

# The Rejection of Consequentializing

Daniel Muñoz

Forthcoming in *The Journal of Philosophy*

Consequentialists say we may always promote the good. Deontologists object: not if that means killing one to save five. “Consequentializers” reply: this act is wrong, but it isn’t for the best, since killing is worse than letting die. I argue that this reply undercuts the “compellingness” of consequentialism, which comes from an outcome-based view of action that collapses the distinction between killing and letting die.

## 1. Introduction

Act consequentialists think that acts are right when they have good outcomes, so that it’s never wrong to make things go best. The classic complaint: there seem to be counterexamples. Five lives are better than one, for instance, and yet you may not kill me even if my organs will be used to save five sick people (Foot 1967). Best outcome, wrong act—what does that mean for consequentialism?

Many ethicists, though they accept that killing is wrong, still feel “compelled” to remain consequentialists. So they say that killing really *does* make things worse, because of something bad about the act itself: five deaths are a worse outcome than one, but not worse than one death plus a killing. Of course, we don’t normally think of an act as part of its own outcome. Our normal concept is that of a *narrow* outcome: the sum of acts’ downstream effects (Thomson 2003: 8, fn. 2)—it’s “what happens minus the fact that the agent made it happen.” But there is also the kosher notion of a *broad* outcome (Portmore 2009: 330), which includes the act itself, and so consequentialists may insist that killing one to save five is wrong because its broad outcome is suboptimal.

This reply—known as “consequentializing”—is now itself a classic. But there is something funny about it. Just ask: how could it be worse *by consequentialist lights* to kill and save five? The killing’s narrow outcome is the best available: less death. And, given consequentialism, we can’t say that the act is worse because it’s wrong (that gets the explanation backwards), or because it violates a non-consequentialist right to life (there isn’t any). So what could be bad about an act *narrowly aimed at*

*the very best?* The consequentializers' answer must be that the act is bad fundamentally. I argue that, if this were the case, it would undercut their own reason for being consequentialists.

This is a new objection. I am not just wheeling out the intuition that one killing doesn't seem as bad as four deaths (Scheffler 1994: 109). Nor am I rehashing Nozick's (1974) point that, if killings are only wrong because they are bad, then we ought to minimize murders, even when that means committing one ourselves—which is intuitively wrong. I'm not here to insist on my intuitions. I want to bring out a tension in the consequentializer's overall view.

What “compels” the consequentializer is the idea that it can't be wrong to produce the best outcome. But what *makes* this idea compelling, I argue, is a conception of action: the view that to act *just is* to produce an outcome. Consequentialism is only compelling if we are already inclined to see action as production.<sup>1</sup> Now, some philosophers have directly attacked the “compelling idea” and this view of action (Hurley 2017, 2019b). I will not. I grant that action might well be production, and that consequentialism could be genuinely compelling. But for the consequentializer, who draws brute distinctions between acts with the same effects, it makes no sense to insist that outcomes are the only point of action. The consequentializer can't be compelled towards consequentialism.

After some background (sections 2–4), I argue for my two premises: that consequentialism is only compelling given the view that action is production (section 5), and that consequentializing undercuts that view of action (section 6).

## **2. Consequentialism**

Let's start by getting clear on the target. Who are the “consequentialists?” I have in mind a thin definition. A consequentialist is anyone who believes that an act's deontic properties (rightness,

---

<sup>1</sup> I borrow the name “Action as Production” from Korsgaard 1996: 217 and Schapiro 2001: 95.

wrongness, etc.) are explained exclusively by the evaluative properties of its outcome (goodness, desirability, etc.). The right is grounded in the good.

At the core of consequentialism is the so-called “compelling idea” that it is never wrong to make things go best. At the periphery, consequentialists disagree about *which* things are good (just pleasure, or something more?); they may also disagree about whether *maximizing* the good is always required (Yetter Chappell 2019); whether rightness is *sensitive* to the agent’s imperfect knowledge (Jackson 1991); whether goodness is *relative* to an agent at a time (Schroeder 2007); and whether the theoretical role played by “goodness” should be played by a close relative (like the *reasons to prefer* one outcome to another; see Dreier 2011: 101, Portmore 2011, Setiya 2018). Whatever you think about these issues, you can’t be a consequentialist if you see rightness as fundamental, or as grounded in something other than lovely outcomes—responding to rights, expressing respect, etc.

So I will abstract away from some insider debates when using the label “consequentialist.”<sup>2</sup> But let me emphasize that I use this label just for convenience. My point is that anyone who grounds the right in the good should resist the idea that it’s bad to kill the few to save the many. Nothing turns on whether we use “consequentialist” broadly, as a name for all of these authors, or narrowly to name a special subset.

### 3. Consequentializing

In the beginning, consequentialists agreed that acts could have value only extrinsically, as a means to producing the really intrinsic goods. Some thought the sole good was pleasure (Sidgwick 1874); others believed in a plurality of goods including beauty and friendship (Moore 1903). But no one thought of acts as good in themselves.

---

<sup>2</sup> I am talking only about act consequentialism. On rule consequentialism, see Hooker 2000. For more on the definition of consequentialism, see Sinnott-Armstrong 2015.

This would change in the 20<sup>th</sup> century, as consequentialists partook from the Tree of Counterexamples. Some acts smacked of wrongness despite their wholesome effects—framing the innocent to stop a riot, breaking a promise for expedience, etc. Here is a modern case:

*Footbridge*

A runaway trolley is headed toward five innocent workers. You are on a footbridge over the tracks beside a large man. Push him to his death, and his body will halt the trolley, saving the lives of the five. There is no other way to save them. (Thomson 1985)

The man on the bridge does not consent to being pushed, and his life promises to be just as good as the life that any of the workers would enjoy if saved. Common sense tells us that it would be for the best if the large man happened to fall in front of the trolley, as if blown by the wind—and yet it is intuitively wrong for you to shove him there.

Old-school consequentialists want to resist this intuition. I won't object. My target is the new school, who try to absorb the intuition of wrongness by *consequentializing* common sense. To consequentialize a moral view is to cook up a consequentialist counterpart that preserves the original's claims about which acts are right and which are wrong by changing what the view considers good and bad.<sup>3</sup>

How does this work? The standard “recipe” is for the consequentializer to say that it is *intrinsically bad that the act was performed*—in this case, the act of killing one to save five (Dreier 2011: 9–99; cf. Portmore 2014: 251). Thanks to the badness of the act itself, killing is wrong, after all.

It makes perfect sense that this would be the recipe of consequentializers. To account for the wrongness of killing, they need to locate something in the outcome that is unexpectedly bad. But what? Not the narrow outcome. An act's narrow outcome, as a reminder, consists in its downstream

---

<sup>3</sup> For recent work in the consequentializing tradition, see Dreier 1993, 2011; Louise 2004; Peterson 2010: 155; Brown 2011; and Portmore 2007, 2011. The classic precursor is Ewing 1939, 1947. On the differences between consequentializers, see Schroeder 2017. (My target, in Schroeder's terms, is the “intuitive argument” for consequentializing, not the “assimilation” and “pragmatic” arguments.) For objections, see Sachs 2010; Hurley 2013, 2019b; and Tenenbaum 2014.

effects. But the effect of pushing the man off of the Footbridge is that he falls and *only one dies*; the effect of refraining is that the man doesn't fall and *five die*. Clearly, less death is better.<sup>4</sup> So there is only one place left to look within the "broad outcome," and that is the act itself (or in the fact that it was done; this distinction doesn't matter for us).<sup>5</sup>

So the consequentializer has a recipe, and it seems to handle Footbridge fine, by putting badness in the act of killing. Indeed, Mark Schroeder thinks that the consequentializers can "easily accommodate" cases like Footbridge; the recipe might not lead to "a plausible view about how bad murders and deaths are," but it nonetheless "shows that consequentialism as such is not directly threatened by such cases" (2007: 266). The real trouble only starts, he thinks, when we bring in more elaborate examples, such as:

*Villainous Footbridge*

You are on the footbridge as before, but this time, the trolley was started up by a villain. Kill the large man, and you will reduce the number of murders on the tracks from five to one.

Here killing seems wrong, but there is no easy way to accommodate this judgment within consequentialism, except by saying that killing is bad *from the point of view of the agent*—one must reduce one's own murders. And even that would not help with:

*Redemption Footbridge*

Same as Villainous footbridge, but the villain was *you*. Kill the large man, and you will reduce the number of murders *you* end up committing from five to one.

To treat this case, we would have to make the badness of killing relative to the time as well as the agent.<sup>6</sup> And so Schroeder thinks that the heart of the debate is here in the realm of relativity. The

---

<sup>4</sup> It's not essential that the choice involves a difference in the number of people harmed. We could also imagine a choice of whether to harm one person to prevent a bigger harm to another.

<sup>5</sup> Even if we include the past in an act's broad outcome, clearly the badness doesn't belong in the past. For one thing, nothing is particularly bad about the situation in Footbridge until the trolley hits. More importantly, the past is the same no matter what the agent does, so facts about the past can't make the outcome of pushing *worse* than the outcome of not pushing.

<sup>6</sup> Setiya (2018) argues that consequentialists can handle the Footbridge cases without appeal to

question is: should we believe in values relativized to times and moments?

That's a good question. But I think the consequentialist is "directly threatened" well before relativity enters the picture. There is something unstable about a consequentialism on which it's bad to cause good results, as does the killer in Footbridge. The core problem here isn't about whether killing is bad in a relative or a neutral way. For consequentialists, the act should not be bad at all.

#### 4. Why Not Consequentialize?

I hope you are feeling puzzled. We all learn in Ethics 101 that consequentialists, who ground the right in the good, are supposed to start with a list of Goods and Bads. Sometimes the list of Bads is as short as simply "pain," but of course the list might be longer, featuring a variety of things—vice and error, injustice and inequality.

What's wrong with adding one more thing to the list? Why can't the consequentializer just say that killing is another fundamental Bad?

One answer, mentioned earlier, is that killing one to save five isn't intuitively so awful—at least not by consequentialist lights. I think there is something to this point. It is worth reminding ourselves that the consequentializer is cramming badness into an act with an unimpeachable narrow outcome: four fewer deaths. This can't be justified by a prior concern for death-reduction. On the contrary: the act that produces *less* death has *more* badness. (Letting the five die is intrinsically *better* than killing one.) Old-school consequentialists will scratch their heads.

But this isn't news to the consequentializers. They agree that wrongful killings can have good effects, and they aren't trying to derive the badness of killing from anything else. Their point is that killing is bad fundamentally. This is not meant to be instantly obvious, but only to recapture,

---

relative values. It is not clear to me whether he is in the scope of my arguments, since he does not commit to the (more standard) view that the killing is wrong because of the badness of the act itself.

from a consequentialist point of view, our sense that it's wrong to kill the one.

So why not consequentialize? My answer is that the consequentializer is destabilizing their own view, since there is a tension between their initial motivation—the “compelling idea”—and the claim that killing is fundamentally bad.

The consequentializer's view has three parts: a theory of what makes things good, a theory of what makes acts right, and an account of what actions are. They are compelled by the idea that it's never wrong to produce the goods. But this idea is only compelling given Action as Production—the view that to act *just is* to produce an outcome—and this view flies out the window once we open the door to a distinction between killing and letting die.

My argument has two premises:

*Compelled by Production*

Consequentialism is only “compelling” given Action as Production.

*Undercut by Consequentializing*

If killing is worse than letting die, then Action as Production is false.

The conclusion is that, if the consequentializer is right about killing, consequentialism isn't really compelling. I'll now argue for the premises in turn.

## 5. Why Consequentialism Compels: Action as Production

“Ever since its inception,” writes Tamar Schapiro (2001: 91), consequentialism has been “reinforced if not entailed by a particular picture of action”—Action as Production. On this picture, the point of action is to make things happen, to change or preserve states of the world.

Action as Production is a view about what actions are: for any act A, there is a state S such that doing A *just is* to aim at producing S.<sup>7</sup> Acts cannot differ unless they aim at producing different

---

<sup>7</sup> I won't need to distinguish an *action's* aiming at producing S from the *agent's* aiming to produce S. But perhaps, in the case of unintentional acts, an action has an “aim” that the agent doesn't.

outcomes. What is it to “produce” an outcome? In the ordinary sense, producing means causing.<sup>8</sup> I produce the death of the big man if I push him off the bridge, but I don’t “produce” five deaths on the tracks when I refuse to push, nor does my refusal “produce” the remainder of the big man’s life.

But consequentializers construe “production” more broadly. An act produces whatever happens when the act is done, no matter whether the act was a direct cause, a constitutive ground, or a failure of prevention. In this broad sense, acts even “produce” themselves. Portmore explains this very nicely:

It is through our actions that we affect the way the world goes...Moreover, whenever we act intentionally, we act with the aim of making the world go a certain way. The aim needn’t be anything having to do with the causal consequences of the act. The aim could be nothing more than to bring it about that one performs the act. For instance, one could intend to run merely for the sake of bringing it about that one runs. The fact remains, though, that for every intentional action there is some end at which the agent aims. (2011: 56; cf. 2014: 246)<sup>9</sup>

To act is to aim at an end, but the end can include the act itself.

That’s Action as Production. Now, why is it so crucial for the consequentializer? One simple reason is that it directly supports consequentialism. The picture of action nudges us toward the moral theory. Thus Mill slides from action to ethics:

All action is for the sake of some end, and rules of action, it seems natural to suppose, must take their whole character and colour from the end to which they are subservient. (1861: 2)

If acts aim at ends, we “naturally” infer to an end-based morality. And here is Portmore:

...since our actions are the means by which we affect the way the world goes, and since our intentional actions necessarily aim at making the world go a certain way, it is only natural to suppose that our reasons for action are a function of our reasons for preferring some of those possible worlds to others, such that what each agent has most reason to do is to bring about the possible world, which of all those that she can actualize through her actions, is the one that *she* has most reason to want to be actual.<sup>10</sup> (2011: 56)

---

<sup>8</sup> Some traditional consequentialists, like the utilitarian Richard Cumberland (1727: 196), explicitly say that action is causation (Schapiro 2001: 94–95).

<sup>9</sup> Another common way to put it: an act’s outcome is the world that would be actual if the action were performed (see e.g. Portmore 2009: 330, Brown 2011: 752).

<sup>10</sup> Portmore’s views here are subtler than I’m letting on. In this passage, he is introducing “the



This is just like Mill, right down to the phrase “natural to suppose.” We even find the same current running through the the work of non-consequentialists, such as Judith Jarvis Thomson:

given that for a person to act just is for the world to go in a way that it otherwise would not go, surely the question whether he ought to act had better turn on a comparison between how it will go if he acts and how it will go if he does something else—to repeat, there seems to be nothing else for it to turn on. (2003: 8)

I agree. If action *isn't* production, then consequentialism loses an easy source of luster.

But as Thomson hints, there is a deeper rationale for Action as Production: it is part of what makes consequentialism feel inescapable (“nothing else”), part of what makes it so compelling.

The consequentialist’s “compelling idea” is that we may always make things go best: there is no such thing as a wrong act with the best outcome.<sup>11</sup> This idea is what steels the consequentializers’ will as they persevere through hard cases like Footbridge. No wonder. A natural reaction, on first hearing the compelling idea, is to ask how on earth it could be false.

Now, I won’t claim that the compelling idea is false, as the deontologists think, or that it’s dangerously ambiguous, as Paul Hurley argues.<sup>12</sup> I am happy to grant that the consequentialist’s compelling idea is clear and attractive in the abstract. My claim is that the idea loses its luster if we drop Action as Production.

But how are we supposed to know that this idea is what makes the idea so compelling? I

---

teleological conception of practical reasons” (TCR), which is what I’ve called “consequentialism.” For him, “consequentialism” is the conjunction of TCR with “moral rationalism,” the view that, if something is morally required, then one has most reason to do it, all things considered.

<sup>11</sup> “According to most consequentializers [the compelling idea] is, roughly, the idea that it is always permissible to bring about the best outcome” (Schroeder 2017: 1477). See also Foot 1983: 198, 1985, Dreier 1993: 24–5, Smith 2003, Thomson 2003: 8, Portmore 2009: 332–33, Schroeder 2007: 279–80, Hurley 2013, 2017: 30.

<sup>12</sup> Hurley (2017) argues that the compelling idea can be interpreted either as: (i) it’s always permissible to do the best *action*, or (ii) it’s always permissible to do the action with the best *outcome*. The first is compelling but not consequentialist; the second is consequentialist but not compelling. I agree with Hurley’s point, but my argument is meant to have even weaker premises. I grant that the “compelling idea,” even interpreted as (ii), might be compelling *if* we don’t consequentialize.

think it helps to start with an opposite question. How could the idea *fail* to be compelling?

Consider what happens when we apply the idea where it doesn't belong—in the realm of attitudes like belief and resentment. The result is a pair of underwhelming ideas:

<i>Epistemic Idea</i>	It is always permissible to adopt a belief if doing so would make things go best.
<i>Emotive Idea</i>	It is always permissible to feel resentment if doing so would make things go best.

Suddenly, the idea that producing the goods must be right does not seem irresistible, not even to some consequentializers.<sup>13</sup>

To illustrate: suppose a despot offers me a fortune to believe that  $2 + 2 = 5$ , or to resent you for the lovely present you just gave me. It doesn't follow that I should believe an obvious falsehood or resent a good deed.<sup>14</sup> At most, as Parfit (2011: 1, 50–1, 420–32) would say, the despot's bribe means that I should *cause* myself to have these attitudes, e.g. by hypnosis.<sup>15</sup>

If the “compelling idea” seems more plausible for action than for belief and emotion, that raises a question. What's the source of the difference? Clearly, there must be an underlying difference in the nature of actions as opposed to these two attitudes. Unlike actions, these attitudes are *not* aimed at changing the world: belief is essentially a matter of fitting one's mind to the truth, and resentment is by nature a response to a wronging. So to recommend resenting someone *because it would make things go well*, or recommend a belief *because that would prevent some nastiness*, is to give the a

---

<sup>13</sup> See e.g. Portmore (2011: 59) on “fittingness reasons” for preference and Parfit (2011: Appendix A) on “object-given reasons” for attitudes. (On Parfit's sympathies towards consequentializing, see his 2017: Chapter 57).

<sup>14</sup> The case of desires is tricky, since they are in a sense directed at an outcome—their content. Still, I don't think that a despot's bribe can be a reason to desire something bad for its own sake.

<sup>15</sup> Rowland 2019. Of course, some think bribes are reasons to believe (e.g. Rinard 2017). But even if this view is defensible, it isn't widely felt to be “compelling.” It would be a huge blow to consequentializers if their “compelling idea” turned out to be on a par with the Epistemic Idea.

reason of the wrong kind—something that isn’t even a candidate reason.<sup>16</sup>

The right kinds of reasons for belief and emotion, I think, are “responsive:” the reasons are given by facts about things as they are, not as they *would* be if one did this or that. Whereas “telic” reasons come from outcomes, responsive reasons come from the *situation*, which consists in facts about things as they stand: who made which promises to whom, who has rights to what, who is under threat, etc.<sup>17</sup>

Now, back to actions, we are ready for the main point. *If we give up Action as Production, we are opening our minds to the possibility of non-telic reasons for action, which aren’t grounded in the value of outcomes.* If some actions are, like beliefs, by nature a kind of response to one’s situation—as retribution is supposed to be to wrongdoing, and respect is to others’ rights—then nothing “compels” us to think of the reasons for these actions as exclusively telic. There will be something else, besides good results, for permissibility to turn on.

That’s why I think the compelling idea relies on Action as Production. If some acts essentially fit the world as is, rather than trying to change it, then it makes perfect sense why optimific changes could be wrong. Unless action is aimed entirely at production, we shouldn’t be “compelled” to think that producing the goods must be permissible, since there may be other considerations—like responding to rights—that are of the right kind to make a moral difference. Without Action as Production, the compelling idea may turn out to be no more compelling for action than for belief and other attitudes.<sup>18</sup>

---

<sup>16</sup> For more on the wrong kind of reasons problem (for buck-passing theories of value), see Rabinowicz and Rønnow-Rasmussen 2004, Hieronymi 2005, Darwall 2006: 66, and Way 2012.

<sup>17</sup> On telic vs. responsive reasons, see [omitted]. See also Broad 1930 and Howard forthcoming on fittingness, Müller 2011 on backward-looking reasons, and Pettit 1989 on honoring value. What unites all these concepts—responding, honoring, looking back, fitting—is the emphasis on acting in light of the given facts of the situation, rather than in light of the prospect of a certain outcome.

<sup>18</sup> Many authors already agree that the compelling idea depends on Action as Production (e.g.

That is my defense of my first premise. It was a bit dense. To make it more intuitive, it might help to imagine a dialogue with my opponent here—namely, the consequentializer who denies Action as Production. This person loves the distinction between killing (wrong!) and letting die (not so wrong!), but they think that the problem with killings is mere intrinsic badness. Suppose we ask them: why aren't you just a deontologist? Why insist that killing is wrong because of its badness, rather than because people have rights against killing (a responsive reason), or because killing expresses heinous disrespect, etc.? It would be strange for them to say: "But those *can't* be reasons! How could anything be a reason if it's not part of an action's outcome?" This just doesn't make sense coming from someone who thinks that there is more to action than producing outcomes. If actions can morally differ in virtue of what they express, or which things they respond to, why can't these factors be relevant to rightness and wrongness directly, as normative reasons? (This question has extra bite if the consequentializer believes in responsive reasons for attitudes. Why not actions, too?) Without Action as Production, consequentialism sounds less like a "compelling idea" than an arbitrary preference.

There are two key points in this argument. The first is that we should spell out the "compelling idea" using the distinction between right and wrong kinds of reasons. To be compelled by consequentialism is to see telic reasons as the only reasons of the right kind for action. (To be merely *convinced* by consequentialism is to think that, once the philosophical dust settles, reasons for action will *turn out* to be telic.) I think this captures the aura of inevitability, not just plausibility, around consequentialism. But—point two—it is a truism that the right kinds of reasons for

---

Scheffler 1985: 414–15, Portmore 2009: 332–33, Hurley 2017: 188–89). But not everyone. Betzler and Schroth (2019: 127–30) say that only agent-neutral consequentialism can be "compelling." My argument doesn't need this strong claim. I can grant that agent-relative consequentialists may have coherent, compelling views, too—so long as they don't consequentialize, and thereby hang on to Action as Production. (Consider, e.g., the egoist view that promoting self-interest is all that matters.)

something depend on what that something *is*. That is why, unless action *is* the production of outcomes, the right kinds of reasons for action do not have to come exclusively from outcomes.

## 6. Why Consequentializing Undercuts

Consequentialism is only compelling given Action as Production. Now I'll argue for:

*Undercut by Consequentializing*

If killing is worse than letting die, then Action as Production is false.

The claim is that the consequentializer, who takes killing to be worse than letting die, is recapturing intuitions but betraying consequentialism.

But *where* exactly is the betrayal? Why can't consequentializers embrace Action as Production? Consequentialists come in so many stripes and dots—hedonistic and pluralistic, agent-relative and agent-neutral—and we are granting that all of these varieties can gel with Action as Production, so long as they don't consequentialize. What makes the consequentializer in particular unable to think of acts in terms of outcomes?

The answer is that the consequentializers are drawing a moral distinction that can't be traced to any difference in outcomes. There is nothing *about* the act of killing that can *make* it worse than letting die, given Action as Production, so long as the narrow outcomes are the same.

Let me lay out my argument more carefully. Premise one is a principle about value:

*Resultance*

If killing is worse than letting die, then something about the killing must make it worse.

(I am talking, as usual, just about the case of killing one to save five vs. letting one die to save five.)

This premise should be uncontroversial. Everyone agrees that moral properties are based in other properties; they are “resultant,” in Ross's phrase (1930: 28). Pain, for example, is worse than pleasure because of a difference in how they feel. But what difference could make killing worse than letting die? Given Action as Production, the only place to look is within the broad outcome. That leaves only two possibilities:

*Only Two Options*

If Action as Production is true, then only two things could make killing worse than letting die: a difference in narrow outcomes or a difference in the acts themselves.

But as the consequentializers themselves seem to assume:

*Same Narrow Outcome*

Killing and letting die have the same narrow outcome.

After all, why else introduce the concept of broad outcomes? But if the narrow outcomes are the same, then there is nothing to make the acts themselves any different, and so there can't be any difference at the level of broad outcomes, either. (I'll defend this in a moment.)

*Narrow Principle*

Given Action as Production, if two acts have the same narrow outcome, then they are intrinsically the same.

Thus consequentializing undercuts Action as Production. If killing is worse than letting die, there has to be something about killing *besides* its outcome that makes it so. And so there must be more to killing than its aiming at an outcome.

I know of two objections to this argument.<sup>19</sup> The first targets the "Narrow Principle."

Look, you are ignoring the importance of switching to the concept of *broad* outcomes. Don't forget: just as the act of running produces the outcome that one runs (Portmore 2011: 56), killings produce killings, and lettings-die produce lettings-die. Surely the consequentializer can use this fact to single out killings as worse, while staying true to Action as Production.

But this is a mistake. Unless there is *already* a difference between killing and letting die, you can't show that two acts are different by pointing out that one produces a killing whereas the other produces a letting die. The "difference" won't make a difference.

To illustrate, imagine an old-school consequentialist who accepts Action as Production but rejects any metaphysical killing/letting-die distinction. (Perhaps they think the difference is merely

---

<sup>19</sup> I owe thanks to an anonymous referee for suggesting both objections, and for pushing me to clarify my argument. (The objections are written in my own words.)

verbal or conceptual.)<sup>20</sup> On their view, killing and letting die have the same broad outcomes, because they have the same narrow outcome, and the acts are themselves the same.

Of course, if we *do* think these acts are different (even when they have the same narrow outcome), then we must *also* believe that they have different broad outcomes. But the difference in acts would have to be prior, due to some other source. Suppose, for example, that we accept an “expressive” conception of action, on which killing differs from letting die in virtue of what the acts express; killing “says” to the victim that they lack a high moral status (Schapiro 2001: 95–98). On this view, killing and letting die have different broad outcomes (they produce different acts). But that is not, and cannot, be the reason *why* they are different acts. The broad outcomes are downstream of a more fundamental difference in expressive content. Without that more basic difference, nothing would be there to make the broad outcomes different, and the metaphysical distinction between killing and letting die would collapse. The same goes for responsive views of action. Acts that differ in how they respond to the situation (e.g. revenge vs. mere harm) will have different broad outcomes, even if the narrow outcomes are the same, but only because there is something other than outcomes that makes the difference.

What if we are overlooking some hidden difference in the broad outcome, something special about killing? Could this be an opening for the consequentializer to bring in the distinction between killing and letting die? That is the idea behind the final objection:

There is always going to be *some* metaphysical difference between killing and letting die. When I kill, typically, at a minimum, I have to move my body (as when I push the man off the bridge). When I let die, typically, I don’t have to lift a finger (as when I watch icily as the man is pushed by the wind). Won’t killing and letting die always be distinguished by some such physical difference?<sup>21</sup>

---

<sup>20</sup> To be sure, traditional consequentialists might well accept a real metaphysical distinction between killing and letting die. They could just say that the difference isn’t morally weighty.

<sup>21</sup> We could spell this out in terms of worlds: the world in which I kill one to save five will be somehow physically different from the world where I let one die and instead save five.

This is a retreat from the consequentializer's "recipe," which takes for granted that she can help herself to opponents' distinctions. Now we have a consequentializer trying to *retrieve* the killing/letting die distinction within the limits of Action as Production.

Does it work? I don't think so, but to explain why, I will need to clarify the objection.

One way to hear the objection is as follows. Killing and letting die have the same narrow outcome, but killing is *intrinsically* different because it involves some further actions, like the moving of my hand towards the big man as I topple him over the footbridge (there is no such movement involved if I let the wind do the toppling).

But this cannot be quite right. If killing involves moving my hand, then the narrow outcome must include something that is missing from the outcome of letting die—*my hand's moving*. This is not something that I do. What I do is *move my hand* (transitive verb), and as a result, *my hand moves* (intransitive).<sup>22</sup> This latter fact is part of the narrow outcome. And indeed, it seems that for any pair of a killing and a letting die, we can point to some such movement that is in the narrow outcome of killing but not of letting die.

Now we can state the above objection more clearly. The question is: since the killing has extra movements in its narrow outcome, why can't the consequentializers use this fact to recover the killing/letting die distinction within the Action as Production framework?

My answer is that killings do not essentially have these extra movements. They do in the actual world, but that is only because killing is not actually, for human beings, a *basic action*—something that we do but not by doing something else. Only supervillains have the power to kill "with their mind." Humans must kill by pulling triggers and pushing men off of bridges. Letting die, even for us, does not require any such means. And so these means contingently allow for Action as

---

<sup>22</sup> See Hornsby 1980 on the distinction between transitive and intransitive uses of "movement."



Production to distinguish killings from lettings-die.

If we are to find a case where killing and letting die really do have the exact same narrow outcome, we will need some thought experiments. Suppose that a supervillain has the power to kill a certain person basically—disappearing them at will, without lifting a finger. Unless he kills the one, five others will die by being disappeared. He saves them by killing the one. Now contrast this with the case where all six are to be disappeared, and the villain can use his powers either to *save* the five or save the one. He saves the five, letting the one die. These actions have the same narrow outcome.<sup>23</sup> Given Action as Production, there can be no difference between them. But deontologists will say that there is a difference—a metaphysical difference giving rise to a moral one—and so we have a case where we see the limits of the view that makes consequentialism so compelling.<sup>24</sup>

I conclude that, if killing is worse than letting die, then Action as Production must be false. The consequentializer is drawing moral distinctions that can't be given a metaphysical basis within Action as Production, no matter how broadly “production” is construed. But without Action as Production, consequentialism isn't compelling; telic reasons needn't be the only reasons of the right kind. That completes my argument. Consequentializing undercuts consequentialism.

## 7. Conclusion

---

<sup>23</sup> To avoid some complications—won't the villain's memory be different, depending on what he does?—we can suppose that the villain disappears the moment he acts.

<sup>24</sup> But won't there be some facts *before* the time of action that distinguish the world where the villain kills from the world where the villain lets die? For example, before the killing, only five were disposed to die, whereas before the letting-die, all six were. Surely this kind of fact can distinguish the broad outcomes of the two actions (if we take “broad outcomes” *very* broadly to include the past—see fns. 5, 9, above). Here I have a new objection. If the feature of the “outcome” that makes a difference is a fact about the past, then the same point could be made equally well using talk of situations and responsive reasons. If you are comfortable appealing to past “outcomes,” you are in effect admitting that situations can make a direct moral difference, and so you cannot claim to be “compelled” to see responsive reasons *based* in situations as being reasons of the wrong kind.

We began with a classic problem for consequentialism, which is that killing is intuitively wrong even in some cases where it has the best outcome (such as Footbridge). The standard line is that this problem is shallow. We can “consequentialize,” insisting that the act of killing is surprisingly bad, and recovering the wrongness of murder by trolley. I have argued that the problem runs deeper.

By consequentializing, we are positing a moral difference between acts that can’t be reduced to prior difference in outcomes. Anyone comfortable with this cannot claim to be “compelled” to see outcomes as the only reasons of the right kind for action. The consequentializers should be open, for example, to the possibility of responsive reasons, like the fact that the person on the bridge has a right not to be thrown off (or expressive reasons, like the fact that killing is disrespectful). If these aren’t reasons of the right kind, then the consequentializer must say *why*—which is no easy take, if they cannot appeal to Action as Production.

How could the consequentializers respond? One option would be to find something other than the act itself to put the badness of killing in—something like the violation of rights. To say this, however, is to admit that the consequentializing “recipe” isn’t a sure thing. It also raises a dilemma. If “the violation of rights” just refers to the violating act, we are back to stuffing badness in actions. But if “violation” is supposed to signify that rights have intrinsic moral *oomph*, so that they are unfit to disregard, we are positing responsive reasons to respect people’s rights and giving up the consequentialist game.

Another option is to tweak the consequentializer’s view of action. Again, there is a dilemmatic tight squeeze. Action as Production is too restrictive for the consequentializers, but if we loosen the view to allow for responsive actions, consequentialism will not seem so compelling; responsive reasons will not seem like reasons of the wrong kind. (*Mutatis mutandis* for expressive

reasons, etc.)<sup>25</sup>

And so we are left with a stark choice. We can have the compelling idea of consequentialism, or we can have the intuitive judgements of commonsense deontology—but not both. In the end, consequentialism is at its most compelling when it lets us get away with murder.<sup>26</sup>

---

<sup>25</sup> I can think of two more options. One is to find a new positive argument for consequentializing (something in the genre of Portmore 2017). Another is to find negative arguments that cast doubt on the deontologist's concepts, e.g. that of a responsive reason (such arguments might be found in the literature on the "paradox of deontology," on which, see Scheffler 1994: Chapter 3 and Chappell 2011.)

<sup>26</sup> My warmest thanks to two anonymous referees for their insightful, charitable, and challenging comments, which led to several improvements. In particular, their comments helped me clarify the role of narrow outcomes in my argument, and explain why I don't think the consequentializer can get away with saying that killing is bad fundamentally. The paper is much better thanks to the referees' suggestions. For other very helpful comments, I would like to thank Nathaniel Baron-Schmitt, Tamar Schapiro, Kieran Setiya, Caspar Hare, Thomas Byrne, Kelly Gaus, Cheryl Abatte, and my audiences at Macquarie University and CU Boulder's Rocky Mountain Philosophy Conference.

## Works Cited

- Betzler, Monika and Schroth, Jörg (2019). "The Good of Consequentialized Deontology," in Christian Seidel (ed.), *Consequentialism: New Directions, New Problems*. Oxford: Oxford University Press: 115–135.
- Broad, C.D. (1930). *Five Types of Ethical Theory*. London: Routledge and Keegan Paul.
- Brown, Campbell (2011). "Consequentialize This," in *Ethics* 121 (4): 749–771.
- Chappell, Timothy (2011). "Intuition, System, and the 'Paradox' of Deontology," in Lawrence Jost and Julian Wuerth (eds.), *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics*. Cambridge: Cambridge University Press, pp. 271–288.
- Cumberland, Richard (1727). *A Treatise of the Laws of Nature*. John Maxwell (trans.). London: R. Phillips. Originally published in Latin in 1672.
- Darwall, Stephen L. (1995). *The British Moralists and the Internal 'Ought': 1640: 1740*. Cambridge: Cambridge University Press.
- (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: Harvard University Press.
- Dreier, James (1993). "Structures of Normative Theories," in *The Monist* 76 (1): 22–40.
- (2011). "In Defense of Consequentializing," in Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 1*. Oxford University Press.
- Ewing, A.C. (1939). "A Suggested Non-Naturalistic Analysis of Good," in *Mind* 48 (189): 1–22.
- (1947). *The Definition of Good*. London: Macmillan.
- Foot, Philippa (1967). "Abortion and the Doctrine of Double Effect," in *Oxford Review* 5: 5–15.
- (1983). "Utilitarianism and the Virtues," in *Mind* 94 (374): 196–209.
- (1985). "Morality, Action, and Outcome," in Ted Honderich (ed.), *Morality and Objectivity: A Tribute to J.L. Mackie*. London: Routledge & Kegan Paul: 23–38.

- Hieronymi, Pamela (2005). "The Wrong Kind of Reason," in *Journal of Philosophy* 102 (9): 437–457.
- Hooker, Brad (2000). *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford: Oxford University Press.
- Hornsby, Jennifer (1980). *Actions*. London: Routledge & Keegan Paul.
- Howard, Chris (forthcoming). "The Fundamentality of Fit," to appear in Mark Timmons (ed.), *Oxford Studies in Normative Ethics, Volume 14*.
- Hurka, Thomas (2014). *British Ethical Theorists from Sidgwick to Ewing*. Oxford: Oxford University Press.
- Hurley, Paul (2013). "Consequentializing and Deontologizing: Clogging the Consequentialist Vacuum," in Mark Timmons (ed.), *Oxford Studies in Normative Ethics* 3: 123–153.
- (2017). "Why Consequentialism's 'Compelling Idea' is Not," in *Social Theory and Practice* 43 (1): 29–54.
- (2019a). "New Consequentialism and the New Doing-Allowing Distinction," in Christian Seidel (ed.), *Consequentialism: New Directions, New Problems*. Oxford: Oxford University Press: 176–97.
- (2019b). "Exiting the Consequentialist Circle: Two Senses of Bringing it About," in *Analytic Philosophy* 60 (2): 130–163.
- Korsgaard, Christine M. (1996). "From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action," in Stephen Engstrom and Jennifer Whiting (eds.), *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*. Cambridge: Cambridge University Press.
- Louise, Jennie (2004). "Relativity of Value and the Consequentialist Umbrella," in *Philosophical Quarterly* 54 (217): 518–536.
- Mill, John Stuart (1861) [2001]. *Utilitarianism*. Second Edition. George Sher (ed.). Indianapolis: Hackett Publishing Company, Inc.

- Moore, George Edward (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- Müller, Anselm Winfried (2011). “Backward-looking Rationality and the Unity of Practical Reason,” in *Essays on Anscombe’s Intention*, ed. Anton Ford, Jennifer Hornsby, and Frederick Stoutland, Cambridge, MA: Harvard University Press: 242–69.
- Nozick, Robert (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Parfit, Derek (2011). *On What Matters, Volumes 1 and 2*. Oxford: Oxford University Press.
- (2017). *On What Matters, Volume 4*. Oxford: Oxford University Press.
- Pettit, Philip (1989). “Consequentialism and Respect for Persons,” in *Ethics* 100: 116–26.
- Portmore, Douglas W. (2007). “Consequentializing Moral Theories,” in *Pacific Philosophical Quarterly* 88: 39–73.
- (2009). “Consequentializing,” in *Philosophy Compass* 4 (2): 329–347.
- (2011). *Commonsense Consequentialism: Wherein Morality Meets Rationality*. New York, USA: Oxford University Press USA.
- (2014). “Replies to Gert, Hurley, and Tenenbaum,” in *Philosophy and Phenomenological Research* 88 (1): 241–255.
- (2017). “Uncertainty, Indeterminacy, and Agent-Centred Constraints,” in *Australasian Journal of Philosophy* 95 (2): 284–298.
- Prichard, H.A. (1912). “Does Moral Philosophy Rest on a Mistake?” in *Mind* 21: 21–37.
- Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni (2004). “The Strike of the Demon: On Fitting Pro-attitudes and Value,” in *Ethics* 114 (3): 391–423.
- Rinard, Susanna (2017). “No Exception for Belief,” in *Philosophy and Phenomenological Research* 94 (1): 121–143.
- Ross, William David (1930). *The Right and the Good*. Oxford: Clarendon Press.
- Rowland, Richard (2019). *The Normative and the Evaluative: The Buck-Passing Account of Value*. Oxford:

Oxford University Press.

Sachs, Benjamin (2010). "Consequentialism's Double-Edged Sword," in *Utilitas* 22 (3): 258–271.

Schapiro, Tamar (2001). "Three Conceptions of Action in Moral Theory," in *Noûs* 35 (1): 93–117.

Scheffler, Samuel (1985). "Agent-Centered Restrictions, Rationality, and the Virtues," in *Mind* 94: 409–19.

———(1994). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Oxford University Press.

Schroeder, Mark (2007). "Teleology, Agent-Relative Value, and 'Good'," in *Ethics* 117 (2): 265–295.

Schroeder, S. Andrew (2017). "Consequentializing and Its Consequences," in *Philosophical Studies* 174 (6): 1475–1497.

Setiya, Kieran (2018). "Must Consequentialists Kill?" in *Journal of Philosophy* 115 (2): 92–105.

Sidgwick, Henry (1874). *The Methods of Ethics*. London: MacMillan, 7<sup>th</sup> Edition.

Sinnott-Armstrong, Walter (2015). "Consequentialism," in *The Stanford Encyclopedia of Philosophy* (Winter 2015 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2015/entries/consequentialism/>](https://plato.stanford.edu/archives/win2015/entries/consequentialism/).

Smith, Michael (2003). "Neutral and Relative Value After Moore," in *Ethics* 113: 576–598.

Tenenbaum, Sergio (2014). "The Perils of Earnest Consequentializing," in *Philosophy and Phenomenological Research* 88 (1): 233–240.

Thomson, Judith Jarvis (1985). "The Trolley Problem," in *The Yale Law Journal* 94 (6): 1395–1415.

———(2003). *Goodness and Advice*. Amy Gutman (ed.). Princeton: Princeton University Press.

Way, Jonathan (2012). "Transmission and the Wrong Kind of Reason," in *Ethics* 122 (3): 489–515

Yetter Chappell, Richard (2019). "Willpower Satisficing," in *Noûs* 53 (2): 251–265.