Saving and Letting Live

(draft. under review.)

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. When person B wouldn't have died had person A not acted as she did—when B's death is a *consequence* of A's actions—in some cases A kills B and in other cases she doesn't. If I run you down with my car, then I kill you; if I ignore your drowning cries, I don't kill you, I merely let you die. That is a very familiar metaphysical distinction (I set aside, for now, whether it gives rise to a moral distinction). In the same vein, when person B would have died had person A not acted she did—when B's survival is a consequence of A's actions—in some cases A saves B's life and in other cases she doesn't. Many familiar cases fall into the former camp: A drags B from the burning building; A resuscitates B; A shoots C just before C shoots B, etc.; and each time A saves B's life. In contrast, here are three examples where B would have died but for what A did, yet A doesn't save B's life:

> A simple case of commission: Annie brakes at the red light while Ben crosses the road in front of her. If Annie hadn't braked, then she would have run Ben down; Ben would have died, but for what Annie did, yet Annie didn't save Ben's life.
>
> A simple case of omission: a runaway trolley is heading towards a trapped worker. Thankfully, its momentum has disturbed a boulder which is about to fall onto the tracks, shielding the worker. Annie is about to stop the boulder, but decides otherwise. If Annie hadn't decided otherwise, then the trolley would have killed the worker; the worker would have died, but for Annie's deciding not to stop the boulder, yet Annie didn't save the worker's life.

> An interesting case of omission: a second runaway trolley is 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. If Annie hadn't kept quiet, then Ben would have steadied himself and the trolley would have killed the five workers; the workers would have died, but for Annie's keeping quiet, yet Annie didn't save their lives.

(I add that scepticism about these (metaphysical) claims would be inappropriate: it's plain that Annie doesn't save those lives—imagine Annie boasting otherwise. Just as it's plain that A saves B's life when she, e.g., resuscitates him; just as it's plain that I kill you by running you down with my car; just as it's plain that I don't kill you when I ignore your drowning cries.)

So, when B's survival is a consequences of A's actions, it might be that A saved B's life or it might merely be that B would have died but for what A did. That phrasing—"merely that B would have died but for what A did"—is cumbersome and it will be useful to have something that better rolls of the tongue. As I said above, I have adopted *letting live*, where

> A lets B live $=_{df}$ B's survival is a consequence of something A does, yet A doesn't save B's life.

So there is a metaphysical difference between saving and letting live, but 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 that our reasons not to kill another are, other things equal, stronger than our reasons not to let another die. And those same people have thought that that difference shows up in cases where we should let five people die rather than kill one person (e.g. when the one person is blocking our path to the five).[1] And 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 survives, it also makes a difference whether A saves B or A lets B live.

---

[1] This originated with Foot (1967, 2002), was discussed by Thomson (1976), has been objected to by, e.g., Bennett (1998) and Kagan (1988), and defended by, e.g, Woollard (2008). Exactly what that difference amounts to is controversial—and I don't take a stand on it here. Foot said it amounted to it being impermissible to kill one to save five. Thomson (1985) objected to that, but Thomson (2008) defended it—as did my REDACTED.

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 survival). 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, other things equal, we have more reason to not kill someone than we do to not let someone else die; and the (supposed) moral importance of the distinction between saving and letting live holds that, other things equal, we have more reason to save someone's life than we do to let someone live.[2]

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? I say that there is. More precisely, I say that we have just as much reason to think there is such a difference as we have for thinking that there is a moral difference between killing and letting die: the two differences come as a package. I set aside, for elsewhere, whether, as many have thought, there is in fact a moral difference between killing and letting live.

In the next section I argue for that moral difference; in §§4-5 I theorise about what it is for A to save B's life; and in the remainder I consider the moral significance of (merely) letting someone live.

3. We've already seen one case where the moral difference between saving and letting live shows itself. Recall the interesting case of omission, above:

> A runaway trolley is 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.

For ease, I will call this case WIND. To my mind, Annie acts impermissibly in this case: when I put myself I her shoes, I conclude that I would alert Ben to his peril. And everyone I've asked about WIND says the same, but they've also found that verdict puzzling. They scratch their heads and say

[2] I needn't have stopped there: if we have more reason not to kill than we do not to let die, then presumably we also have more reason not to paralyse than we do to not let be paralysed; and more reason not to break a leg than to not let a leg break; etc.

"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?"

And they might have continued. Along with consequentialist principles and familiar deontological differences between, e.g., killing and letting die, they might also have considered the Kantian requirement that we not treat people as means—that Annie not treat Ben as a means. But Annie doesn't treat Ben as a means, she doesn't treat him as anything, she doesn't do anything to him; as far as Ben is concerned, Annie might not have even been there.[3] So they might continue to scratch their heads and wonder why it's impermissible for Annie to keep quiet. What are they missing?

They're missing two important factors about the case. The first is that:

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

And that fact contrasts with another, 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—she let them live.

(Just as, in deciding against stopping the falling boulder, Annie didn't save the life of the trapped worker in the simple case of omission, above.)

And those facts, when combined with the moral difference between saving and letting live explain why it's impermissible for Annie to keep quiet in WIND: because there is more reason to save another's life than there is to let another person live; sufficiently much more reason that it's impermissible for Annie to fail to save Ben's life in WIND, even though, by doing so, she will let the five workers live.

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

---

[3] "But Ben's life was a means to the survival of the five, and Annie let that happen. Isn't that sufficient for Annie to treat him as a means (and thus sufficient explanation of why it was impermissible for her to keep quiet)?" Setting aside how this would extend the concept of "treat as a means" beyond any natural usage, it still wouldn't be helpful here. Suppose that the track loops around behind the workers and continues over the footbridge such that either (i) Ben falls and his life is a means to the survival of the five, or (ii) he doesn't fall and then the lives of the five workers are a means to the survival of Ben (since they would block the trolley from reaching his position). Whatever Annie does, someone's life will be a means to the survival of another, and she will be letting that happen, yet it's still impermissible for Annie to keep quiet.

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." The answers to these two questions stand or fall together. To my mind, their answers are not deep. When we want to know what the moral significance of my doing such-and-such is, it matters—obviously! plainly!—what I will *do* if I such-and-such: it matters whether I will *kill* someone or not (even if the total number of deaths is constant); and it matters whether I will *break* someone's arm or not (even if the total number of broken arms is constant); and it matters whether I will *save* someone's life or not (even if the total number of survivals is constant). In a credo: what someone does is not exhausted by the consequences of they bring about; action is not production.[4]

So there is a moral difference between saving and letting live and that difference explains why it's impermissible for Annie to keep quiet in WIND.

That difference also neatly solves a puzzle, from Re'em Segev (2015). He considers the following sort of case:

> A runaway trolley is again heading towards five trapped workers. A consequentialist is about to push the fat man off the footbridge into the trolley's path (the only possible way of stopping the trolley). Annie, a champion wrestler, is able to to harmlessly restrain the consequentialist.

Segev says (surely, rightly), that Annie's required to restrain the consequentialist, but is puzzled as to why that would be so. After all, as he says,

> common theoretical accounts of deontological wrongdoing forbid positive agency, for example doing harm or intending harm. Yet not preventing harm does not in itself amount to doing or intending harm—it is possible not to prevent harm without doing or intending harm. Therefore, standard deontological accounts do not entail a reason in favor of preventing such wrongdoing (2051).

That's because those common deontological accounts have ignored the difference between saving and letting live. After all,

> if Annie restrains the consequentialist, then she will save the fat man's life,

but

[4] See Tamar Schapiro (2001).

> if Annie does nothing, then she will save no one's life and merely let the five workers live (and let the fat man die)

and the difference between saving and letting live is such that Annie has more reason to save the fat man's life than she has to let the five workers live.[5]

Segev's case and WIND are both of a similar flavour.[6] I close this section by turning to another reason for thinking there is a difference between saving and letting live—one that has been staring at us for ages.

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

[5] For more on this problem of so-called "preventing deontological wrongdoing," see Mogensen (2016) and Sinclair (2017).

[6] Another problem of that same flavour was raised 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.

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. Suppose I had donated $4000 to the Against Malaria Foundation last year, sufficient for them to prevent a death in the aggregate. I'm certain that I wouldn't thereby have saved a live—just imagine me boasting otherwise! Regardless, the next section will explain why I couldn't have saved a life via that donation.)

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.

The moral difference between saving and letting live has been staring at us for ages: it is the difference between Singer's two cases.

4. Over the preceding sections I have asserted of various cases that so-and-so did (or didn't) save someone's life by doing such-and-such. In §2 I commented, after some such assertions, that scepticism about those particular claims would be misplaced. That's because those assertions concerned clear cases: if A resuscitates B, then A plainly saves B's life; if, as in the simple case of

omission, A doesn't interfere with the boulder that is tipping onto the track between the trolley and B, then A plainly doesn't save B's life. (Just as if A shoots B, then A plainly kills B, etc.) What goes for those cases, goes for the majority of other cases too, since the majority of cases are similarly clear, and much progress can be made with just them.

But what about unclear cases? Just above, I claimed that Annie wouldn't, by donating $4000 to Oxfam, save anyone's life (even though, $4000 is what it costs Oxfam, in the aggregate, to save a life). And, in conversation, some people have said that that is an unclear case. What are we to do about such cases?

One response is to search for an analysis of saving (and of killing and of breaking, etc.—or perhaps a single analysis covering them all) which will tell us whether A saved B for any particular case. I have little hope for such analyses—both here and more generally (REDACTED). Instead, I think the best we can do is take the unclear cases as they come and, slowly, build up some theory—no small task. I get that task underway here by focusing on donations to Oxfam; I continue with shifted focus in the next section.

Last year, Agatha donated $4000 to Oxfam (which was sufficient, in the aggregate, for them to prevent a death); and last year Annie dragged a drowning child from a pond. I claim that Annie thereby saved a life, but Agatha did not. Why is that?

Here is a simple principle: if A saved someone's life, then there exists some B such that A saved B's life. If this principle is right, then it would account for why Annie did save a life last year, but Agatha did not. After all, Annie saved the drowning child, but there is no determinate person that Agatha saved by donating that money to Oxfam. Instead, her $4000 went into a pot of millions, which Oxfam then distributed (likely months later) to its many beneficiaries: what could possibly fix whether Agatha saved this beneficiary or that beneficiary or that other beneficiary?

But is the principle right? It certainly seems right, particularly compared with its kin: if A broke a window, then there exists some window that A broke; if A killed someone, then there is some B that A killed B; if A saw someone, then there is some B such that A saw B. And it certainly seems right when we consider the semantics of 'A saved someone's life': what else could that be but 'there is some

person B such that A saved B.'And it certainly seems right when measured against the truths its denial would give rise to: 'I saved someone's life last year, but there is no particular life that I saved' (and, relatedly, 'I saw someone in the room, but there is no particular person that I saw'). Is there good reason to reject it, nonetheless?[7]

Here's one case that might give us pause (borrowed from Parfit 1986). Suppose persons B1 and B2 are trapped in a flooding mineshaft and can be lifted to safety only if persons A1 and A2 stand on particular platform. Isn't it the case that A1 saved someone's life by standing on the platform, even though there is no particular person that A1 saved (since it would be arbitrary to pick between B1 and B2)? It's not: A1 (together with A2) saved both B1 and B2; A1 saved two lives that day, she just did it together with A2 (just as Annie and the helicopter pilot each saved the drowning child's life together). It is a mistake of metaphysical mathematics to calculate how many lives so-and-so saved by dividing the total number of lives saved by the total number of savers—two surgeons operating all night do not each save half a life, but, instead, each of them saved a single life.

(I suppose the zealot might want to insist that what goes for A1 here also goes for Agatha: *viz*. that, by donating $4000 to Oxfam, she saved, e.g., a million lives (or however many lives it is that Oxfam saved this year). There's little to recommend that conclusion: Agatha plainly didn't save a million lives. Why should Agatha come apart from A1 in this way? This is a tricky question, but one difference between them is that while it's true that fewer people would have been saved had A1 not stood on the platform—or if the first surgeon had not operated—it's not true that fewer people would have been saved had Agatha not donated her money.[8])

---

[7] This principle shouldn't be confused with its counterfactual cousin: *viz*. if A would have saved someone's life by ϕ-ing, then there is some particular person B such that A would have saved B's life by ϕ-ing. That principle is falsified by cases in which who A would have saved would have been determined by, e.g., the roll of a die. However, even in a such a case it would be true that: if A had ϕ'd then there would have been a particular person such that A saved him (i.e. the person whose number came up on the die).

[8] Why isn't it true? The number of lives Oxfam saves this year is heavily influenced by this year's weather and, in turn, the natural disasters that weather causes. Different weather would have resulted in different disasters, different needs, different costs of satisfying those needs and, in turn, a different number of people saved by Oxfam. But which natural disasters strike this year is *extremely* sensitive to everything that happened last year—including whether Agatha donated that money. And if she hasn't donated it, it would be sensitive to exactly what she would have done instead (moving in *this* particular way, or *that* particular way, etc.)—yet there's just no fact of that matter and, in turn, no fact of the matter as to how many lives Oxfam would have saved this year had Agatha not donated. See Hare 2011 and REDACTED.

So why didn't Agatha save anyone's life with her $4000 donation to Oxfam? Because there's no person that she saved and if there's no person that she saved, how can it be that she saved someone's life?

Charities are, in some sense, live to all this. It's not a coincidence that they're keen to connect donors to particular beneficiaries ("Here is so-and-so with her new malaria net, purchased with your money") and doubtless this is because it leads to more donations. The preceding says that that's because it makes donors feel like they are saving that beneficiary's life, even though they aren't. Just because a charity can, after the fact, assign to each of its donors someone that they (the charity) has saved, that doesn't mean that the donors each saved the beneficiary to whom they are assigned—and they didn't.

This might make us all wonder whether there is any way that Oxfam could run such that its donors would save the beneficiaries of their donations. Here's one way (though I don't endorse it). Oxfam generates a list of people who will die without immediate assistance. Oxfam's canvassers then pick a person from that list, B, and pitch to potential donor as follows (speaking truly, throughout): "Excuse me, Agatha. This person, B, is in dire need of assistance. We have the infrastructure to provide it to him, but we don't have the funds to first purchase the goods. If you donate today, then we will use that money to purchase those goods and deliver them to B. If you don't, then B will be forever struck from our list and will certainly die. You are the only one who can, with our help, save his life." Under such an operation, Agatha's donation would go to a determinate beneficiary and, in turn, Agatha would save that beneficiary's life (much like Annie would have saved the child's life by buying the helicopter pilot his fuel).

5. We can expand our theory of saving by tying saving to something more familiar—*viz.* killing. By means of introduction, recall:

> A simple case of commission: Annie brakes at the red light while Ben crosses the road in front of her. If Annie hadn't braked, then she would have run Ben down; Ben would have died, but for what Annie did, yet Annie didn't save Ben's life.

While it's plain that Annie didn't save Ben's life here, we might wonder why she didn't. After all, she did stop several tonnes of metal from running Ben down and why, we might wonder, isn't that sufficient for Annie to have saved his life?

It's not sufficient because saving and killing are *duals*: you do not save someone's life by merely failing to kill them and nor do you kill someone merely by failing to save them. More precisely:

> if A would have killed B by ϕ-ing, then A does not save B's life by not ϕ-ing.

Annie would have killed Ben by running him down and thus, by the preceding, it's not the case that Annie saves Ben's life by not running him down. Similarly, I would have killed you had I refused to move my car out of the fire engine's path (supposing that only those firefighters can rescue you from the flames), and therefore I don't save your life when I do move my car (thus enabling the firefighters to rescue you). And the converse:

> if A would have saved B's life by ϕ-ing, then A does not kill B by not ϕ-ing.

Annie would have saved the child's life by dragging her from the pond and, therefore, Annie does not kill the child by failing to drag her from the pond.

So saving and killing are duals.[9]

I said in the previous section that our best response to tricky cases—to those cases in which it's unclear whether A saves B's life—is to build up a theory of saving. That's no small task, but this section and the previous one have gotten it underway.

6. When B dies, it might be that A killed B or it might be that A let B die, where

> A lets B die $=_{df}$ B's death is a consequence of something A does, yet A doesn't kill B.

We now know that that isn't the end of the story since, when A lets B die it, might also be the case that A failed to save B's life. That is, for example, what happened in WIND: Annie didn't merely let B die, by keeping quiet, she failed to save his life. Where we thought there was only a two-way metaphysical distinction there is in fact a three-way metaphysical distinction, giving rise to the following principles:

> SLL (saving-letting-live): other things equal, we have more reason to save someone than we have to let someone live.

[9] In fn.12 this theory is put to work.

> KFS (killing-failing to-save): other things equal, we have more reason not to kill someone than we have to save someone.
>
> KLD (killing-letting-die): other things equal, we have more reason not to kill someone than we have to not let someone die.

The first principle, SLL, explains why, e.g., it's impermissible for Annie to keep quiet in WIND. The second principle, KFS, explains why it's impermissible to cut up a healthy patient in order to distribute his organs into five, otherwise terminal, patients. (Deontologists have traditionally explained that fact by instead appealing to KLD, but we now know that that's a mistake.) We haven't yet seen cases that fall under KLD's scope, nor do they arise naturally, but here is one example:

> Only Hero can save five hostages, however, he is too preoccupied to do so. To spur him into action, Annie kills Hero's partner—framing the bad guys. Hero then rescues the five.

KLD explains why it's impermissible, as surely it is, for Annie to kill Hero's partner: because she does so to let the five hostages live (she doesn't save them, Hero does), yet she has more reason not to kill Hero's partner than she does to not let the others die.

I said in the introduction that there is a metaphysical difference between saving and letting live, that that metaphysical difference corresponds with a moral difference and, in turn, that we can do a lot with it. Much of what we can do with it stems from the three principles above: they explains what's going on in cases like WIND and they also explain the difference between Singer's two cases. My suspicion is that this is just the beginning and that now we're live to the difference between saving and letting live we'll see that difference cropping up in various places throughout moral theory.

7. There's a moral difference between saving and letting live, such that we have more reason to save someone's life than we have to let another live—or, even, to let five others live. But what is the moral significance of (merely) letting someone live?

Asking around, everyone says (assumes?) that letting live has some moral significance. 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 significance 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 significance and if that significance aggregates (as presumably it would) such that there is a number *n* for which it's permissible to fail to save one person in order to let *n* people live, then we're similarly clueless as to whether it's permissible, e.g., for Annie keep quiet in WIND. (I present this as a problem for SLL, but I could just have easily have done so for KFS or KLL.)

We're clueless because we're clueless as to whether Annie will let *n* people live by keeping quiet. To adapt James Lenman's canonical example, suppose that it's 100 BC and an analogue of WIND is playing out on the banks of the Rhine, this time starring Agatha and Burt:

> a boulder is heading towards five trapped villagers and a gust is about to blow Burt into its path, lest Agatha alert him to his peril.

Agatha does alert Burt, but was it permissible for her to do so?[10] With 2100 years of hindsight, we now know that Burt was a many-times-great grandfather of Hitler. Thus, if Agatha had kept quiet (and Burt had fallen), 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 was permissible for Agatha to keep quiet: after all, if Agatha had kept quiet and the five villagers had lived, then perhaps one of their descendants would have been worse still than Hitler (e.g. Malcolm the *Truly* Terrible) such that, by letting those villagers die, Annie would have let all of Malcom's would-be victims live—we just don't know (nor will we ever know).

And what goes for Agatha and Burt goes for Annie and Ben: just as we're clueless now as to whether it was permissible for Agatha to alert Burt to his peril, so too are we clueless whether it's permissible for Annie to keep quiet in WIND.[11]

At least, we are if letting someone live has some moral significance. And it's easy to see why that significance 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

---

[10] Lenman (2000).

[11] 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.

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 significance 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 significance."* 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 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 considers the simple case of omission's counterfactual, in which A does stop the boulder from tumbling into the path of the runaway trolley. 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 prevents a barrier from protecting B, then A kills B).[12] [13] 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*

---

[12] A reviewer asks about a variation of this case in which it isn't a boulder tumbling onto the tracks, but instead A's (expensive! shiny!) car, after its parking brake fails. The reviewer (rightly) says that, in that case, A wouldn't kill B if she manages to stop the car in time, but, instead, that she would merely fail to save him. The reviewer then asks how that conclusion is supposed to fit with my claim that A kills B if she stops the boulder. Interestingly, the duality of saving and killing is diagnostic here. After all, since A would save B's life by letting her car roll into the trolley's path, it follows from the duality that she wouldn't kill B were she to stop the car from rolling into the trolley's path. Yet the same doesn't hold of the boulder: A doesn't save B's life by letting the boulder tumble into the trolley's path and that is compatible with her killing B when she stops it tumbling. Roughly, what's making the difference must be whether A brought the barrier (the car, the boulder) to the situation or not. I set aside exactly how that is to be made more precise.

[13] 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.

shouldn't alert Ben in that case and surely that's because letting someone live has some moral significance 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 significance 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 significance of letting the worker live since even if letting one person live has some moral significance, 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 doing so.[14] 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

[14] 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.

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.[15] 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 significance of letting live.

Objection 3: *"But if letting live has no moral significance, 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, 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 significance (indeed, having the same moral significance 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

[15] 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).

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.

Objection 6: "*If Annie can either ϕ and let one person live or Ψ and let five people live, then surely, other things equal, Annie is required to Ψ. But if that's so, then letting live must have moral significance, after all.*"

As with the first objection, this schema is not as easy to satisfy as it might appear. Here is one attempt that won't do: one person is drowning over here while five people are drowning over yonder and A can make it to only a single group in time. Yes, A should save the five, but she would be saving them—not letting them live.

Another sort of case is one in which ϕ-ing and Ψ-ing are donating to less- and more efficient charities, respectively. Like with the third objection, this is another strange objection since it again relies on an intuition that is not held since we don't think people are always obliged to donate to the most effective charities: we think it is permissible for someone to donate to their local community centre, even when they could instead have donated to the Against Malaria Foundation. Singer (and others) might have tried to convince us otherwise on this point too, but those arguments similarly rely upon letting live having moral significance.

A final sort of case is tricker. Consider this variation of WIND:

> two runaway trollies on parallel tracks are heading towards one trapped worked and two trapped workers, respectively. B is about to be blown off the bridge into the first trolley's path, while C is about to be blown off the bridge into the second

> trolley's path. A can either alert B to his peril or alert C, but not both. If she alerts B, she will save one life and let two live (the two who will be shielded when C falls); if she alerts C, she will similarly save one life, but she will let only one live (the one who will be shielded when B falls).

If letting live has no moral significance then, other things equal, A is free to save either B or C (perhaps tossing a coin to decide). I say that's right. Some readers will find that objectionable. They'll say that—surely! plainly!—A is required to alert B and thereby save one life and let two people live (as opposed to only one) and, therefore, they'll conclude that letting someone live does have moral significance, after all. I say that's mistaken. I say that such readers have ceded too much to the consequentialist and that they have stopped seeing B and C as people and started seeing them as means—in this case, means to blocking runaway trollies. I say they have had one thought too many. (Compare: an oncologist and a postman are each about to fall and you can only save one. I say you have similarly ceded too much to the consequentialist if you reason towards saving the oncologist because, in doing so, you'll let more people live over the coming years.)

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 save and whom we kill (and, by extension, whom we fail to save and, occasionally, whom we fail to kill). The result is that our sphere of 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 ϕs*.

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 who, 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 KFS 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.[16]

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.

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 is itself a moral notion, these friends ring the death knell for KFS, KLD and SLL. 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.[17]

---

[16] 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."

[17] REDACTED.

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 whoever is standing behind.)

In what way is it sensitive? Roughly, it seems to be that whether A kills B in certain cases 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 both rough and of limited scope. After all, A surely kills B when she throws him into the path of a trolley, regardless of whether it was already heading towards another victim.

How to make it more precise and more general is, I think, a deep question (and not one that obviously has a good answer). Regardless, it's not one that I will address here.

Jonathan Bennett, *The Act Itself*, (Oxford University Press, 1998).
Christopher Boorse and Roy A. Sorensen, "Ducking Harm," *Journal of Philosophy* 85 (1988): 115–34.
Philippa Foot, "The Problem of Abortion and the Doctrine of the Double Effect," *Oxford Review* 5 (1967).
Matthew Hanser, "Why are Killing and Letting Die Wrong?," *Philosophy & Public Affairs* 24 (1995): 175–201.
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.
Derek Parfit, *Reasons and Persons*, (Oxford University Press, 1986).
Tamar Schapiro, "Three Conceptions of Action in Moral Theory," *Noûs* 35 (2001): 93–117.
Re'em Segev, "Should we prevent deontological wrongdoing?," *Philosophical Studies* 173 (2016): 2049–68.
Thomas Sinclair, "Preventing Optimific Wrongings," *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.