "cause to die" and "kill" can't line up perfectly: Suppose a gunsmith's defective work causes a sheriff's gun to jam, with the consequence that the sheriff is gunned down. "The gunsmith caused the death of the sheriff" sounds a lot better than "The gunsmith killed the sheriff." (An example adapted from Byrne [2021, 23] is also very suggestive: It is very dark. Jones turns on her headlights to help find Jackson, who is lost. The illumination allows Johnson, a villain, to see Jackson and to shoot Jackson. Jones can say "I caused Jackson's death" but can't say "I killed Jackson.") A common reaction among linguists is to suppose that "kill" requires "direct causation," but as Neeleman and Van de Koot (2012) point out,

29. Note that the canonical use of "x hastens y" is one where y is an event and that this is certainly not the case for "x boils y," "x melts y," and "x kills y." We also note that the substantial literature in linguistics on the relation between causatives and so-called periphrastic causation may be very relevant to understanding how "hasten" (and "accelerate") work. Periphrastic causative constructions are ones in which a clause that specifies an effect serves as grammatical object (e.g. "The sergeant made him march," "The director caused her to quit"). The consensus seems to be that periphrastic constructions are also less demanding than paradigmatic causatives. For a helpful survey of this and related issues, see Martin and Schäfer 2014. The reason that this literature on periphrastics is relevant to "hasten" is that it is plausible that hastening can be analyzed in terms of the periphrastic "make": while hastening someone's death is not, in general, a matter of making the death happen, it *is* a matter of making it happen sooner.

this can't be right: If we add to the story that the gunsmith plotted the death of the sheriff and deliberately tampered with the gun, the causal connection is no less indirect, but the "The gunsmith killed the sheriff" is much more acceptable. (Recall also that Stalin killed millions.) It is also not clear that a lack of intention systematically induces a refusal to apply "kill" in more indirect cases: A recent discussion by a lawyer of product liability contains the turn of phrase "when a company kills through its negligence."[30] And similar examples are easy to find.

How great a mismatch *is* there between "kill" and "cause to die"? Is the kind of causal relation implicit in causatives like "kill" and "boil" *vastly* more demanding than the relation typically expressed by "cause" (in which case, "cause," though a causative, is perhaps a bit more like "hasten").[31] This is something like Lewis's view, according to which causing people to die is a commonplace achievement: "I am sure that I—and likewise you, and each of us—have caused ever so many people to die, most of them yet unborn. But I have never killed anyone—I hope . . . so killing must be a special kind of causing to die." Or is the relation expressed by "cause" in ordinary discourse much more demanding than that suggested by approaches like David Lewis (1973), so that the extension of "caused y to die" and that of "killed y" is at least much closer than Lewis's discussion would suggest?[32] Because we don't wish to try to sort all this out here, we shall continue to focus on the distinction between

30. See Fieldfisher, "New Law to Prosecute Negligent Companies," August 22, 2006, https://www.fieldfisher.com/en/injury-claims/insights/new-law-to-prosecute-negligent-companies.

31. Note, though, that if one likes the analysis of "hasten" in terms of "make happen sooner" proposed in fn. 29, whatever contrast there is between the periphrastic "cause" and the periphrastic "make" will make for interesting potential contrasts between "hasten" and "cause." For more on this, see the appendix.

32. Note that Lewis (1973) famously took counterfactual dependence between events x and y to be sufficient (though not necessary) for the truth of "x caused y." If he is correct, then "Jones's dropping of the brick caused Jones's death" is true in both versions above. Once we have got that far it is hard to resist accepting "Jones caused Smith's death." (A natural bridge principle is that if x is the agent of an event that caused some event y, then "x caused y" is true.) It bears emphasis that "Jack caused Jill to die but Jack didn't kill Jill" doesn't exactly roll off the tongue, which one might expect it to, given Lewis's vision. However, examples such as Katz's Wild West story suggest, prima facie, that the relation between "cause to die" and "kill" is not quite entailment. Knobe's (2009) picture is that folk judgments of causality are driven by a selective filtering of counterfactuals—it's a matter of the right counterfactuals breaking a certain way. Knobe's and Lewis's approaches can be made compatible if we suppose that the truth of causal claims in English come apart quite dramatically from ordinary folk's willingness to accept them. It is beyond the scope of this paper to inquire as to which aspects of ordinary practice should be semanticized.

cases where one kills and cases where one is a factor in a death but doesn't kill, and will not pass judgment on the relation between "killing" and "causing death."

It is also worth making salient one reason why our taxonomy can only give rather soft predictions about folk judgments using causatives. Suppose that Jones makes a defective television that then explodes in the presence of Smith. Smith dies. Candidate uses of "kill" include "The television killed Smith" and "Jones killed Smith." But such contexts have a "pick one" flavor. Here there is a choice between an "automatic agent" (recalling Thomason's ideology), namely the television, and a bona fide agent, Jones. But it is hard to pick both: The speech "The television and Jones both killed Smith" is very awkward.[33] For the purposes of product liability practice in tort law—one kind of setting which animates Hart and Honoré—questions framed directly using causatives like "kill" are often sidestepped in settling liability in favor of expressions along the lines of "is responsible for," a phrase that figures prominently in Hart and Honoré's own discussion. But if one is in a context where one wants to deploy the causative "kill," then one will have to pick either the television or Smith to fill in "X killed Smith." And the above discussion doesn't really provide that much guidance as to how ordinary folk resolve the competition between the television and Smith. For what it's worth we think in this case it is much more likely they will go for "The television killed Jones," but the factors resolving the competition between animate and inanimate agents as the preferred agent of the causative seem quite subtle and we cannot hope to fully catalogue them here.[34] When the activity of the inanimate object is subsequent to the animate agent and is itself an abnormal intervention, there will be strong pressure to point to the inanimate agent. Thus, when a freak wind or lightning bolt carries the brick to Smith's demise, we say it is the wind or the lightning bolt that killed Smith, not the person who dropped the brick out the window. But where there is no abnormal intervention of that sort we may often *still* choose

33. Contexts in which one says "the cause" also have a "pick one" flavor: in this case, the "pick one" aspect is particularly salient because of the appearance of the definite. And the plain "caused" also plausibly takes on this profile.

34. Consider the following pair, which also point to some subtleties, this time involving animate agents: (a) Smith throws Jones through a pane of glass. Here, "Smith broke the glass" and "Jones broke the glass" both seem passable. (b) (Adapting a case from Hart and Honoré, though their focus is on "caused the glass to break" rather than "break the glass"): Smith hits Jones, "who staggers" and then falls against the glass. Here, even though "the second agent's role is hardly an 'action' at all" (1985, 76), "Smith broke the glass" does not sound acceptable.

the inanimate factor over the animate one on account of its temporal proximity.[35]

Of course one might think that "Smith killed Jones" is true, even if not asserted, in a context when one says "The television killed Jones." But that makes it puzzling why "The television and Smith killed Jones" is not felicitously assertable. Granted "The television but not Smith killed Jones" is awkward: But "The television and not Smith was the cause of Jones's death" is awkward too—and in the latter case it seems fairly clear that there is a semantic requirement of uniqueness. Notice also that the "pick one" phenomenon even applies to many cases where there is a pair of candidate *animate* agents. Suppose Stalin sends a note to Jones instructing Jones to kill Smith. There is a natural context in which we say "Stalin killed Smith" and one in which we say "Jones killed Smith." But "Stalin and Jones killed Smith" is only natural in contexts where neither's contribution was sufficient and it took teamwork to kill.[36] We shall not pass final judgment on the full extent of context-dependence here. It is enough to notice that when it comes to making causative judgments, folk often find themselves choosing between certain animate agents and temporally more proximate agents, and the resolution of this competition is quite delicate.

We now have a preliminary taxonomy of cases where x is a factor in y's death but does not kill y: Following the subsection labels, let us call the first kind of case above "mediating agent cases," the second "abnormal event cases," the third "bad action cases," and the fourth "accelerated process cases." Perhaps

35. Notice our contrasting inclinations between a case where Smith pushes Jones on a spike and Jones dies and a case where Smith sends Jones on a diplomatic mission to Liechtenstein and Jones gets killed by a meteor. The spike is already in place when the push occurs. But the meteor may well be heading to Liechtenstein when the diplomatic decision is made. So the cases cannot be straightforwardly distinguished by saying that in one case but not the other, there was an abnormal intervention posterior to the act that was a factor. Some more folksy conception of what is and isn't already "part of the scene" may be playing a role. The fact that the inanimate "agent" is in motion in one case and not the other may also be making a difference to our judgments. In a case where Smith pushes Jones and an enormous hailstone falls on Jones a few seconds after—a hailstone that was twenty meters above the place where Jones died when the push occurred—we are already less inclined to say that Smith killed Jones.

36. There also seem to be subtle pragmatic differences between causatives. As Sarah Moss pointed out (in correspondence), even if everyone killed by Stalin was killed by gunshot, it is not very natural to try to get at the relevant number by asking "How many people did Stalin shoot?" This kind of pragmatic messiness makes it especially difficult to expect any very clean general answer about the demandingness of causatives.

the taxonomy is not completely exhaustive. But it will provide us enough to work with when probing the moral significance of killing.

In what follows we shall focus our discussion on the second and fourth of these categories. In a mediating agent case, it is very natural to suppose that the badness, responsibility, level of guilt, etc., of the original agent is somewhat diminished, since the mediating agent takes on at least some of the guilt and so on. Meanwhile, if badness of the agent contributes to whether an action counts as a killing then it is quite obvious that in *this* way, being a killing rather than a mere factor can have moral significance, but for uninteresting reasons, since in bad action cases it is the badness of certain agents that explains why they—rather than other relatively good agents—count as killing. But how about the second and fourth subcases? In the next section we look at these two subcategories, with an eye to exploring the moral importance of the distinction between those cases and cases of killing.

Is the Distinction Morally Important?

Is there a morally important distinction between the class of abnormal event cases and the class of cases that are killings? And is there a morally important distinction between accelerated process cases and the class of cases that are killings? In this section we attempt to probe these questions, tentatively arriving at negative conclusions. Of course we haven't provided anything like informative necessary and sufficient conditions for "to kill" and there may also be some measure of semantic context-dependence and vagueness. Yet we have provided what we hope is a helpful guide to the sorts of considerations that help contour the extension of "to kill" in the mouths of English speakers, and our investigation will proceed on the assumption the truth conditions of "kill" sentences are sensitive to the presence and absence of abnormal interventions and to the folk distinction between initiating and accelerating a process.[37]

We think it might be helpful to look both at cases where the potential victim's early demise is regarded positively and also ones where it is not. We begin with cases where the protagonist does not want an early demise for the potential victim but performs an action where there is a risk of early demise. Call these "mere risk cases." We then look at pairs of cases where an earlier demise

37. We are aware that one might think that the relevant factors contour use but do little to contour truth conditions. Readers are free to think matters through from this perspective, though it will matter what the relevant truth conditions (ones that do not closely align with use) turn out to be.

is aspired to but not foreseen. Call these "intentional but unforeseen cases." We finally look at a choice between actions that are known to be a factor in a victim's death but where only one is a killing. Call such cases "knowledge cases."

Mere Risk Cases

If one takes a small risk of producing an outcome but is not in any way motivated by that outcome, one neither produces that outcome knowledgeably nor intentionally. Is there an important ethical difference between risking killing someone and risking being a factor in their death?

Consider the following case. (It is obviously extremely stylized, but extremely stylized cases are often useful for probing intuitions):

Ski Slope

> One is skiing out of control down a ski slope. One can't but collide with Smith but one can control whether one knocks Smith down the right or left slope of a hill. At the bottom of the left slope it is known that there are no spikes but just very occasionally, trees fall onto that area.[38] At the bottom of the right slope there is no chance of being killed by a falling tree, but just very occasionally, people fall on spikes. Suppose, as one's credences go, if one knocks Smith to the left, there is zero chance Smith will die from spikes and a one in a million chance that Smith will die from a tree falling on them just after they reach the bottom, and if one knocks Smith to the right, there is zero chance they will die from a falling tree but one in a million chance they will die from a spike.

If one attached special disvalue to killing, one would expect to have a clear preference for risking death from tree fall to death from a spike. Indeed, if one attached special disvalue to killing, one would presumably go for the tree-risk option over the spike-risk option, even in certain circumstances where the risk of death from tree fall is *higher than* the risk of death from spikes. (Of course, how much higher a risk one would tolerate would depend on how much extra disvalue one attaches to killing in particular.) Suppose, for

38. We acknowledge that it is important that the trees only fall occasionally. If the trees are falling left and right to make a gauntlet of trees (picture an Indiana Jones movie), pushing the skier toward the trees now seems more like pushing them downhill toward a tornado. Meanwhile, if for any reason you are having difficulty with the intuition that you kill the skier if things go badly with a spike, then make the temporal gap between push and spiking short, so that it is as similar as possible to the spike scenarios described earlier.

example, that we had the choice between giving Smith a one in a million risk of death from a spike or instead one in half a million risk of death from a tree falling on them. We submit that it seems quite strange to go for the tree risk over the spike risk on the grounds that the eventuation of the tree risk would not constitute a killing and the eventuation of the spike risk would. It's not that we can't imagine a person who attached special disvalue to killing per se in this way. Such a person would say "Granted, it's twice as likely that Smith will die if I avoid the spike risk in favor of the tree risk. But there is no risk of my killing them with the latter option, only of my accelerating their death by pushing them to where a tree is about to fall; so that is obviously preferable." But we submit that this is a rather strange way to think in a scenario like this. This preference for tree risks seems quite fetishistic when it involves putting someone's life at greater risk.

The pattern persists even when the risk differential is reduced. Suppose a morally good skier was in the unfortunate position of having to either expose Smith to a one in a million chance of death from spikes or a one in eight hundred thousand chance of death from tree fall. It would be rather shocking for that skier to expose Smith to a greater risk. This indicates to us that when one looks head on at the distinction between killing and being a mere factor (thanks to an abnormal event), it is hard to give it moral weight.

Let us turn to a case that involves the choice between a risk of killing and a risk of accelerated process (without killing).

Cannonball

> A heavy cannonball is rolling down a chute from the top of a hill to the bottom, placed there by a malevolent agent who has ill designs for Smith. One knows it is heading toward Smith and that if one does nothing, Smith, who is tied up at the bottom of the hill, will die at 2 p.m. One is motivated to save Smith. One has a choice. (a) One can pull a lever that simultaneously lifts the top of the chute at the top of the hill and moves it slightly rightward. One is very confident that this will save Smith by redirecting the chute away from him. But there is a one in one hundred chance that it will simply accelerate Smith's death by making the chute steeper, thus accelerating the descent of the ball but failing to redirect it away from Smith. In this case Smith will die at 1:55 p.m. (b) One can direct a second cannonball down a second chute. One is 99/100 confident that things are set up so this cannonball will, farther down the hill, collide with the first chute, redirecting that chute and saving Smith. But there is a one in one hundred chance that things are set up so it will instead head toward Smith at an even faster rate, beating the first cannonball to the bottom of the hill, so that Smith dies at 1:55 p.m.

Here the dialectic is rather similar to that of the previous case. Suppose one attached disvalue to killing vis-à-vis death by accelerating a process. Then one should presumably opt for the lever over the second cannonball and maintain that preference even if the risk of death is a bit higher. Disvalue killing enough and one should prefer the lever over the cannonball even if the risk of Smith dying is double on that option. And more generally, if one attaches special disvalue to killing, one would expect there to be some extra risk such that one would prefer the lever over the second cannonball even though the former adds that extra risk of death vis-à-vis the second. Once again, it seems very strange to imagine opting for the riskier option on such grounds.[39] Wouldn't we look askance at an agent who exposed Smith to the greater risk when a second cannonball, with lesser risk, was available?

Let us consider a few ways that one who attached special importance to killing might respond.

First, there is logical space for a view that attaches special disvalue to killing but that opts for the least risk of death when there was a risk of death mismatch among options—on this view, any special disvalue of killing only comes into play when the risks of death *exactly match*. This is a lexical priority view where only a perfect match in expectation of death would lead one to proceed down to the lexically secondary question of whether an event, if a death, is a killing.[40]

39. Of course, the fact remains that we nevertheless are inclined to feel negatively about someone who shoots another, even if they would otherwise die soon, and even if the motive is benevolent. One thing that may be going on here is that given the constraints of human architecture, a willingness to shoot here may be a bad portent for someone's dispositions elsewhere. We also note that there are moral theories that give such architectural constraints a constitutive role in foundational moral theorizing. In this connection we invite readers to explore what the motive consequentialism discussed in chapter 4 would have to say about section 3 of this chapter.

40. Here is an example from Sarah Moss (in correspondence) where lexical priority style thinking is tempting and which may more generally be suggestive: "Consider the fact that I value helping Oliver with homework, pushing him on the swing set, and so on. To me, there's an important moral difference between me helping him and a babysitter helping him; I really value it being me who helps. But if you tell me, 'Hey, you have a choice: Either a babysitter will have time to help Oliver with eleven math problems, or you will have time to help him with ten problems,' I feel guilty saying, 'I value being the one who helps Oliver, so I'll pick the latter.'" (For the record, Moss was inclined to think there is an important contrast between first- and third-person perspectives between the deliberator and the external assessor, both in the math problem scenario and in our cases above.) If you were looking on and someone you cared about was at the bottom of the chute, would you feel very differently about one of the cannonball choices? Readers are invited to explore the differential impact of such perspectives. (For what

It is clear enough, though, that this view risks severe costs by the standards of common-sense morality, and so cannot serve as its bulwark.[41] This view suggests, for example, that if you can throw a live grenade into someone's apartment that has a slightly less epistemic chance of inducing an early death than leaving things to an assassin lying in wait, then one must throw the grenade. (Imagine things are set up so the assassin will walk away if one throws the grenade without checking whether the relevant person died, and that the assassin's grenade is a bit more powerful.)[42] Note moreover that the natural way of cashing out the view under discussion requires one to prioritize whichever action generates more longevity, so long as there are not obvious other asymmetries as far as pain and suffering are involved. But that means that one can be positively required to kill in cases where it is clear that the person would have lived a bit less long otherwise. And this will be disturbing. We can use variations on the canyon/assassin case from earlier to make the point, so long as we tweak the case so there is no serious suffering involved in the canyon option (perhaps the person is drugged and sleeping). Note, finally, this view risks giving killing hardly any explanatory work to do, since expectations of death are very rarely exactly the same between options.

A second idea is to appeal to interaction effects: Perhaps (recalling the phenomenon of "secondary permissibility") when the alternative option risks hastening death, that negates what would be the normal moral drawbacks of an option that carries a similar chance of killing. But that again may have unwelcome results. Suppose that one has the forced choice between the quite

---

it's worth, we don't feel much impact when shifting to a third-person perspective in the cannonball case.) See also our brief discussion of reactive attitudes below.

41. That said, the lexical priority view might be offered as a replacement to common-sense morality rather than as a way of trying to vindicate it. Similar remarks apply to various other ideas canvased below.

42. In this way common-sense morality arguably does not recommend doing what the endangered person would want you to do if they knew what you know. If the potential victim knew you know they would surely want you to throw the grenade as that gives them a better chance of survival. (This is at least true if they had no special relation to you rather than the assassin. If they were friends with you and sufficiently captivated by common-sense morality they might have a self-sacrificing thought along the lines of "I wouldn't want my friend to have blood on their hands.") Perhaps common sense gets a bit more confused if we set things up so that the potential victim does know what you know and actually asks you to throw the grenade. But here many would still be horrified at someone who acceded to the request. This illustrates the potential conflict between concern for others and special prohibitions on killing that is also explored in Hare (2016), though his focus there is on variants of the famous trolley case where one has the option to kill one to save many.

normal practice of withholding food from someone in a hospice so as to accelerate death, and injecting them with a lethal poison. The fact that there is a hastening as an alternative does not make one comfortable injecting. So again, whatever its merits, this appeal to interaction effects is not a route to securing common-sense morality.

But there are many other ways in principle of trying to exploit interaction effects. The line of thought (involving ski slopes and cannonballs) we have just explored involves what Shelly Kagan (1988, 5) calls a "contrast argument" that "proceeds by offering a pair of cases that differ only in the factor in question." As he notes, the most straightforward way of advancing a contrast argument presumes that if "we judge the cases to be similar . . . despite the variation in the given factor, this shows that the factor is not morally relevant at all" (1988, 6). But as Kagan makes clear, it is dangerous to assume that if a factor sometimes makes a difference, it always makes a difference.[43] He is certainly right. Claims of relevance outside ethics are instructive here: Being fork-shaped is relevant to having utility in the kitchen. The fact that there is a pair of microobjects that differ in that one is fork-shaped and the other isn't, both of which are equally useless in the kitchen, does not destroy that relevance thesis. The size of a painting is relevant to how much it will fetch in auction.[44] This hypothesis is not negated by finding a completely botched painting and noting that its size-enhanced counterpart has no greater auction value.[45] For

43. One thing not to do, Kagan warns us, is to commit the "Additive Fallacy" (1988), whereby one hopes to determines the value of a situation be adding up separate positive and negative contributions of the relevant factors. (This would be to ignore interaction effects.) Kagan makes many helpful observations in this connection. One helpful analogy he offers is chemical: an element might only play a certain role in certain chemical environments, so that in other environments its presence or absence may be neither here nor there.

44. Perhaps quite a bit of the confusion stems from the fact that claims like "The size of a painting is relevant to how much it will fetch in auction" can sometimes convey "sometimes relevant," sometimes "always relevant," and perhaps more often something along the lines of "generically speaking, relevant." More clarity is needed on what it means to say a factor is relevant (ethically, aesthetically, etc.). These issues do not go away when we switch to "intrinsically/constitutively relevant." See Appendix B.

45. In this case a certain aspect of the situation nullifies the contribution of size. In other cases, some such aspect enhances the import of a property. For an illustrative example within ethics, recall Thomson's (1976, 204) well-known example where the import of punching someone in the nose is upgraded (again, designed to show the fraught nature of contrast arguments). The key example is one where, owing to an unusual physical condition of the victim, the assailant knows that punching the victim in the nose will kill the victim. Thomson observes correctly that in such circumstances there is no real contrast between punching and decapitation. But to conclude that nose-punching is no worse than decapitation is improper.

now we don't want to fuss too much about quite what extra is added to a claim of moral relevance by such qualifiers as "intrinsic" and "constitutive" (for more on this see the second appendix). But however these play out, it seems pretty clear that we may want to count a property as ethically significant even if it only made a difference in the presence of certain other factors and so could not be relied on to make a positive moral contribution in any environment. We thus need to tread cautiously.[46] We do not take contrast cases to be demonstrative, and so what we say above and below can at best be evidence against the ethical significance of killing.

Intentional but Unforeseen Cases

One can intentionally do things without knowing that one will do them. Suppose, for example, that Jones wants Smith dead. Jones has a crossbow. They know they are a bad shot—at the kind of distance that Smith is, they tend to miss nine times out of ten. But Jones gives it their best try. They aim as carefully as they can, pull the trigger, and the arrow pierces Smith's heart, to lethal effect. Jones doesn't know that they will succeed in killing Smith but Jones does so intentionally.[47,48]

46. And there are obviously pairs of cases where the probabilities of death are the same but where we have very different reactions. Pulling the trigger in a Russian roulette situation where there are an enormous number of chambers and just one bullet feels a lot worse than sending someone out to pick up a burger, but the chances of the relevant person dying may be very similar. Assuming there is a moral difference, what is its source if not the killing/non-killing distinction? This challenge is not hard to answer—on a natural elaboration of the Russian roulette story, the perpetrator attaches special value to the death of the agent, but the natural elaboration of the burger example is not like this.

47. A more controversial thesis is that in these cases one *intends* the outcome. In the theory of intentional action, the "Simple View" that intentionally F-ing entails an intention to F is hotly contested. (The label the "Simple View" is due to Bratman [1984], who contested it.) For proponents of the Simple View (including Adams 1986 and McCann 1991), the inference from intentionally killing to having an intention to kill is utterly straightforward in any particular case, because it is underwritten by entailment. For the denier, things are less straightforward.

48. Seen through the lens of a Hart-and-Honoré-style framework, the reason this intuitively counts as a killing is that while it is abnormal for her aim to be accurate, the fact that she is pointing in the right direction just as she is finishing pulling the trigger is part of the scene, not an abnormal event subsequent to the pulling of the trigger. If she aimed in the wrong direction but a bolt of lightning or a freak tornado redirected the arrow, an attribution of killing would be much more dubious. But when there is no freak intervention, killing can be seen as one of many normal continuations (it is important to think of normality in an existential way here—recall fn. 13).

One might take the view that while killing does not in general have disvalue compared to accelerated processes and abnormal event cases, there is special disvalue to *intentionally* killing vis-à-vis intentionally accelerating death via abnormal events or by accelerating lethal processes. (Note that this is an example of a candidate interaction effect. Maybe the contrast between killing and merely accelerating death is unimportant when it comes to unintentional cases, but takes on considerable significance when the relevant act type is intentional.)

Here we can look at some pairs that manifest this contrast.

Cannonball 2

A heavy cannonball is rolling down a chute from the top of the hill to the bottom, placed there by someone who has ill designs for Smith. Jones knows it is heading toward Smith and that if Jones does nothing, Smith, who is tied up at the bottom of the hill, will die at 2 p.m. Jones wants Smith's death to come sooner rather than later. Jones has two ways of trying to achieve this. (a) Jones can pull a lever that lifts the top of the chute. She is .2 confident that it will accelerate the cannonball by sharpening the steepness of the chute, so that Smith dies five minutes earlier than if Jones had done nothing, and is .8 confident that it will make no difference. (b) Jones can direct a second cannonball down a second chute. Jones is .2 confident that it will travel much faster to the bottom of the hill than the first cannonball, so that Smith dies five minutes earlier than if Jones had done nothing, and .8 confident that it will be slower and that the first cannonball will still kill Smith at 2 p.m.

Ski Slope 2

Jones wants Smith dead. Jones is deliberating between two options. Jones can push Jill down the right slope of a hill into Avalanche Valley where Jones is 1 in 10,000 confident that an avalanche will occur shortly after Smith gets to the bottom of the slope from which she will not survive.[49] Or Jones can push Jill down the left slope of a hill into Spike Valley where she is 1 in 10,000 confident that Jill will fall on a spike and die. (Jones is certain there will be no avalanche in Spike Valley and that there are no spikes in Avalanche Valley.)

49. Avalanches are exchanged for falling trees to provide a bit of variety—there is no significance to that choice here. Avalanche Valley is so called to help the reader remember the case. The idea isn't that avalanches are completely normal there. Let us imagine that there is only an avalanche only once every year, though when there is one it kills everyone at the bottom of the valley.

Again, holding Jones's desire for Smith's early demise fixed, we struggle to see a significant moral difference between the two options in each of these cases. Suppose, for example, Jones opts for the lever and Bill dies at 1:55 p.m. It would sound rather hollow, as an attempt to argue for a lesser gravity of sin, for Jones to make the speech, "Well, at least I didn't kill Smith." Now, of course, the moral gravity of the cases can be varied by varying Jones's motivational structure. A more sinister (and natural) version of the case is one where Jones has ill will toward Smith and wants Smith dead sooner rather than later for malicious reasons. But alternatively, we could fill in the details so that a fast-moving cannonball produces a more painless death and Jones wants to accelerate death because the death then involves less pain.[50] Varying the cases along a parameter like this makes a difference to the cases. But so long as the cases are symmetric along this parameter, the mere contrast between intentionally killing and intentionally accelerating death does not seem to have much of a moral grip.

Knowledge Cases

One could in principle opt for the view that there is an important moral contrast between knowingly bringing about a killing and knowingly bringing about a death that is not a killing. (This is another candidate interaction effect: The property of killing only really comes into its own when the killing is knowingly performed.) Does a suggestion along those lines hold up?

When it comes to the knowledge cases, it is hard to find versions of the mere acceleration type that fit the "abnormal event" type. Recall that when the agent has advance knowledge of the relevant abnormality, the conjunction of the death and the abnormality will typically be "designed" by human agency, and so the abnormality downgrade from killing to mere factor will be hard to achieve. But it is easy to contrive foreknowledge cases that are accelerated process cases and not killings. For example, one can easily tweak the cannonball case so one has a choice between pressing a lever that one knows will accelerate the initial cannonball and rolling a second cannonball that one knows will arrive ahead of the first cannonball. Once again, we can fill out the case either by imputing wholly malicious motives ("I want Smith dead as soon as possible because I can't stand Smith") or more noble ones ("If I do nothing

50. In this variant, the claim that the death is accelerated intentionally is perhaps a *bit* more tendentious. But for what it's worth, our judgment is that it is still straightforwardly correct. (Note that only in the presence of the Simple View—see fn. 27—does this judgment conflict with Double Effect sensibilities.)

this death will be a slow and agonizing one while a fast-moving cannonball will be far less painful"). But holding fixed the motives, we find it extremely difficult to see a moral difference between the two cases. As we vary motives across cases, we get markedly different reactive attitudes: Our reaction to someone who is concerned about agony is very different to someone who values the shortening of a life as an end in itself. But if we hold motives fixed and the known duration of acceleration fixed, and merely vary whether there is known acceleration without killing or known acceleration with killing, we find it much harder to discern a palpable moral difference.

Now, of course, for many warm-up illustrations of the contrast between killing and mere counterfactual dependence, there is a marked epistemic contrast. When Jones takes Smith to the beach and Smith dies from sniper fire, Smith's death is utterly unforeseeable by Jones. By contrast when Jones shoots Smith through the head, Smith's death is utterly foreseeable. Nevertheless, once we ensure epistemic parity by finding cases of killing and cases of mere acceleration when the outcomes conditional on each choice are known to be the same, a moral difference between killing and mere acceleration in these cannonball setups is difficult to discern.

One might wonder whether a moral contrast between killing and being a factor might show up more vividly if we asked ourselves "Would you do this to save a life?" We can set a cannonball case up so that a button toward the bottom of a chute will release a life jacket that will save someone who would otherwise drown in a pool below. In one version of the story we need to raise the chute in order to release the life jacket, since the button needs a fast-moving cannonball to make it work. The person at the end of the chute will then die a little earlier. In a second version of the story, we need to send a cannonball down a second, slicker chute to release the life jacket, since it is the second chute that has the button. In the first version you don't kill, but in the second you do. But again, we find it hard to muster a difference in inclinations between the cases.

Let us take stock. We have provided at least some reason to think the distinction between being a mere factor in a death and killing someone is not a distinction with any straightforward moral significance. Granted, our ordinary practice drives a wedge between killing and being a mere factor. It is unsurprising that we have a concept like killing: The phenomenon of being a factor is so ubiquitous as to make it rather challenging to use as a go-to concept for practical decision-making. But having probed the moral import of that wedge, we have so far come up relatively empty-handed. We realize that some will be unconvinced. Perhaps they will say the intellectual character of our contrast cases somehow dulls our moral sensibilities to the true moral import of killing. And one might somehow hope that there are other environments where the

special moral importance of killing will be thrown into sharp relief. But the challenges to maintaining a conservative view should not be underestimated.

It is worth remarking that our (albeit tentative) judgments about the moral insignificance of the killing/factor distinction are plausibly reinforced by the realization that the distinction between killing and being a factor implicit in ordinary practice is extremely soft from a metaphysical point of view. If Jones pushes Smith onto a spike, Jones kills Smith, even if Jones does not see the spike and even if the push was not in malice (perhaps Jones and Smith were engaged in a friendly bout of outdoor judo). The spike may have been left by a malicious person who was titillated by the idea of someone falling on it. Nevertheless, even when informed of this, we suspect people will judge that Jones killed Smith. Suppose instead that Jones kidnaps Smith, turns the key of the car, and a car bomb that Jones was unaware of explodes, placed by a person who was titillated by the idea of the people in the car getting killed. Here we suspect that ordinary people will be less inclined to say that Jones killed Smith. But the contrasts driving these judgments seem insubstantial. Hart and Honoré's discussion indicates why the judgments might be breaking this way: The explosion is an abnormal event subsequent to Jones doing what they did. But the spike was "an existing state of affairs rather than an intervening event." The thought driving all this is presumably that the spike didn't have to do anything subsequent to the push—*it just sat there,* whereas the bomb had to do something subsequent to pushing Smith into the car, namely explode. But we find it very difficult to think that this kind of folk distinction can stand up to much scrutiny. Suppose Jones had pushed Smith into a car that happened to be full of lethal gas (again with no lethal intent). We can easily imagine a significant portion of ordinary people assimilating the gas to the "existing state of affairs" category and judging Jones to have killed Smith. But there is a very natural sense in which the gas had to do something subsequent to the pushing in order for Smith to die—it had to enter Smith's lungs and cause chemical burns, etc. Another example: Suppose Jones takes Smith to Liechtenstein and there is a large asteroid that, unbeknownst to Smith, is already on a collision course with Liechtenstein. Again, many will demur from saying that Jones killed Smith when the asteroid kills Smith (along with the rest of the population of Liechtenstein). But to think that the spike just sat there and didn't have to do anything while the asteroid had to do something does not seem very respectable, even if folk judgments are driven by initial takes on the cases along those lines. Moreover, sensibilities about what is and is not part of the scene may also be doing a lot of work. Let's change the spike example so the spike is rapidly alternating between poking out of the ground and then retreating into the ground, with some electric motor driving the process along. Every two seconds the spike pops out of the ground. If Jones

pushes Smith and Smith dies because the spike is accelerating out of the ground, it is still much more natural to say that Jones killed Smith than if Jones takes Smith to Liechtenstein and then Smith dies as a result of an asteroid or nuclear missile heading that way. But it's very hard to think that there is a contrast with moral weight here.

Hart and Honoré are themselves certainly alive to the possibility that the key folk ideas about causation that drive doctrines about causation in the law will not look good once fully understood:

> It may, of course, well be that when we thoroughly understand the commonsense notions of causation we should no longer wish our thought on any matters, let alone legal judgments of responsibility, to be dominated by them: we may think that they are vague, crude, or anthropomorphic, or all of these: they may be "the metaphysics of the Stone Age" which should be replaced by modern notions of probability or "risk." (1985, 1–2)

In our view, such concerns are far from misplaced.

Let us briefly address two objections to the pessimistic drift of our discussion. First, one might worry that the preceding reflections suggest that there is something wrong with hoping not to kill people (as opposed to hoping to not be a factor in a death). Even if we are certain, or almost certain, that we will be a factor in some death, is it not still reasonable to desire/hope and so on that we will not be killers? If the preceding reflections make such hopes and desires misplaced, this would certainly be jarring. In response, note that our central normative claim was that there is no difference in badness between killing and being an abnormal event or accelerated process factor when things are otherwise *epistemically and motivationally symmetric*. But that does not mean that the news that someone has killed tells us nothing about their epistemic and motivational situation. It is very easy to be a factor in someone or other's death and almost inevitable that each of us will be. Conditional on the news that we will be a factor, that doesn't do much to increase our expectation that we will be bad people in the future. But conditional on the news that we will kill, matters are different. A factor is much more likely to count as a killing if the relevant death is proximate to the action. Close the temporal gap between the action and the death and there is likely to be less reliance on abnormal intervening events. Make the gap great, and it is more likely that abnormal intervening events play a role. Of course, this isn't always the case (imagine we plant a bomb set to go off in a year). But killings *tend* to be much more proximate to the relevant death than that. Further, conditional on a death being proximate to our actions, it is more likely that it is intended or at least foreseen—and in those latter scenarios our moral status tends to be compromised. And let us not forget our first and third themes (mediating agent cases

and bad action cases): For those reasons as well, the conditional probability of our being bad conditional on killing is higher than conditional on being a mere factor but not killing. Compatible with everything we have said, it is reasonable to hope that we don't kill.

Second, one might fall back on the idea that even if the killing/not killing distinction does not have moral significance per se, there is a distinction in the vicinity that does have moral purchase, namely the distinction between causing someone to die on the one hand, and being a mere factor in their death on the other. Now in the context of discussions of killing and letting die some have suggested that we have a basic aversion to causing death and that this underlies our moral sensibilities.[51] And one might be concerned that the above reflections in effect try to show too much. Even if "cause to die" doesn't quite line up with "kill" (as the Katz vignette discussed earlier would suggest), certain cases above are plausibly asymmetric with regard to "cause." After all, it is much more natural to say that one caused Smith's death if he dies at 1:55 p.m. from a second cannonball than if he died from one's accelerating a cannonball that had been initially released by someone else. Of course, there are those like Lewis (1973) who will, on reflection, insist on a causal symmetry between those cases.[52] But what if one opts for a view that comports with our ordinary disposition to ascribe causality in one case and not the other?

We certainly don't want to be too dismissive to those who insist that causality is of fundamental moral significance. And that is part of why our conclusions are offered in a somewhat tentative spirit. But we do think it worth pressing whether we really do feel a difference between, say, the lever and second cannonball versions of our stories, where both involve a positive intervention by the agent, and so where there is no act/omission contrast to muddy the waters. Think back to the risk versions of the cases, for example. Do we really expect good agents to take on a greater risk of Smith's dying on the grounds that this is compensated for by a lesser risk of the agent's counting as causing the death (as opposed to merely accelerating it)? It seems to us quite odd to countenance any increased risk of death on these grounds: If the lever is more likely to end in death than the second cannonball, that would seem to settle the matter, assuming the relevant symmetries in time of death. Cases like this, with a pair of positive interventions to choose between but some asymmetry in how folk would classify the outcomes using causal language, may indicate that causal asymmetries do not quite carry the moral weight that the

51. See, for example, McMahan (1993, 277–79), where he argues that we care about the "form and degree of an agent's causal responsibility."

52. See fn. 23.

(continued...)

INDEX