

Saving and Letting Live [draft: 17th May 2022]

Thomas Byrne

There is a difference between killing and letting die. In this paper, I introduce the corresponding difference between saving and letting live and I argue that it's morally important in much the same way and for much the same reasons. I then put that difference to work in moral theory. I first use it to give a general account of those cases in which agents face a choice between one person living and five people dying, or *vice versa*. I then argue for an (objective) moral theory that disregards the (mere) consequences of our actions and, in turn, is immune from cluelessness.

10000 words

1. There is a difference between person A killing person B and it merely being the case that B wouldn't have died but for what A did—*viz.* the difference between killing and letting die. There is also a difference between A saving B and it merely being the case that B would have died but for what A did. If I move my car out of the way, enabling the fire engine to reach your house in time, it might be that you would have died but for my moving my car, but I don't save you—the fireman carrying you through the flames does. That latter difference has no name: I will call it the difference between *saving* and *letting live*. We can do a lot with it.

2. I begin by considering those cases in which an agent does something that results in
one person dying and five people living,
when she could instead have done something that would have resulted in
that same one person living and those same five people dying.

I will call any such case a *five-or-one case* since in choosing between performing the first or the second action, the agent faces a choice between five people dying or one person dying. (Of course, the exact size of those numbers isn't important, but only that the bigger one is just slightly bigger than the smaller one.)

We're all familiar with one particular five-or-one case: in FOOTBRIDGE a runaway trolley is heading towards five trapped workers and, in order to save their lives, Annie pushes Ben into the trolley's path, stopping it and saving the five.¹ Here, Annie's pushing Ben results in

Ben dying and the five workers living,

but if Annie had instead stayed back, then that would have resulted in

Ben living and the five workers dying.

So, as I say, FOOTBRIDGE is a five-or-one case—and one we're familiar with.

We're familiar with others too. The case in which the transplant surgeon chops up her healthy patient in order to save the lives of five others is a five-or-one case. And the case in which the coastguard runs down a pedestrian stuck in the road so that she can make it to the beach in time to rescue five drowning swimmers is another.²

The cases just mentioned are all alike in two important respects. The first is that, in each of them, the death of the one is not merely a result of what the agent does—it's not merely that the one wouldn't have died but for what the agent did—but instead that the agent *kills* the one: Annie kills Ben by pushing him off the footbridge; the surgeon kills the healthy patient by chopping him up; and the coastguard kills the pedestrian by running him down.

The second respect in which the preceding five-or-one cases are alike mirrors the first. In each of them, the survival of the five isn't merely a result of what the agent does—it's not merely that, each time, they wouldn't have survived but for what the agent did—but instead it's that the agents each *save* the five: Annie saves the five workers by stopping the trolley; the surgeon saves the five by transplanting the organs they each need; and the coastguard saves the five swimmers by dragging them to shore. (I add that

¹ Thomson (1976), (1985).

² Both from Foot (2002).

scepticism about these claims would be inappropriate: it's obvious that Annie saves those lives. There is something in the vicinity of this scepticism that is appropriate and that is that in tricky cases we should be reluctant to adjudicate whether A saves B, but these are not tricky cases—though we shall see some, shortly.)

That the agents each kill the one and save the five in these cases makes appropriate a certain principle of why they each act impermissibly—and I take it as a datum that they do act impermissibly. The principle I have in mind goes back to Philippa Foot (2002) and it says that

other things equal, it's impermissible to kill one person in order to save the lives of five others.

Since, e.g., Annie kills Ben in order to save the five trapped workers, this principle—*Foot's principle*, as I will call it—rightly accounts for the impermissibility of Annie's doing so.³ The same goes for the surgeon, the coastguard, and other five-or-one cases in which the agent saves the five and kills the one. I will call all such cases *kill-save* (five-or-one) cases.

Not all five-or-one cases are kill-save cases and here are some examples of that.

REVENGE: the tide is rising and five clammers are trapped in a cove. Annie can make it to the cove in time to rescue them, but Villain wants them dead. Annie learns that if she rescues the five, then Villain will kill an innocent victim in retribution. Annie rescues the five, Villain kills Ben.

Here, Annie's rescuing the five results in

Ben dying and the five clammers living,

but if Annie had instead stayed back, then that would have resulted in

³ Thomson's famous case in which a bystander diverts a trolley away from five trapped workers and onto a single worker is ostensibly a counterexample to Foot's principle (1976, 1985). Note that it's a counterexample if only if (i) the bystander acts permissibly, (ii) the bystander kills the one, and (iii) the bystander saves the five. To my mind it isn't a counterexample. REDACTED argued that (ii) is false, while Thomson (2008) argued that if (ii) is true, then (i) must be false. So long as one of us is right, then the case isn't a counterexample, after all: I proceed on that basis. Others might process on the basis that the trolley problem is a fascinating, standalone puzzle.

The principle has also been objected to on more general grounds (e.g. Bennett (1998) and Kagan (1988)). For responses see Woollard (2008) and REDACTED.

Ben living and the five clammers dying.

So REVENGE is a five-or-one case, but it's not a kill-save case. It's not a kill-save case because while Annie saves the clammers, she doesn't kill Ben: Villain kills Ben, the bullet kills Ben, but Annie certainly doesn't. Since Annie saves the five but doesn't kill the one, I will say that REVENGE is a *save-only* (five-or-one) case.

Here is another.

WIND: a runaway trolley is again heading towards five trapped workers. Ben is leaning precariously over the footbridge rail observing the situation and, if he falls, he will stop the trolley before it reaches the five. Annie sees a harsh wind gusting towards Ben. She could alert him, giving him time to move to safety, but she keeps quiet. The wind knocks Ben off the bridge, stopping the trolley.

Here, Annie's keeping quiet results in

Ben dying and the five workers living,

but if Annie had instead alerted Ben, then that would have resulted in

Ben living and the five workers dying.

So WIND is a five-or-one case, but it's not a kill-save case—nor a save-only case. Instead, it's what I will call a *neither-kill-nor-save* (five-or-one) case since Annie neither kills Ben nor saves the five workers.

One further example is harder to categorise along these lines.

NAPPING: as with FOOTBRIDGE, but this time Annie is too far away to do anything herself. However, she spots a committed consequentialist napping on the footbridge. She knows that if she yells, then the consequentialist will wake up, see what's going on and push Ben off the footbridge in order to save the five workers. Annie yells.

This five-or-one case is harder to categorise because it's unclear exactly what Annie's role is here. On the one hand, the consequentialist kills Ben: does that give Annie immunity? On the other hand, Annie "unleashes" the consequentialist on Ben: if I unleash a caged tiger as you sneak past, which then eats you,

don't I kill you? And the corresponding questions arise for whether Annie saves the five. I set these matters aside, but will return to them later.

(For completeness, I note that along with kill-save cases, save-only cases, neither-kill-nor-save cases and as-yet-uncategorisable cases, there are also *kill-only* cases—ones in which agents kill the one, but don't save the five. We'll see such cases later.)

Foot's principle—that it's impermissible to kill one in order to save five others—has little to say about the preceding three cases. It has little to say because they're not kill-save cases and, therefore, they fall outside its scope: that it's impermissible to kill one to save five is irrelevant to a case in which the agent either doesn't kill one or doesn't save five. What we want is a theory that covers all five-or-one cases, not just the kill-save cases. Foot's principle might be part of such a theory, but it isn't the whole theory.

What verdicts would we want that theory to return? I said that I took it to be a datum that it was impermissible for Annie to push Ben in FOOTBRIDGE, but what about in the three cases just introduced—does Annie act impermissibly there? To my mind, Annie acts permissibly in REVENGE. When I put myself in her shoes and consider what I would do, I conclude that I would recuse the clammers—in full knowledge and with a clear conscious that my doing so will result in Villain's vengefully killing Ben.

On the other hand, I think Annie acts impermissibly in both WIND and NAPPING. When I put myself in her shoes in those cases, I conclude that I would alert Ben to his peril and that I wouldn't yell to awaken the consequentialist—again, in full knowledge and with a clear conscience of what will result.

Of course, that I would do such-and-such when in Annie's shoes does not entail that it's impermissible for her to do otherwise, but it's a starting point. I will return to these verdicts later—beginning with WIND in the next section.

So what we want is a theory that covers all five-or-one cases. We've already seen the makings of such a theory and the central idea is simple enough. Above, I introduced two distinctions. The first was between killing and letting die. Many people have thought that that distinction is morally important: that it makes a moral difference whether A kills B or whether B merely dies as a result of something A does

(cf. Foot's principle). The second distinction was a distinction between saving and letting live. People have not thought that distinction morally important. But it is morally important—in the same way, and for the same reasons as the first distinction. We can do a lot with those *two* distinctions: starting with a theory of five-or-one cases (§§3-6), and ending with a moral theory that disregards the (mere) consequences of our actions (§§7-9).

3. Recall

WIND: a runaway trolley is again heading towards five trapped workers. Ben is leaning precariously over the footbridge fence observing the situation and, if he falls, he will stop the trolley before it reaches the five. Annie sees a harsh wind gusting towards Ben. She could alert him, giving him time to move to safety, but she keeps quiet. The wind knocks Ben off the bridge, stopping the trolley.

Just now, I said that Annie acts impermissibly in this case, and everyone I've asked about it says the same, but they've also found that verdict puzzling. They scratch their heads and say "but Annie doesn't kill Ben, she just keeps quiet, and the result of her doing so is that five people live instead of one person; and given all that, why should it be impermissible for her to do so?" Where does that reasoning go wrong?

It goes wrong by omitting two important facts about the case. The first is that:

- (a) if Annie had alerted Ben, it's not merely that Ben would have lived, but that Annie would have saved Ben's life.

And that fact contrasts with another, similarly omitted (although, mentioned above):

- (b) in keeping quiet, Annie didn't save the lives of the five, but instead it's merely the case that they would have died but for her keeping quiet.

That phrasing—"it merely being the case that B would have died but for what A did"—is cumbersome and it will be useful to have something that better rolls off the tongue. As I said above, I have adopted *letting live*, where

A lets B live =_{df} B's not dying is a result of something A does, yet A doesn't save B's life.

("Not dying" rather than "living," lest reproduction be counted.) For example, the five's not dying in WIND is a result of Annie's keeping quiet, yet Annie doesn't save the five: thus she lets them live.

The distinction brought out by (a) and (b) is therefore the distinction between saving and letting live. Does it make a moral difference?

If it does, then it would be of a familiar sort. Many people have thought that the distinction between killing and letting die makes a moral difference. In particular, they think that killing someone is sufficiently worse than letting someone die such that it's impermissible to kill one person even when the alternative is letting five people die (cf. Foot's principle). If it makes a moral difference, when B dies, whether A kills B or A lets B die, we shouldn't be surprised at the thought that, when B doesn't die, it also makes a difference whether A saves B or A lets B live.

Why shouldn't we be surprised? In both cases, there is some event: a bad event in the former (B's dying) and a good event in the latter (B's not dying). And, in both cases, the relevant distinction concerns the metaphysical relationship that A stands in to that event. The moral importance of the distinction between killing and letting die holds that it's worse when A stands in that particularly *direct*—for lack of a better term—relationship to B's death that one stands in when and only when one kills B; and the (supposed) moral importance of the distinction between saving and letting live holds that it's better when A stands in that particularly direct relationship to B's not dying that one stands in when and only when one saves B.⁴

We can imagine what that moral difference would look like, too. It would hold that saving someone is sufficiently better than letting someone live such that,

other things equal, it's impermissible to fail to save someone's life in order to let five others live.

Since to save someone is to give them a helping-hand, I'll call this *Hand's principle*.

⁴ REDACTED.

We can also see how that principle would neatly account for neither-kill-nor-save cases, like WIND—although that it does so is no surprise. Given (a) and (b), it follows immediately from Hand's principle that it was impermissible for Annie to keep quiet: she failed to save Ben's life (even though, in doing so, she let the five workers live).⁵

Some readers will bristle at all this: “why would it make a difference whether Annie saves someone's life or merely lets them live? Surely all that matters, at the end of the day, is how many people are alive and how many are dead.” Such readers might substitute “kill” for “save,” and “let die” for “let live,” and ask themselves the question that results: *viz.* “why would it make a difference whether Annie kills someone or merely lets them die? Surely all that matters is how many people are alive and how many are dead.” To my mind, the answer to that second question goes little deeper than “because it makes a difference whether *I* kill him,” but, in any case, its answer will stand or fall with the corresponding answer to the first question.⁶

Now, we might be able to imagine what the moral difference between saving and letting live would look like; and we shouldn't bristle at the thought of such a difference, but...is there such a difference? That it neatly accounts for neither-kill-nor-save cases is some reason to think there is. And later I'll argue that the conjunction of Hand's principle and Foot's principle gives us a neat theory of all five-or-one cases: that's more reason to think there is a moral difference between saving and letting live. However, there's another reason to think there is such a moral difference that has nothing to do with five-or-one cases and, instead, has been staring at us for ages. Let's turn to it.

⁵ Hand's principle also solves a puzzle presented by Segev (2016). Segev asks why we are required to prevent, e.g., a consequentialist from pushing Ben in FOOTBRIDGE (or, more generally, why we are required to prevent “deontological wrongdoing”?). Because, if we don't prevent the consequentialist, then we fail to save Ben's life in order to let the five workers live. See also Mogensen (2016) and Sinclair (2017).

Hand's principle also fits nicely with cases introduced by Hanser (1995). Hanser asks why, e.g., A isn't required to kill a shark that is about to eat B when, if killed, a different shark will eat B instead (which A won't be able to stop)? Because A wouldn't save B's life in doing so.

⁶ “But people have a right to not be killed but no right to be saved.” Yes, but it's not merely that some right is being violated, but that *I* would be violating that right that matters (otherwise we could simply “consequentialise” on rights violations—which we can't).

Peter Singer (1972) introduced us to two cases: in the first, Annie fails to save a child drowning in a pond because it would ruin her fancy outfit; in the second, Annie fails to donate the cost of that same outfit to Oxfam, despite knowing that with every such donation the charity is able, in the aggregate, to prevent a child's death from malaria. Singer then argued from the fact that Annie acts terribly in the first case (as surely she does), via the premise that there is no moral difference between the two cases, to the surprising conclusion that Annie also acts terribly in the second case. Ever since, undergraduates have vainly sought to respond to Singer's argument by pointing to some ostensible moral difference between the two cases; and lecturers have, on Singer's behalf, adjusted one of the two cases to show that the difference in question doesn't make a moral difference.

Singer's own example of this back-and-forth concerns the fact that, in the pond case, Annie is uniquely positioned to save the child, whereas there are millions of others who are equally able to donate, but decide not to—that is the ostensible moral difference between the two cases. Singer then adjusts the first case so that many other people are also able to save the drowning child, but don't. Since, in that case, it would still be terrible for Annie to refuse to rescue the child, Singer concludes that that uniqueness can't make a moral difference, after all.

Other ostensible moral differences include: that the drowned child is close to Annie, yet Oxfam's beneficiaries are far away; that the drowned child wouldn't have drowned but for Annie's refusal, yet there is no determinate child who wouldn't have died had Annie given to Oxfam; that there is only a single drowning child, but many children dying of malaria; and so on. I leave it to the reader to adjust the relevant cases in order to generate Singer's responses to these other differences—I suspect it's all very familiar.

Here is a simple explanation of the moral difference between Singer's cases: there is a moral difference between saving and letting live, and only in the first case does Annie fail to save a child's life—you save those children you drag from ponds, but not those that are the ultimate beneficiaries of your few-thousand dollar charitable donation. (In conversation, some people have questioned that latter claim, which puzzles

me because it strikes me as plain: I donated over \$4000 to the Against Malaria Foundation last year, sufficient for them to prevent a death in the aggregate, but I'm certain that I never saved anyone's life. Just as, for example, a rich tax payer might have contributed enough tax to the treasury to end a life, in the aggregate, in the Iraq War, but that tax payer never killed anyone. However, I do stress that I picked Oxfam and Against Malaria Foundation advisedly, due to the size and complexity of their operations; but charities needn't be like that, as we'll soon see.)

Testing this explanation is simple enough. We need only consider variations of Singer's two cases and see whether, for each of them,

Annie acts terribly by refusing to ϕ iff by ϕ -ing, Annie would have saved a child's life.

Simple, but tedious. Instead of considering variation after variation, we might instead consider a single convoluted case that incorporates all of the ostensible moral differences mentioned, above:

A number of children are drowning in a Russian lake. An automated lifeguard system protects the lake, but is currently deactivated. Once activated, it will rescue the first child that passes under its sensor, before deactivating permanently. The lake is very choppy and the children are chaotically bobbing around in the vicinity of that sensor. The system's activation switch lies in the middle of a pond in Boston. Hundreds of people surround that pond, Annie included. Annie could swim into the pond and activate the switch, but she doesn't for fear of ruining her outfit. Nor does anyone else.

Here, Annie isn't uniquely positioned to act (given the many others by her pond); nor are the drowning children close to Annie (they're in Russia); nor is there any determinate drowned child who wouldn't have died had Annie acted differently (because, given the choppy weather, there is no fact as to which child would have first passed under the sensor after Annie had activated it); nor is there only a single drowning child.

Despite all that noise, the following is clear: if Annie had activated the switch, then she would have saved a child's life. I trust that it's also clear that Annie acts terribly by refusing to activate that switch. I suggest that we can continue to convolute the details, but so long as it remains the case that Annie fails to save a child's life, then it will remain the case that she acts terribly—and *vice versa*. I add that, on its face, that sounds rather sensible.

That was a variation of Singer's pond case—let's consider a variation of the charity case. I assumed, above, that the charity in the original case operates like, e.g., Against Malaria Foundation and therefore that it's not the case that Annie would have saved a child's life by donating that few-thousand dollars. It's instructive to imagine a donation in virtue of which the donor does (or would) save a child's life. However, to do so, we must travel a long way from Against Malaria Foundation—passing cases in which it remains clear that the donor wouldn't save a child and others in which it's increasingly unclear either way. In any case, we eventually arrive at something like the following:

A child is drowning in a distant lake. She can only be reached in time via helicopter, but Pilot has run out of fuel and money. Annie is the only other person around and Pilot asks her for the requisite funds. Annie refuses since she wants to buy herself a new outfit. Pilot doesn't make it to the lake and the child drowns.

It's a daft case, but Annie surely acts terribly in refusing to donate the money. Why is that? Because, in doing so, she failed to save the child's life. (I mention now that that latter fact will prove useful when we return to the as-yet-uncategorisable metaphysics of cases like NAPPING.)

This diversion has gone on long enough and I won't consider any more variations of Singer's cases. What we wanted to know is whether there is a moral difference between saving and letting live. To my mind, Singer's cases wonderfully illustrate that there is, but I won't press the point. Let's instead turn to what that moral difference would give us, since those fruits are more reason to think it exists.

4. When B's death is a result of something A does, there are different relationships that A might stand in to that death. The difference between killing and letting die picks out two such relationships, but we now know that that's not the end of the story: when A lets B die, it might also be the case that A failed to save B's life. We can order these different relationships in terms of directness (again, for lack of a better term) —from most direct to least:

A kills B,

A fails to save B (and lets B die),

A lets B die (and it's not the case that A failed to save B).

For simplicity, I will omit the parentheticals and thus refer to that middle entry as "A fails to save B," and refer to that bottom entry as "A lets B die."⁷

Foot's principle says that it's impermissible for A to kill one person when A could instead have stood in a less direct relationship (*viz.* failing to save) to five deaths. Hand's principle says that it's impermissible for A to fail to save one life, when A could instead have stood in a less direct relationship (*viz.* letting die) to five deaths. There's a pattern here which we might generalise as follows:

(*) other things equal, A acts impermissibly in a five-or-one case iff the relationship that A stands in to the death of the one is more direct than the relationship A would have stood in to the deaths of the five had she acted otherwise.

We now see that Foot's and Hand's principles are just specific instances of the more general (*).

Those specific instances of (*) account for why Annie acts impermissibly in FOOTBRIDGE and WIND, but (*) also accounts for why Annie acts permissibly in REVENGE (recall: Annie saves the five clammers and Villain randomly kills Ben in revenge for her doing so). It does so because the relationship that Annie stands in to Ben's death (neither killing him nor failing to save him) is *less* direct than the relationship that

⁷ Hanser (1999) also introduces another relation—*preventing someone from being saved*—which superficially looks similar to what I'm doing here. He's interested in cases where, e.g., A removes the barrier that would otherwise stop B and B's car from toppling off the cliff. To my mind, those are plainly killings. His argument otherwise relies upon cases structurally akin to Thomson's famous bystander case and the metaphysical confusion such structure elicits. See REDACTED.

Annie would have stood in to the deaths of the five clammers had she stayed back (she would have failed to save them), and thus (*) rightly returns that Annie acts permissibly.

REVENGE is what I called a save-only case, because while Annie saves the five clammers, she doesn't kill Ben. Its converse—a *kill-only* case—would be one in which Annie kills Ben, but doesn't save the five: perhaps Annie kills Ben knowing his death will motivate Hero to save five others. (*) says that Annie acts impermissibly since the relation she stands in to Ben's death (killing him) is *more* direct than the relationship she would have stood into the death of the five had she done otherwise (letting them die). That is surely correct.

For completeness, I'll mention that (*) also returns the right verdict in cases where the agent has a choice between saving one or saving five (or between killing one or killing five). If A saves the five, then the relationship that A stands in to the death of the one (failing to save) is *equally* direct as the relationship she would stand in to the deaths of the five (failing to save) had she instead saved the one; and therefore (*) says it's permissible for A to save the many (and, similarly, to kill the few).

That's all rather tidy and simple.

It's worth stressing just how tidy and simple. There are three different metaphysical relationships that one might stand in to a death—from most direct to least direct: killing, failing to save, and letting die. And whether the agent acts permissibly in a given five-or-one case is fixed by whether the relationship she stands in to the death of the one is more direct than the relationship she would have stood in to the deaths of the five. That's it.

(I add that while I have connected killing, failing to save, and letting die in terms of metaphysical “directness,” I needn't have done so. I might instead have said that killing is worse than failing to save, failing to save is worse than letting die, and killing is worse than letting die—and have left them as three independent claims. Readers who are sceptical of this directness might proceed as if I had. To my mind, it's intuitive that when A kills B, A stands in a more direct relationship to B's death than if, e.g., A had merely failed to save B—but I won't press that point.⁸)

⁸ REDACTED.

5. I just said that, when B dies, there are three morally important metaphysical relationships that A might stand in to his death. Readers might have expected me to add a fourth—namely, that A caused B to die—but it doesn't belong. It doesn't belong because it draws distinctions in morally irrelevant places, regardless of which view of causation one holds.⁹ Some people think that causation is largely a matter of counterfactual dependence such that E causes C iff C wouldn't have occurred but for E: if they're right then causation doesn't distinguish between five-or-one cases because there is always such dependence, by definition.¹⁰ On the other hand, some people think that causation is narrower and think, for example, that while Annie does cause Ben's death in REVENGE (by causing Villain to kill him), Annie doesn't cause Ben's death in WIND (when she keeps quiet): if they're right, then A's causing B's death cuts across the permissibility facts since Annie acts permissibly in REVENGE but impermissibly in WIND. So, as I say, regardless of how we think about causation, it draws distinctions in morally irrelevant places and that's why it's omitted, above.

6. Recall

NAPPING: as with FOOTBRIDGE, but this time Annie is too far away to do anything herself. However, she spots a committed consequentialist napping on the footbridge. She knows that if she yells, then the consequentialist will wake up, see what's going on and push Ben off the footbridge in order to save the five workers. Annie yells.

I said that Annie acts impermissibly here, but I also said that it wasn't clear whether Annie killed Ben, nor whether Annie saved the five. I remarked:

On the one hand, the consequentialist kills Ben: does that give Annie immunity? On the other hand, Annie “unleashes” the consequentialist on Ben: if I unleash a caged

⁹ This point has been made more generally, elsewhere: e.g. Carolina Sartorio (2010) and REDACTED.

¹⁰ E.g. David Lewis (1973).

tiger as you sneak past, which then eats you, don't I kill you? And the same questions arise for whether Annie saves the five.

With (*) in hand, progress can be made. After all, if Annie acts impermissibly then, by (*)'s lights, it must be the case that Annie's relationship to Ben's death is more direct than the relationship that she would have stood in to the deaths of the five had she kept quiet (*viz.* letting them die). Since Annie certainly doesn't fail to save Ben's life, she must therefore kill Ben. Can that be right? I've come to think that it is right.

To see why, let's return to the caged tiger, just mentioned, and suppose that Annie opens its cage while Ben is sneaking past. In which case, it's not only that the tiger kills Ben, but also that Annie kills him—they both kill him. That preceding clause—"they both kill him"—seems unremarkable, but its meaning actually isn't transparent. When we say, for example, of two playground bullies that they both hit the victim, what we mean is that the victim was hit twice-over, once by the first bully and once by the second, but we don't mean anything like that when we say that Annie and the tiger both killed Ben: Ben certainly wasn't killed twice-over, once by Annie and once by the tiger. So what do we mean?

Here's a grizzly legal case to consider alongside the killing case and against the hitting case. In *Dudley and Stephens*, Dudley held Parker down while Stephens slit his throat.¹¹ The courts found that Dudley and Stephens murdered Parker, but of course Parker wasn't murdered twice over. I trust there is no mystery here: Dudley and Stephens murdered Parker *together* (even though it was Stephens wielding the knife). And such cases abound: when the bricklayers build a wall, it's not that the wall is built many times over, but that the bricklayers build the wall together (even though there might be a specific bricklayer who lays the final brick); or when the villagers fish the pond to depletion, it's not that the fishpond is depleted many times over, but that they deplete the fishpond together (even though one specific villager might have

¹¹ (1884) 14 QBD 273 DC. All three were lost at sea in a lifeboat, starving, and Parker was closest to death's door.

caught the final fish); or when the relay team win the race, they win it together (even though only one runner crosses the line first).¹²

And we saw another case like this in §3, when Pilot needed money from Annie in order to reach the drowning child. If Annie had donated the money to Pilot, then both she and Pilot would have saved the child's life, but the child wouldn't have been saved twice over—instead, Annie and Pilot would have saved the child together. (I add that that's the case even though Pilot's role was more proximate and "bigger" than Annie's; just as it would remain the case that Pilot saved the child even if there was, e.g., a third person operating the winch that lifted the child to safety.)

So what do we mean when we say that Annie killed Ben and the tiger killed Ben? We mean that they killed him together—and kill him together they did. (And that's the case even though the tiger's role was more proximate and "bigger" than Annie's.) It's in that same vein that Annie killed Ben in NAPPING: she and the consequentialist killed Ben together.

Other cases are like this, too. Suppose that in a variation of FOOTBRIDGE, Annie pays her underling to push Ben; the thoughts of the preceding paragraphs return that Annie and her underling killed Ben together. (And that fits with the moral facts: there isn't any moral difference between Annie's pushing Ben and Annie's paying her underling to push Ben and that Annie does the same thing in both cases—*viz.* killing Ben—fits with that.)

Now, there are deep questions here concerning the limits by which A and B can kill C together (or, similarly, the limits by which they can save C together). It's plain that Dudley and Stephens killed Parker; and a little reflection reveals that the same goes for Annie and the tiger (they killed Ben together); and still more reflection, combined with the conclusions of the previous section, reveals (I claim) that the same goes for Annie and the consequentialist in NAPPING. But good sense says it doesn't also go for, e.g., the consequentialist and the consequentialist's mother (they don't kill Ben together in NAPPING—at least, they don't in virtue of her giving birth to him). Why is that? Similarly, why is that Annie and Pilot save

¹² Certain theories of joint action (e.g. Michael Bratman 1992) think that the intentions of the group are central to whether the groups acts together. Since it strikes me as plain that, e.g., the villagers deplete the fishpond even when they're acting independently, I set such theories aside. Other theories are in agreement (e.g. Sara Rachel Chant (2006)).

the child, but Annie doesn't save anyone when she donates to Oxfam?¹³ Similarly, just how far through a chain of command can one kill? As I say, these are deep questions and they're not easily answered.

But nor are they questions that need to be answered here. What we wanted was a theory of five-or-one cases and that we now have. Yes, the precise extension of the verdicts returned by that theory will, in part, be fixed by the answers to those questions, but that is a job for another day.

7. There's a moral difference between saving and letting live, and the moral status of saving someone's life is sufficiently great that it's impermissible to fail to do so even when the alternative is letting five others live (where, recall, A lets B live iff B's not dying is a result of something A does, yet A doesn't save B's life). That difference gave us the theory of five-or-one cases introduced in the preceding sections. Now I turn to a different question: what is the moral status of letting someone live?

Asking around, everyone says (assumes?) that letting live has some moral status. Surely, they say, if A's ϕ -ing will let B live then, other things equal, A should ϕ and surely, they continue, it's the moral status of letting live that explains why that's the case. Perhaps that's right, but it comes at a high price: we're almost always *clueless* as to whether it's permissible to act.

This cluelessness is a well-worn problem for consequentialism, but if letting live has some moral status and if that moral status aggregates (as presumably it would) such that there is a number n for which it's permissible to kill one person in order to let n people live, then we're similarly clueless as to whether it's permissible, e.g., for Annie to push Ben in FOOTBRIDGE. (I present this as a problem for Foot's principle, but the same would go for Hand's principle and others like it, too.)

We're clueless because we're clueless as to whether Annie will let n people live by pushing Ben. To adapt James Lenman's canonical example, suppose that it's 100 BC and we're on the banks of the Rhine: a boulder is heading towards five trapped villagers and Agatha can save their lives only by pushing Burt

¹³ "Why doesn't a donation to Oxfam *unleash* its aid in the same way that opening a cage unleashes a tiger—such that you and Oxfam save the child together?" Because that's plainly not how Oxfam and its peers work. Instead your donation goes into a pot, that generates a budget for next year, which goes through committees as to its allocation, etc. Compare that with the Pilot who Annie does "unleash."

into the boulder's path. Agatha doesn't push him, but would it have been permissible for her to do so?¹⁴ With 2100 years of hindsight, we now know that Burt was a many-times-great grandfather of Hitler. Thus, if Agatha had pushed Burt, then there would have been no Hitler and, in turn, none of his victims would have died by his hand: Agatha would have let millions live (surely a number far greater than n). But even with that hindsight we still don't know whether it would have been permissible for Agatha to push Burt: after all, if Agatha had saved the five villagers, then perhaps one of their descendants would have been worst still than Hitler (e.g. Malcolm the *Truly Terrible*)—we just don't know. And what goes for Agatha and Burt goes for Annie and Ben: just as we're clueless now as to whether it would have been permissible for Agatha to push Burt, so too are we clueless whether it's permissible for Annie to push Ben (by the lights of Foot's principle).¹⁵

At least, we are if letting someone live has some moral status. And it's easy to see why that moral status is the culprit: letting live is metaphysically cheap; so cheap that we can do it over any distance—both geographical and temporal—far beyond our ken. It's that cheapness which results in it being the case that the vast majority of those that we let live exist after we die, which, in turn, results in this cluelessness.

In what remains, I push for the alternative: there is no moral status whatsoever to (merely) letting someone live; when it comes to life and death all that matters is whether A kills B or A saves B (or fails to save B). The result is a moral theory that disregards (mere) consequences.

8. I start by considering some objections since, in conversation, that is where people immediately head. In the next section, I sketch a positive picture.

Objection 1: "*If A's ϕ -ing will let B live then, other things equal, it's impermissible for A not to ϕ ; and that's because letting live has some moral status.*" Put schematically, this objection is compelling, but the challenge is finding particular cases that support it. A first attempt: a trolley is heading towards B and A

¹⁴ Lenman (2000).

¹⁵ There's an argument via chaos that there is nothing for us to be clueless about since it's *indeterminate* how things would have been this century had Agatha pushed Burt (e.g. Hare 2011). However, it's being indeterminate whether it's permissible for Annie to push Ben is just as pricey as us being clueless as to whether it's permissible. Regardless, what I say about cluelessness would also go for indeterminacy.

can divert it down an empty track. Now, it's impermissible for A to fail to divert the trolley, but that doesn't support the objection since that impermissibility is explained by A's failing to save B (as certainly she would). A second attempt: a trolley is heading towards B, but an unbalanced boulder is about to tumble into the trolley's path; A stabilises the boulder, stopping its tumble. As with the first attempt, A acts impermissibly here, but this similarly fails to support the objection since that impermissibility is explained by A's killing B (if A removes the barrier that protects B from X, then A kills B).¹⁶ My suspicion is that every such attempt to find a particular case which supports this objection will fail like the previous two: it's impermissibility will be explained by either A's killing B or A's failing to save B.

Objection 2: *“Consider a variation of WIND in which there are hundreds (or thousands...or millions...) of people on the track—all of whom will be killed by the runaway trolley. Surely Annie shouldn't alert Ben in that case and surely that's because letting someone live has some moral status and, when the numbers are big enough—as they are here— it's impermissible to save the one instead of letting the many live.”*

I agree that it's impermissible for Annie to alert Ben in such a case, but I disagree that that impermissibility is explained by the moral status of letting live.

Consider another variation of WIND in which the person on the footbridge is a Villain who has trapped a lone worker onto the tracks, below (and, as before, the wind is gusting towards Villain and will blow him into the trolley's path unless Annie alerts him). It's presumably impermissible for Annie to alert Villain in this case— I certainly would keep quiet in her shoes. However, that impermissibility cannot be explained by the moral status of letting the worker live since even if letting one person live has some moral status, it certainly doesn't have enough to trump saving one person's life.

What can explain that impermissibility? It's the fact that Villain is required to save the worker—and I take it as a datum that he is so required. And that explanation makes sense: morality speaks in a consistent voice, but it wouldn't if it required Villain to save the worker, yet permitted Annie to prevent him from

¹⁶ That Annie would kill here contrasts in a puzzling way with WIND's counterfactual in which Annie does alert Ben to his peril. I discuss this puzzle in an addendum.

doing so.¹⁷ More generally, it wouldn't speak in a consistent voice if it required B to ϕ , yet permitted A to prevent B from ϕ -ing. Thus, the fact that B is required to ϕ entails that A isn't permitted to prevent B from ϕ -ing.

When we return to the case with Ben on the footbridge and hundreds (or thousands...or millions) on the tracks, the question is whether Ben is required to save the many. Here's a simple argument that he is required. There is some number n such that it's impermissible for Annie to alert Ben when n people are on the tracks (and, by stipulation, n people are on the tracks). Presumably, there is also some number m , such that, if m people are on the tracks, then Ben is required to save them—even if it costs him his life. Is it the case that $n > m$, that $m > n$, or that $m = n$? The thoughts of the preceding paragraph rule out $n > m$, since that would permit cases in which Ben is required to save those on the tracks, yet Annie is permitted to prevent him from doing so. A related thought rules out $m > n$ since that would permit cases in which Ben isn't required to give his life for those on the tracks, declines to do so—as he is at liberty to do—yet it's impermissible for Annie to save him, regardless: Annie is required to let Ben give his life, despite Ben not being required to give it. That won't do: morality's voice is perverse no more than it is inconsistent.

That leaves only $m = n$: the number of people on the tracks that makes it impermissible for Ben not to save them is the same number that makes it impermissible for Annie to prevent him from doing so. That strikes me as sensible: morality should hang together in that way.¹⁸ And that morality hangs together in that way responds to this second objection: it's being impermissible for Annie to alert Ben entails that Ben is required to save the many trapped workers and it's that latter requirement that explains the former impermissibility—and not anything to do with the moral status of letting live.

Objection 3: *“But if letting live has no moral status, and we don't save those that are the ultimate beneficiaries of our donations to, e.g., Oxfam, then we have no obligation to donate to, e.g., Oxfam.”* This is a strange objection since it relies upon the intuition that we have an obligation to donate to Oxfam,

¹⁷ In one sense, Annie isn't preventing Villain since he is still able to jump himself. It's the other sense that's important here: given he won't, in fact, jump, he will only save the worker if the wind pushes him off the footbridge.

¹⁸ Thomson (2008) highlighted a related way in which morality should hang together. She said that it can't be that A is permitted to force B to pay a given cost unless B is required to pay that cost (it can't, e.g., be that A is permitted to push B off the footbridge unless B is required to jump).

despite there being no such intuition. After all, we take it to be permissible for Annie to not donate the cost of her outfit to Oxfam. Singer (and others) have argued otherwise, but we now know that that argument relies upon letting live having some moral status (indeed, having the same moral status as saving). So relying upon Singer's conclusion to mount this objection begs the question.

Objection 4: *“Earlier, you said that it was in-principle possible for us to save lives via a donation to charity (§3's helicopter pilot case was of that sort). If we grant that we're not obligated to give to charities like Oxfam, this returns that whether we're obligated to give to a particular charity is fixed by the charity's causal structure. What a thing for morality to care about!”* That ‘thing’ is whether I can save a life and that makes a difference. Suppose that Annie is on one side of an uncrossable river and Ben is starving to death on the other. As it stands, Annie isn't obligated to save Ben's life (after all, it's not possible for her to do so), but if I transport Ben to Annie's doorstep, then Annie becomes obligated to save Ben's life; similarly, if I transport Annie and her pantry to Ben's location; similarly, if I open a ferry route across the river. When whom we are able to save changes, our obligations change in turn; some charities bring about such changes, some don't.

Objection 5: *“If Annie can either ϕ , the result of which will be that one person doesn't die, or Annie can Ψ , the result of which will be that five people don't die, can it really be right that we need to know about the causal structures of ϕ -ing and Ψ -ing before we know what Annie ought to do?”*

It certainly can. Suppose a child is drowning in a lake and Annie can either donate the \$20,000 strapped to her chest to Oxfam (which will result in five deaths being prevented, in the aggregate) or Annie can save the drowning child (which will ruin the money since there's no time to remove it, first). It's those causal structures that explain why Annie is required to save the drowning child.

9. What does a moral theory that disregards (mere) consequences look like? When it comes to life and death, it says that what matters is whom we kill and whom we save (and, by extension, whom we fail to save and, occasionally, whom we fail to kill). The result is that our sphere or moral concern is limited. But

it's not limited temporally nor geographically since we can, with some effort, kill and save at great distance: if, e.g., I plant a bomb with a thousand year fuse, then I kill whomever it kills in 3022. Nor is that sphere limited arbitrarily. Instead our sphere of moral concern is limited metaphysically, by what we (can) *do*. Saving is something we do; so too is killing; but letting live isn't—it's not something we do, but instead it's a mere consequence of something else that we do.

Of course, some moral theorists say that that's a distinction without a difference. They insist that what fixes whether it's permissible for A to ϕ are the consequences of her ϕ -ing. The view proposed here says that that is a profound mistake: what fixes whether it's permissible for A to ϕ is *what A will do if she ϕ s*.

This paper has focused on cases where only lives are at stake, but that credo neatly generalises. When, for example, hunger is at stake, what matters is whom you feed and whom you starve (and not whom you merely let starve or be fed); when arms are at stake, what matters is which arms you break and which you heal (and not which arms you will let break or heal); and, more broadly, when it comes to aid, what matters is whom you help and whom you fail to help.

This theory shifts the tonnage of moral theorising towards metaphysics: since it matters whom we kill (or save...or starve... etc.,) we need to know when we kill (and save and starve...) and, in turn, exactly what it is to kill (and save and starve...). That metaphysical project is an important project for moral theory—and an exciting one.

This theory also shifts our moral obligations (or what we take them to be)—particularly our obligations towards strangers. It isn't the case, as certain people would have us believe, that our obligations to strangers are fixed by the total amount of happiness we can bring about (in the loosest sense of the phrase). Instead, they are fixed by whom, in the strictest sense of the word, we are able to help.

ADDENDUM

Consider the following case schema:

(S) A runaway trolley is heading towards five trapped workers. If Annie stays back, then X will fall into the trolley's path, stopping it from reaching the five. Annie instead intervenes and prevents X from falling into the trolley's path. The trolley kills the five.

I briefly considered one version of this schema under Objection 1, in §8. In that case, X is an unbalanced boulder that's about to fall into the trolley's path. I said of that case that, by stabilising the boulder, Annie kills the five trapped workers. Another instance of this schema is WIND's counterfactual in which Annie does alert Ben to his peril, saving his life—in that instance, X is Ben. I never explicitly said whether Annie kills the five in that case, but since it's a direct implication of (*) that she doesn't kill them, it'd better be that she doesn't.

This is puzzling: whether Annie kills the five in a given instance of (S) is seemingly sensitive to whether X is a boulder or X is Ben. Can that be right?

I think it is right and that it's just one instance of a more general sensitivity. Elsewhere, I introduced the following pair of cases:

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

As before, but this time it's not Agatha on the tracks, but a boulder. Agatha removes the boulder and the trolley instead hits and kills a worker who is tied to the tracks behind it.¹⁹

I introduced the pair as part of a demonstration of our expertise at determining when A kills B, even across pairs of cases which causation itself doesn't distinguish between (as it seemingly doesn't here). This pair was well-suited to that purpose since it's plain—I take it—that Agatha doesn't kill the worker in the first case, but does kill him in the second case.

¹⁹ REDACTED. Boorse & Sorensen (1988) discuss various cases structurally akin to the first member of this pair—one's in which the agent "ducks a harm."

I mention the pair here because it exhibits the same sensitivity as exhibited by (S): in its former member, the relevant X is Agatha herself (and Agatha doesn't kill), in its latter, the relevant X is a boulder (and Agatha does kill).

One response to this sensitivity is to insist that whether A's ϕ -ing kills B must, in part, be fixed by the permissibility of her ϕ -ing. Friends of this response would say that it's permissible for Agatha to jump off the tracks, but impermissible for her to remove the boulder, and it's that difference which accounts for the killing facts. Having thus proved that killing itself is a moral notion, these friends ring the death knell for (*). I have no sympathy for this line of thought since it's plain that whether A kills B isn't sensitive to moral facts: Agatha might have removed the boulder to save the lives of millions (and have done so permissibly), but that wouldn't change that fact that she killed the worker tied to the tracks behind it.²⁰

The correct response is to accept that whether A kills B just is sensitive in the way brought out by the cases, above. (I add that, on reflection, this isn't something that should surprise us: no one ever thought, e.g., that the bodyguard, who tackles his VIP to save her from the assassin's bullet, kills whomever is standing behind.)

In what way is it sensitive? Roughly, it seems to be that whether A kills B in a particular case is sensitive to whether what directly kills B would have killed someone else had A not done what they did. For example, when Agatha moves herself out of the way such that the trolley kills the worker, it is the case that the trolley would have killed someone (*viz.* Agatha) had Agatha not done what she did; but when Agatha moves the boulder out of the way, it's not the case that the trolley would have killed someone regardless.

I stress, however, that that's rough. How to make it more precise is, I think, another deep question—to go alongside those raised in §6.

REFERENCES

²⁰ REDACTED.

Jonathan Bennett, *The Act Itself*, (Oxford University Press, 1998).

Christopher Boorse and Roy A. Sorensen, "Ducking Harm," *Journal of Philosophy* 85 (1988): 115–34.

Michael E. Bratman, "Shared Cooperative Activity," *The Philosophical Review* 101 (1992): 327–41.

REDACTED

Matthew Hanser, "Why are Killing and Letting Die Wrong?," *Philosophy & Public Affairs* 24 (1995): 175–201.

Matthew Hanser, "Killing, Letting Die and Preventing People From Being Saved," *Utilitas* 11 (1999): 277.

Caspar Hare, "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 (2011): 190–206.

Shelly Kagan, *The Limits of Morality*, (Oxford University Press, 1989).

James Lenman, "Consequentialism and Cluelessness," *Philosophy & Public Affairs* 29 (2000): 342–70.

David Lewis, *Causation*, (Oxford University Press, 1987).

Andreas L. Mogensen, "Should We Prevent Optimific Wrongs?," *Utilitas* 28 (2016): 215–26.

Sara Rachel Chant, "The Special Composition Question in Action," *Pacific Philosophical Quarterly* 87 (2006): 422–41.

Carolina Sartorio, "Causation and Ethics," in *Oxford Handbook of Causation*, eds. Helen Beebe, Christopher Hitchcock and Peter Menzies (Oxford University Press, 2010).

Re'em Segev, "Should we prevent deontological wrongdoing?," *Philosophical Studies* 173 (2016): 2049–68.

Thomas Sinclair, "Preventing Optimific Wrongs," *Utilitas* 29 (2017): 453–73.

Peter Singer, "Famine, affluence, and morality," *Philosophy & Public Affairs*, 1972: 229–43.

Judith Jarvis Thomson, "KILLING, LETTING DIE, AND THE TROLLEY PROBLEM," *The Monist* 59 (1976): 204–17.

Judith Jarvis Thomson, "The Trolley Problem," *Yale Law Journal* 94 (1985): 1395–415.

Judith Jarvis Thomson, "Turning the Trolley," *Philosophy & Public Affairs* 36 (2008): 359–74.

Fiona Woollard, "Doing and Allowing, Threats and Sequences," *Pacific Philosophical Quarterly* 89 (2008): 261–77.