

HOW (AND HOW NOT) TO DEFEND LESSER-EVIL OPTIONS¹

INTRODUCTION

We sometimes possess *lesser-evil justifications* for harming one or more victims for the sake of saving others. Consider the example *Trolley*:

Five innocent people are trapped in the path of an oncoming trolley. Bystander could turn the trolley onto a sidetrack, away from the five. If he does so, the five will be saved, but, foreseeably, the trolley will hit and kill one innocent person, Workman, who's trapped on the sidetrack.²

It is normally wrong to direct a trolley toward an innocent person, foreseeably killing him. But it is intuitively justified, here. It is the lesser evil, compared to allowing the deaths of five people.

Lesser-evil justifications are conventionally thought to expand agents' permissible options. Bystander may turn the trolley in the circumstances; alternatively, he may do nothing, sparing the one and allowing the five to die.

¹ This paper owes many debts to many people, too many of whom I fear I've forgotten. But I can at least thank Christian Barry, Kimberly Ferzan, Helen Frowe, Thomas Hurka, Renée Jorgensen, Rahul Kumar, Andrew Lister, Daniel Muñoz, Jonathan Parry, Theron Pummer, Daniel Viehoff, Alec Walen, and two anonymous referees for this journal for invaluable written comments or discussion. The paper was supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Grant.

² Here and throughout, I assume other things are equal: Bystander lacks a special relationship with Workman or any of the five; all the potential victims have roughly same amount of good life ahead of them; they all lack dependents, and so forth.

Helen Frowe rejects this conventional line.³ She is the unique nonconsequentialist proponent of the view that Bystander is morally required to turn the trolley, and, more generally, that lesser-evil justifications confer obligations rather than permissions. For Frowe, the class of lesser-evil options is small to non-existent.

This paper will advocate for lesser-evil options, using Frowe as a foil. In the first part of the paper, I shall test the rejection of lesser-evil options against problem cases. I shall concede — after a fight — that one can accommodate the relevant case-based intuitions without them. In the second part of the paper, I shall advance three independent but compatible accounts of lesser-evil options: the parity account, the prerogative account, and the ‘permissible moral mistakes’ account. The parity and prerogative accounts will rescue, from Frowe’s attacks, strategies for defending lesser-evil options that have appeared, albeit in inchoate form, within the literature. The ‘permissible moral mistakes’ account will offer something new.

³ Helen Frowe, “Lesser-Evil Justifications for Harming: Why We Are Required to Turn the Trolley,” *The Philosophical Quarterly* 68 (2018): 460-80.

PRELIMINARIES

I. The Conventional View and Frowe's Challenge

Lesser-evil justifications adjudicate between the harm with which an endangered group is threatened, and the harm their rescuer will inflict on a victim (or victims), just in case she saves the group.⁴ The rescuer has a lesser-evil justification for inflicting harm if:

- (i) the severity of the harm facing the endangered group is sufficiently great, compared to the harm the rescuer would inflict on her victim(s); or
- (ii) the severity of the harm facing the endangered group is at least comparable to the harm the rescuer would inflict on her victim(s), and the endangered group-members sufficiently outnumber the rescuer's victims.

Lesser-evil justifications rely on evaluative judgments about the badness of harm h befalling n numbers of victims, compared to the badness of harm h_2 befalling n_2 numbers of victims. But they are sensitive to deontic considerations, too. These include, first, the differential justifiability of inflicting vs. allowing harm; and second, the differential justifiability of inflicting (or allowing) harm in different *modes of agency*⁵ – using one's victim's body in a way that leads to his harm vs. harming him as a side effect,

⁴My focus is lesser-evil justifications for harming, although presumably there are lesser-evil justifications for contravening other constraints, i.e., promissory obligations. I am grateful to Christian Barry for pointing this out.

⁵ This language originates with Warren Quinn, "Actions, Intentions, and Consequences: The Doctrine of Double Effect," *Philosophy & Public Affairs* 18 (1989): 334-51.

for example.⁶ Other things equal, inflicting harm is more difficult to justify than allowing harm to occur, and inflicting harm through using the victim's body is more difficult to justify than inflicting harm as a side effect.

For a lesser-evil justification to obtain, the presumptive priority of *not harming* over *saving* (not allowing harm) must be overcome. This requires that the severity of the harm facing the endangered group *substantially* exceeds the severity of the harm the rescuer would inflict on her victim(s), or that the former's numbers substantially exceed the latter's. (To overcome the presumptive difference between *harming as means* and *not saving*, the difference in the severity of harm, or numbers of people harmed, must be greater still.)

The intuitive case against lesser-evil options is simple and powerful. As the high-bar setting for lesser-evil justifications evinces, we take deontic constraints against harming innocent people very seriously.⁷ If the eventuation of the harm threatening a group is bad enough to justify contravening a constraint, it should be bad enough to eliminate the option of doing otherwise. It is implausible that the countervailing consideration – namely, the agent's autonomy-interest in retaining the option of *not contravening the relevant constraint* – defeats the good of saving the endangered group.⁸

⁶ There is an intermediate category between harmful using and side-effect harming, namely, eliminative harming (Quinn, [1989]: 344). Such harmings eliminate threats or obstacles others impose.

⁷ Constraints correlate with presumptive moral prohibitions; the more stringent the constraint, the harder to justify the presumptively prohibited act.

⁸ Frowe, (2018): 465, 471-475. My summary negative case against lesser-evil options diverges from Frowe's, but I have sought to preserve the latter's spirit, and its forcefulness. For detailed treatment of Frowe's argument — including its vulnerabilities — see Kerah Gordon-Solmon and Theron Pummer "Lesser-Evil Justifications: A Reply to Frowe," *Law and Philosophy* (forthcoming).

Trolley is the flagship example of Frowe's attack. She argues that, given Bystander's lesser-evil justification for turning the trolley, he *lacks* justification to *refrain* from turning it. His principle reason for not turning the trolley – that it will kill Workman – has already been defeated. Turning the trolley does not incur Bystander material costs. Nor – assuming a normal psychological profile – does it incur him prohibitive psychological costs. He may want not to kill, but he lacks sufficient *moral* reason not to kill: the balance of moral reasons favours killing.⁹ Even if he is committed to non-violence, this is a rare, one-off case, in which he will be insulated from the gruesome realities of the killing; it is implausible that what remains – his commitment to his policy – licenses him to let the five die.¹⁰

This analysis of *Trolley* generalizes. If an agent has a lesser-evil justification for harming a victim to save an endangered group, then the stakes for the group-members are sufficiently high to defeat the constraint against harming the victim. In most circumstances, it is implausible that the *stakes for the agent* will be sufficiently high, compared to those for the group, to license the agent to privilege her own interest, in not saving its members, over theirs, in being saved. If one has a lesser-evil justification for contravening a constraint, almost always, one will be required to do so.

⁹ Frowe, (2018): 470-71.

¹⁰ Frowe, (2018): 471-474.

PART ONE

II. Two Cases

A seemingly promising line of defense for lesser-evil options says that, without them, we cannot support plausible judgements about the following two cases:

Surgery: Surgeon is trying unjustly to kill Victim because he hates Victim and wants her dead. He will succeed unless Victim kills him. If Victim kills Surgeon, the five patients on whom Surgeon is scheduled to operate over the next twenty-four hours will die.¹¹

Surgery 2: Just like *Surgery*, except Defender is the one positioned to kill Surgeon in Victim's defense.

In these cases, like in *Trolley*, five innocent peoples' lives are pitted against one different innocent person's life. This suggests, if Bystander is obligated to turn the trolley, then Victim and Defender, respectively, are obligated to allow Surgeon to kill Victim in *Surgery* and *Surgery 2*. These implications are counter-intuitive.

Frowe maintains, however, that Bystander's obligation to turn the trolley in *Trolley* is compatible with Victim and Defender each having permission defensively to kill Surgeon.

In this section, I shall argue that Frowe's analysis of *Surgery* succeeds, albeit with emendation, but her analysis of *Surgery 2* fails. The next section — section III — will

¹¹ Adapted from Frowe, (2018): 476.

reassess *Surgery 2*. It will argue that a different analysis of the case accommodates our intuitions without relying on lesser-evil options. Part One will thereby end in a draw. For the paper to succeed as advocacy, it must do so on the strength Part Two. (To the reader disinterested in arguments with negative conclusions: feel free to skip ahead.)

Surgery and *Surgery 2*

I assume, with Frowe, that in both *Surgery* cases, Surgeon is *liable* to be killed to save Victim: he has forfeited his right not to be killed to save Victim, in the circumstances. The only relevant considerations to what Victim or Defender may or must do are the harms to Victim and the five. Defensively killing Surgeon would have the unintended, indirect side effect of preventing five people from being saved. Refraining from killing Surgeon (to allow him to save the five) would, as an unintended, agentially mediated side effect, allow Surgeon to kill Victim.

Compare *Surgery* to *Trolley*. In *Trolley*, Bystander can either allow five to die, or kill Workman as a foreseen side effect of saving the five. Grant for the sake of argument that Bystander is morally obligated to turn the trolley, killing Workman. To assess what Victim or Defender may or must do in the *Surgery* cases, we can compare the options in *Trolley* with their counterparts in *Surgery*. (Since the two versions of the case are relevantly similar, I will compare *Trolley* only to *Surgery* [and not to *Surgery 2*].)

Bystander's options in *Trolley*:

Victim's options in *Surgery*:

Turn the trolley	Save 5	Kill 1	Abstain from self-defense	Allow 5 to be saved	Allow 1 to be killed
	Allow 5 to die	Do not kill anyone		Self-defense	Prevent 5 from being saved

The top row of the tables shows the implications of Bystander turning the trolley (the 1st table) and Victim abstaining from self-defense (the 2nd table). The top left quadrants show, respectively, that in turning trolley Bystander saves five, and in abstaining from self-defense, Victim allows five to be saved. In general, the moral reason to allow people to be saved (or not to prevent their being saved¹²) is *at least as strong* as the moral reason to save people.¹³ Thus, Victim's moral reason to allow the five to be saved is at least as strong as Bystander's moral reason to turn the trolley.

¹² That is, not to block the casual sequence that would result in their being saved from running its course.

¹³ That we have stronger moral reason not to prevent a person from being saved than to save a person is controversial. But no-one argues we have stronger moral reason for latter than the former.

Now look at the top right quadrants. These show that in turning the trolley, Bystander kills one, and in abstaining from defense, Victim allows one (herself) to be killed. In general, we have *weaker* presumptive moral reason against allowing people to be killed than against killing people. Thus, Victim's moral reason to save one (herself) is weaker than Bystander's moral reason not to kill one (Workman).

On this analysis, the balance of moral reasons more decisively favours Victim allowing Surgeon to kill her than it favours Bystander turning the trolley. Victim's moral reason to abstain from self-defense is at least as strong Bystander's to turn the trolley, and Victim's countervailing reason is weaker than Bystander's. Victim's moral reasons for abstaining from self-defense thereby defeat her moral reasons for defending herself *by a greater margin* than that by which Bystander's moral reasons for turning the trolley defeat his moral reasons for not turning it. Bystander's (stipulated) obligation to turn the trolley implies a presumptive obligation for Victim to allow Surgeon to kill her.

That Victim is obligated to allow Surgeon to kill her – and that Defender is likewise obligated to allow Surgeon to kill Victim – is both intuitively unacceptable, and, on this analysis, an unavoidable implication of the view that Bystander must turn the trolley.

Frowe rejects the latter claim, on the basis of her different analysis of the case. She says:

It matters that if Victim is undefended in Surgery, she will suffer an unjustified killing. Surgeon is not killing Victim as some unfortunate but unavoidable side

effect of saving the five. Rather, his killing of Victim is wholly gratuitous.

Requiring someone to let herself be subjected to gratuitous violence is very hard to justify because it is an especially egregious form of treating her as a means.

Being harmfully treated as a means for the good of others is worse than being harmed as a foreseen side effect of benefiting others.¹⁴

It is true that Surgeon killing Victim is “gratuitous”: Surgeon killing Victim does not contribute to Surgeon saving the five; nor does it accomplish any other worthwhile goal. But Frowe’s claim that *requiring* Victim to allow Surgeon to kill her gratuitously *treats Victim as a means* is false, on the colloquial understanding of what is it to treat a person as a means.

To see this, contrast *Surgery* with a variation in which Surgeon wants to kill Victim, not out of malice, but because he needs to harvest Victim’s organs in order to save his five patients.¹⁵ Requiring Victim to allow this is to require her passively to operate in an *opportunistic* mode of agency — to treat someone (in this case, herself) as a tool to be used or an opportunity to exploit.¹⁶ Killing another in this mode of agency is maximally difficult to justify; allowing another to die in the same is barely less so. There is no general duty (nor permission) to behave in this way to allow five others to be

¹⁴ Frowe, (2018): 477.

¹⁵ This variation parallels Frowe’s *Bus*, (2018): 477.

¹⁶ Quinn (1989): 344.

saved;¹⁷ hence, Victim has no such duty. (She might well have permission; we have discretion to do certain things to ourselves that we may not do to others without their consent. But, plausibly, where there is no general permission to do something to another without their consent, there is no general duty do it to oneself.)

By contrast, in the actual *Surgery* case, Surgeon killing Victim does not contribute to the five being saved. Likewise, Victim *allowing Surgeon to kill her* does not contribute to the five being saved. Victim's *abstaining from killing Surgeon*, so that Surgeon can save the five, does so contribute. But Victim's allowing Surgeon to kill her is an unintended side effect of her not killing Surgeon. Familiarly, allowing, as an unintended side effect, someone to be killed is distinct from using that person as means.

There is a different sense of 'treating a person as a means,' which Frowe might have in mind, here, namely, that to treat someone as a means is to treat her as lacking the moral status that she in fact has. Treating someone as a means would thus connote treating her as an *object*, rather than treating her specifically as a tool. The claim might be that when Surgeon commits "gratuitous [lethal] violence" against Victim, he treats her *as if* her moral status is low enough to admit such treatment, thereby treating her as a

¹⁷ If Defender were present, it would be wrong for Defender to allow, for the sake of the five patients, Surgeon to kill Victim for her organs. Defender would be passively using Victim as a means of saving the five. See Jeff McMahan "Intention, Permissibility, Terrorism, and War," *Philosophical Perspectives* 23 (2009): 350.

For excellent analysis of additional cases in which the agent allows harm "on the basis of [the victim's] usefulness to others," see Ketan Ramakrishnan, "Treating People as Tools," *Philosophy and Public Affairs* 44 (2016), especially 147-150.

means.¹⁸ To say that Victim must allow Surgeon to do this (for the sake of the five patients) demands Victim's complicity in her mistreatment.

I am sympathetic to this line (although the language of 'treating a person as a means' is infelicitous, here). I believe Victim is entitled to assert her moral status in the face of an intentional, unjustified attempt on her life by a fully culpable assailant,¹⁹ even at the cost of preventing five people from being saved. I thus believe that Victim has an agent-centred permission to defend herself in *Surgery*.

Frowe does not claim, however, that Victim has special permission to save herself, but that *anyone* — Defender in *Surgery 2*, for example — is permitted (or required) to save her. *This* claim is in tension with the claim that Bystander *must* turn the trolley, which cannot be satisfactorily resolved.

Frowe says of *Surgery 2* that "it is proportionate for... a third party to save Victim's life at the cost of preventing the saving of the five," because the harm with which Victim is threatened — i.e., being killed intentionally by a malign assailant, whose act achieves no offsetting good effects — is especially morally bad: much worse than the harm of being killed as a side effect by an agent serving a just cause.²⁰

¹⁸ I suspect, if pressed, Frowe would commit to the former, 'instrumentalizing' interpretation of 'treating persons as means.' It is the interpretation she develops in subsequent work. See Helen Frowe, "The Limited Use View of the Duty to Save," in David Sobel, Peter Valentine *Oxford Studies in Political Philosophy Volume 7* (Oxford Scholarship Online, 2021, doi: 10.1093/oso/9780192897480.003.0003).

¹⁹ This over and above her agent-centred prerogative to favour her own life.

²⁰ Frowe, (2018): 477. Frowe attributes the badness of Victim's killing to its being "unjustified" and "gratuitous". I've replaced Frowe's language for precision. Gratuitous harms can be intentional or accidental; unjustified harms can be gratuitous, or disproportionate to the good in service of which they're inflicted.

This rationale is implausible. For it to succeed, it must be either worse for the victim, or impersonally worse (or both) to be intentionally killed for immoral reasons than foreseeably killed for moral ones. But the differential badness *for the victim* is comparatively small, given that the baseline for comparison is *the badness of premature death*. It is implausible that being intentionally killed for wrongful reasons is at least *five times* worse for the victim than death *simpliciter* would be.

Suppose by ‘morally worse’ Frowe means impersonally worse. I suspect she is tacitly drawing a 1:1 correlation between the *strength of the constraint* against inflicting harm *h* in mode-of-agency *m* on a victim, and the *badness* of harm *h* being inflicted on the victim in mode-of-agency *m*. But this form of consequentializing does not survive scrutiny. Your reasons not to kill are not symmetrical with my reasons to prevent you from killing. Nor is the relative strength of the general reasons we have to save lives plausibly *that* sensitive to variations in the causes and circumstances of the victims’ deaths.^{21, 22}

In sum: Frowe’s defense of the conjunction that (1) Bystander must turn the trolley in *Trolley*, and (2) Defender may (or must) kill Surgeon in *Surgery 2* does not succeed.

Worse, the conjunction has the implausible implication that, if you can prevent one person from being murdered only by killing a different person as a side effect, you *must* do so. On Frowe’s view, saving five is more important (perhaps significantly so)

²¹ Excepting scenarios wherein the rescuer bears responsibility for the victim’s endangerment.

²² McMahan, (2009): 357-358.

than complying with the constraint against killing one person as a side effect. Likewise, preventing one murder is more important (perhaps significantly so) than allowing five people to be saved. Allowing five to be saved is at least as important as saving five. Thus, transitively, preventing one murder would be more important than saving five, and more important, still, than complying with the constraint against killing one as a side effect.²³ I take this to be a *reductio* of the conjunction.

To conclude, if we accept the third-party obligation in *Trolley*, we must accept the counterintuitive third-party prohibition in *Surgery 2*. By Frowe's own lights, this is a sizeable bullet for proponents of a lesser-evil obligation to turn the trolley to bite.

III. What the Preceding Overlooks

The main argument of the previous section concerned *Surgery 2*. It said: five deaths are worse than one; the moral presumption against preventing someone from being saved is at least as strong as the moral presumption against letting someone die; therefore, other things equal, Defender's moral reasons favour letting Victim die over preventing the five patients from being saved. This reveals the need for lesser-evil options. If Bystander *must* turn the trolley, then Defender *must* let Victim die in *Surgery 2*. The latter is unacceptable, so the former must be false.

I used to find this case compelling. But I now believe it relies on a mistake. Namely, it treats Defender as standing in symmetrical relationships with Victim and with the five patients, the only relevant difference being that Defender's options toward Victim

²³ Theron Pummer pointed out that one might resist the appeal to transitivity, but the burden would be on them to explain its failure, here.

are to save her or to let her die, while her options toward the five are to allow them to be saved or to prevent it. Victim, however, has a presumptive claim against Defender to save her. I suspect the five patients do not have equivalent claims against Defender to allow them to be saved.

Consider: each of five patients has a claim against Surgeon to perform a life-saving operation. They have correlative claims against others wrongfully preventing Surgeon from performing those operations. If a wrongful aggressor were lethally to threaten Surgeon prior to the surgeries, the aggressor would wrong not only Surgeon, but also his patients. But in the actual situation, by credibly and culpably threatening to murder Victim such that the only way to save Victim is defensively to kill Surgeon, Surgeon has *forfeited his right not to be defensively killed*. It is dubious that the patients retain their claims against Surgeon's threateners where Surgeon has forfeited this right. The patients' claims against Surgeon's threateners are, to that extent, sensitive to — or, a stronger formulation, contingent on — Surgeon's claims. If the first, weaker formulation is right, the patients' claims against Defender are diminished. If the stronger formulation is right, the five patients lack even presumptive claims that Surgeon not be killed on the strength of the liability justification for killing him. On either formulation, Surgeon's patients differ from standard victims of collateral harm.

By contrast, Victim has an undiminished claim against Defender to save her; by stipulation it would be an easy rescue for Defender to perform. I submit that saving one person who possesses an undiminished presumptive claim against you to save her is more

important morally than allowing five people to be saved who lack such claims against you to do so.

I said the claims of the five patients against Defender are either diminished or obviated; this wants unpacking. If a claim is *diminished*, it carries less weight (or force) against competing claims than it normally would. If a claim is *obviated*, the interest it normally protects still exists — and presumptively still matters — but it carries less weight against competing interests than it would if protected by a claim. (For this reason, I need not choose between the stronger and weaker formulations of two paragraphs back — both have the same upshot for moral deliberations.)

An achievement of this section is that it spotlights an easily overlooked variable in lesser-evil justifications, namely, the presumptive claims of members of endangered groups against potential contributors (active or passive) to their rescues. To the extent that endangered group members lack such claims against certain rescuers, the good to be achieved in saving them must be discounted in those rescuers' lesser-evil calculations.

With respect to the aim of this paper, however I'm back where I started. If the argument of this section succeeds, it reconciles an obligation to turn the trolley in *Trolley* with a permission (perhaps an obligation) to rescue Victim in *Surgery 2*. The argument is not Frowe's — section II's objections to that were decisive — but it supports the conclusion she needs. Thus, judgments about the cases considered here do not advance the case for lesser-evil options.

PART TWO

IV. The Parity Account

Part One attempted a negative case against eliminating lesser-evil options; it did not succeed. From here, the paper will concentrate on the positive case for lesser-evil options by advancing three accounts thereof. These will be independent of but compatible with one another; I believe each has merit. This section will present the first, parity, account; section V will present the second, prerogative, account; section VI will present the third ‘permissible moral mistakes’ account. I will conclude by sketching how they fit together.

In review, lesser-evil justifications adjudicate between the *badness of the harm* threatening an endangered person or group, and the *stringency of the constraint* against harming a (different) person or group.²⁴ The former is sensitive to the degree of harm the group faces, its numbers, etc. The latter is sensitive to the degree of harm the agent would inflict, the mode of agency in which she would inflict it, and so forth.

Lesser-evil justifications thus pose a conundrum, namely, how to adjudicate between evaluative and deontic considerations (that is, the badness of harm *h* vs. the stringency of constraint *c*). I propose reconciling the two by focusing on the *comparative moral importance* of (1) complying with the constraint against harming one’s potential victim(s), and (2) the good one would achieve (or bad one would prevent) by saving the endangered group. There is a lesser-evil justification for harming just in case complying

²⁴ I am indebted here to Jonathan Quong, “Proportionality, Liability, and Defensive Harm” *Philosophy and Public Affairs* 43 (2015): 144-73.

with the constraint is not more important than the good of saving the group.

One might counter-propose that you have a lesser-evil justification only when saving the group is *more* important than complying with the constraint against harming the victim – perhaps only when it is much more important. Just as the degree of harm threatening the group must be much greater than the degree of harm you would inflict on your victim²⁵ to justify the latter as the lesser evil, the importance of saving the group must be much greater than the importance of complying with the constraint against harming your victim to do so.

But the counter-proposal prompts the question of why, having posited that complying with the constraint is *not more important than* the good of saving the group, we can re-appeal to the importance of complying with the constraint to tip the scales. I doubt there is a satisfactory answer: the importance of complying with the constraint is already fully accounted for; to re-appeal to it as a tie-breaker would be to count it twice.

The claim ‘X is not more important Y’ does not entail that Y is more important X. X and Y could be equally important, or they could be on par.²⁶ The notion of parity yields a promising account of lesser-evil permissions.²⁷

Frowe grants a small subset of lesser-evil permissions, consisting of cases in which complying-with-the-constraint and saving-the-group are morally on a par.

²⁵ ... or the number of third parties must be much greater than the number of your potential victims....

²⁶ I am relying on Ruth Chang’s canonical discussion of parity in “The Possibility of Parity,” *Ethics* 112 (2002): 659-88.

²⁷ I am not the only to think so: for an independently developed parity account, see Jonathan Quong “The Permissibility of Lesser Evil” (ms.).

Suppose, she says, complying with the constraint against killing one person as a side effect were on par with the good of saving five lives. The agent would be permitted either to contravene the constraint, or to allow the five to die. But in that case, complying with the constraint against killing one as a side effect would be *more* important than the good of saving four lives (and *less* important than the good of saving six lives). Saving five is more important than saving four (and less important than saving six); so the constraint against killing one as a side effect cannot be on par with saving five, *and* on par with saving four, or six. If the agent is permitted to save five in *Trolley*, she is forbidden to save four (or fewer), and obligated to save six (or more).²⁸

But we can reject Frowe's reasoning. It is a feature of parity relations that small improvements or diminishments to one member of a pair does not upend the relation. The question "Is Shakespeare greater than Mozart?" would not be resolved by unearthing a lost play, or even a volume of lost plays. Nor is it problematic, generally, for A to be on par with each of B and C, but for B and C *not* to be on par with each other. This can occur if, for example, B and C are qualitatively similar to each other, but dissimilar from A.

Thus, we should not take for granted that adding or subtracting one life will always upend parity relations between constraints against harming and numbers of lives

²⁸ Frowe, (2018): 462. Seth Lazar and Peter Graham assert something similar in "Deontological Decision Theory and Lesser-Evil Options" (*Synthese* [2019], <https://doi.org/10.1007/s11229-019-02496-z>), but with different emphasis. Frowe uses 'on a par' synonymously with 'equal.' For example, she says: "If the moral reasons against harming Workman are exactly balanced by the moral reasons to save five, such that morality is indifferent between harming and not harming, the balance must tip in favour of harming Workman when one will thereby save six;" and "In parity cases... there is equal reason to divert or refrain." ([2018]: 462; 467.) By contrast, Lazar and Graham use 'parity' in Chang's sense, but they claim the difference made by saving one additional person's life is massive enough always to upend it.

saved. This is especially so in the context of a literature rife with rescue cases that involve billions of lives.²⁹ Complying with the constraint against killing one person as a side effect might well be on par with a range of numbers of lives saved, i.e., every number between five and fifteen lives.³⁰ This admits a reasonable scope for lesser-evil permissions.³¹

Some nonconsequentialists, however, might find this scope too narrow. They might believe one is permitted not to kill a person as a side effect, even at the cost of allowing upwards of fifteen people to die.³² Readers who hold this view, but who otherwise find the parity account attractive, would do well to consider endorsing it as one part of a heterogenous account. The next two sections can be read as offering additional parts. I shall present the latter in their freestanding (modular) forms, first, and conjoin them with the parity account in the paper's conclusion.

V. The Prerogative Account

The second account of lesser-evil permissions affirms an *agent-centred prerogative* to comply with constraints over inflicting lesser-evil harms in a subset of cases.

Prefatorily, a note about nomenclature. I'll refer in what follows to a person's *prerogative-protected interests*. By this I mean interests that (i) carry weight or force

²⁹ See, for example, Joe Horton "Aggregation, Risk, and *Reductio*," *Ethics* 130 (2020): 514-29.

³⁰This prompts the question of how to adjudicate between incremental vs. proportionate differences in numbers of lives, which this paper does not address.

³¹ For complimentary treatment of the parity account, see Gordon-Solmon and Pummer (forthcoming).

³² I am grateful to an anonymous referee for pressing this point.

against a person's moral reasons,³³ and (ii) are correspondingly amplified (or fortified) by a prerogative. Prerogative-protected interests are — analytically — eligible to compete with persons' moral reasons; it is an open question which will outperform the other in any particular case.

At first glance, appealing to prerogatives is not a promising route to lesser-evil options. Suppose that, in *Trolley*, the constraint-based reason not to turn the trolley is defeated by the reason to prevent the greater evil. The defeated constraint — against side-effect killing — is stringent. Proponents of the prerogative account are challenged to identify a prerogative-protected interest of Bystander's that is significant enough to outperform said constraint in a matchup against its defeater. Frowe claims Bystander has such an interest only if preventing the evil would “drastically restrict the sphere in which [he] gets to shape [his] own life.”³⁴

But this framing is rigged against the prerogative account. It treats (i) the constraint against killing one, and (ii) the agent's prerogative-protected interest in not inflicting the lesser-evil harm, as sequential competitors against the reason to save the five. In head-to-head competition, Bystander's prerogative-protected interest (however it is specified) does not plausibly fare better against saving the five in *Trolley* than the constraint against killing one does. But if we consider the constraint and prerogative as co-competitors rather than sequential competitors, factoring in prerogatives might well reverse the outcome.

³³ One's interest in experiencing sadistic pleasure, for example, has no force at all against one's moral reasons not to inflict pain on others.

³⁴ Frowe (2018): 472.

In review, when contravening a constraint for the sake of preventing the greater evil is justified, preventing the greater evil is at least as important as complying with the constraint. In a subset of cases, preventing the evil and complying with the constraint are on par. (The size of this subset was the subject of the previous section.) In the rest of the cases, preventing the evil is more important than complying with the constraint. If it is more important, then it is presumptively obligatory. On Frowe's view, preventing the evil becomes supererogatory just in case the agent's prerogative-protected interest is important enough *compared to the importance of preventing the greater evil*.

Contra Frowe, the prerogative-protected interest need not outstrip the importance of preventing the greater evil: it need only outstrip *the difference in importance* between preventing the evil and complying with the (defeated) constraint.³⁵ Preventing the evil can be much more important than complying with a constraint, or only slightly more important than doing so. Preventing the deaths of 100 children overwhelmingly defeats the constraint against twisting one child's arm. Perhaps the former also defeats complying with the constraint against killing twenty-one children, but it would do so by a narrower margin.

The margin of victory matters. In order for (i) the moral reason to comply with the constraint and (ii) the agent's prerogative jointly to defeat the moral reason to prevent the greater evil, the agent's prerogative-protected interests need only be substantial enough to cover the margin. Where the agent can save 100 children by twisting one child's arm, the

³⁵ Alec Walen and David Wasserman suggest this in "Agents, Impartiality and the Priority of Claims Over Duties: Diagnosing Why Thomson Still Gets the Trolley Problem Wrong by Appeal to the 'Mechanics of Claims,'" *Journal of Moral Philosophy* 9 (2012): 554.

agent's prerogative-protected interest must be massive to release the agent from the obligation to twist the arm. By contrast, where to save 100 children the agent must kill twenty-one children, her lesser prerogative-protected interest(s) can change the normative valence of the act from obligatory to permissible.

Prerogatives can most plausibly do this work when not only the margin is narrow, but also the agent has a non-trivial interest in complying with the constraint over promoting the good. In *Trolley*, if we attribute to Bystander such an interest — i.e., a strong aversion to turning the trolley, which stems from a personal commitment to nonviolence — his prerogative-protected interest, together with the constraint against killing, plausibly defeat his moral reasons prevent the greater evil. (This example is problematic; I'll address why two paragraphs down.)

The drawback of the prerogative account is that it is unlikely to support *general* lesser-evil permissions of the sort being sought. Normally, in *Trolley*, Bystander has no substantial interest at stake. He is a generic agent — an everyagent. Accordingly, he has no personal connection to Workman or the five, nor does he have a psychology that makes turning the trolley unusually difficult or costly for him. Thusly characterized, Bystander has no interest I can see that is eligible to compete against doing what he has most moral reason to do. I submit that Bystander-as-everyagent has no prerogative-protected interest in not turning the trolley.

Furthermore — contrary to what I said two paragraphs back — personal moral commitments do not plausibly ground prerogative-protected interests in not doing what the balance of moral reasons says one should. Suppose Bystander's commitment to

nonviolence is predicated on the belief that *inflicting harm is never justified* (or a more sophisticated version of the same). By hypothesis, the belief is false. It is dubious that persons have prerogative-protected interests, generally, in *complying with false moral beliefs*.

This significantly reduces the range of interests agents can cite to release themselves from presumptive lesser-evil obligations. But some such interests remain. An agent might prefer complying with a constraint over preventing the greater evil, not due to misconstruing her agent-neutral reasons, but for agent-relative reasons (of the form ‘*I should do x, given my commitments [or experiences, etc.]*’). The objection I have raised against grounding a prerogative in the former need not apply to grounding one in the latter. If, for example, Bystander’s hypothetical commitment to nonviolence were a pillar of his religious practice — as opposed to being, in the first instance, a moral conviction — it would more plausibly warrant a prerogative’s protection. (I assume a religious commitment to practicing nonviolence oneself is compatible with the belief that persons can permissibly commit violent acts. This is of a piece with [for example] my keeping kosher without believing you should, too.)

I offer the example of a religious commitment because it is a plausible basis for a personal commitment to nonviolence. But I intend it as a token of a broader type, namely, commitments to personal projects that constitute one’s integrity (in Bernard Williams’s sense). To see the contrast between appealing to a *moral* conviction and appealing to a William’s-style commitment, imagine Bystander’s eventual response if pressed upon to turn the trolley. The former says: “But it would be wrong!” (which

again, by hypothesis, is mistaken); the latter says: “But it contravenes who I am!” The latter statement invokes a prerogative-protected interest; the former does not.³⁶

In this section, I advanced a prerogative account of lesser-evil obligations. I claimed that agent’s lesser prerogative-protected interests can work in tandem with constraints to license agents not to prevent the greater evil. But the class of interests that can do this work is limited: interests in complying with an agent’s false moral beliefs are excluded from a prerogative’s protection. Accordingly, we cannot derive *general* lesser-evil permissions via appeal to prerogatives, but only individual permissions, predicated on agents’ personal commitments, histories, relationships, etc.

VI. The ‘Permissible Moral Mistakes’ Account

The third account of lesser-evil permissions will likely be controversial; I offer it more tentatively than I do the preceding two. It proposes that sometimes we are permitted to guide our conduct by the weaker moral reason rather than the stronger one, even without a prerogative-protected interest at stake, *just on the strength of the weaker moral reason*. That is, we are sometimes morally permitted to do what the balance of moral reasons says we should not, just on the strength of the defeated moral reason(s).

³⁶ The claim that prerogative-protected interests and constraints sometimes co-compete against moral reasons to prevent evil prompts the following question: can prerogative-protected interests and moral reasons to prevent evil likewise co-compete against constraints? Suppose, for example, a dedicated effective altruist finds himself in a trolley-type scenario in which the constraint against harming *narrowly defeats* the moral reasons to prevent evil. Could the effective altruist claim a prerogative-protected interest in preventing evil, and thereby special permission to contravene the constraint?

My reflexive answer is “no”: I do not believe prerogative-protected interests carry the kind of weight against constraints that they carry against moral reasons to prevent evil. But proper investigation into the matter is outside the scope of this paper. (I am grateful to an anonymous referee for raising the question.)

The structure of this claim borrows from a way of thinking about prerogatives. Per the previous section, we have prerogative-protected interests that can compete with our moral reasons for acting. When the former are significant enough to be undefeated by the latter, we are licensed to contravene our moral reasons. In a simple example, I have moral reasons to donate all of my savings to effective aid organizations; but my prerogative-protected interest in keeping my savings is strong enough to be undefeated by those reasons. Donating all of my savings would therefore be supererogatory. (Sacrificing an undefeated prerogative-protected interest to do what the balance of moral reasons says one should *just is* supererogation.)

Elizabeth Harman supplements this account of the relationship between persons' prerogative-protected interests and moral reasons with a category she calls "morally permissible moral mistakes."³⁷ Paraphrasing her view, persons *sometimes* retain the prerogative to contravene their moral reasons *even though their prerogative-protected interests are defeated by their moral reasons*. For example, I have moral reason to donate 15% of this month's salary to life-saving charities; my prerogative-protected interest in not doing so is weak enough to be defeated by that moral reason. It might nevertheless be my prerogative not to donate the money. If I exercise that prerogative, I make a permissible moral mistake.

Harman's crucial insight is that we can retain permissions to contravene our moral reasons for the sake of our defeated prerogative-protected interests *so long as the margin of defeat is sufficiently narrow*. I propose that this structure generalizes; that this

³⁷ Elizabeth Harman "Morally Permissible Moral Mistakes," *Ethics* 126 (2016): 366-93. The language of prerogatives is mine, not Harman's.

particular interaction between moral reasons and prerogatives is a token of a broader type.

What I have in mind is as follows.³⁸ Harman's proposal rings plausible because moral reasons and prerogative-protected interests are *incommensurable*.³⁹ It is a marker of pairs of worthwhile incommensurable alternatives that when an agent sacrifices one for the other, she incurs a loss that is not fully compensated. This is so even when the chosen alternative is superior to the unchosen one. In a simple example, I might choose a pint of strawberries over a cookie, endorse the choice as optimal, and still have grounds to feel slightly regretful about not choosing the cookie, instead. By contrast, when choosing between two *commensurable* options, there is no uncompensated loss if I choose the superior one. If I choose the plate with two cookies over the plate with one cookie, I have no grounds for regret about my pick. With commensurable options, there's no rationale for choosing the inferior over the superior. Such a choice would be unreasonable.

With incommensurable options, there is a rationale for making the inferior choice: inferiority need not imply unreasonableness. The caveat is that the greater the margin of superiority, the lesser the comparative significance of the uncompensated loss, and the lesser the appropriate degree of regret. If I choose a third option — a year's supply of ice

³⁸ I am indebted to Daniel Muñoz for discussion of this material.

³⁹ Ruth Chang's account of commensurability and incommensurability — in "Parity, Imprecise Comparability, and the Repugnant Conclusion," *Theoria* 82 (2016): 182-214 — underpins how I'm using the terms, here.

cream — regret over the cookie or strawberries (beyond a passing flicker) would be misplaced. The inferior option stops being reasonable when it is sufficiently outclassed.

The preceding three paragraphs support the following claim. When choosing between two incommensurable alternatives, one of which is superior to the other by a narrow margin, an agent can have *sufficient* reason to choose the inferior option. Or, an agent has discretion when choosing between two reasonable options. This is so even though, in a relevant sense, the agent *should* choose the superior option, as directed by the balance of reasons.

I propose that, just as moral reasons and prerogative-protected interests are incommensurable with one another, different *types* of moral reasons — deontic reasons (to comply with constraints) and telic reasons (to promote the good or prevent evil), respectively — are incommensurable with one another.⁴⁰ Accordingly, when an agent's moral reasons to prevent evil narrowly defeat her moral reasons to comply with a competing constraint, she retains *sufficient* moral reason to comply with the defeated constraint.⁴¹ The agent has discretion, in such cases, to choose between complying with the constraint and preventing the evil. This is so even if she has no prerogative-protected interest at stake.

⁴⁰ One might find this proposal controversial. But it was implicit in section IV's parity account, since parity implies incommensurability. To the extent that one finds the parity account plausible, one should find the incommensurability of telic and deontic moral reasons plausible. For detailed discussion of the relationship between parity and incommensurability, see Chang (2016).

⁴¹ This analysis suggests that when the agent's moral reasons to comply with constraints narrowly defeat her reasons to promote the good, she is nevertheless permitted to violate the constraint. I'm not sure whether this seeming implication is a bug or a feature of what I'm proposing.

This form of analysis applies to *Trolley* as follows. One can grant that the balance of moral reasons narrowly favours turning the trolley, and that Bystander has no prerogative-protected interest in not turning it, but maintain that Bystander is *permitted* not to turn it, just on the strength of his weaker, deontic, reason. Not turning the trolley would be a permissible moral mistake.

The advantage of this account is that it does a nice job with considered judgments about the *Trolley*. I would speculate that, on reflection, many will (i) believe that the balance of moral reasons favours turning the trolley, (ii) deny that Bystander, generically characterized, has a prerogative-protected interest in *not* turning the trolley, and (iii) believe that Bystander is permitted not to turn the trolley. The ‘permissible moral mistakes’ account does an elegant job of reconciling these beliefs; I can see no other way to do so.⁴²

The disadvantage of this account is that it relies on undefended controversial phenomena. Delivering such a defense is outside the scope of this paper.

CONCLUSION

The aim of this paper was to defend lesser-evil options. Part One assessed whether we could support considered judgements about cases without them; it conceded, based on the cases scrutinized, that we can. Part One thus failed as advocacy. Its principal contribution was, instead, to advance our understanding of the factors (and interactions

⁴² Harman (2016) likewise defends the category of morally permissible moral mistakes on the strength of its plausible implications for cases.

among them) to which lesser-evil justifications are sensitive. In particular, it identified a previously overlooked mechanism by which victims can lose their claims to aid, either in full or in part, and the effect on lesser-evil calculations when claims are thusly lost.

Part Two responded to the critical case against lesser-evil options via three positive accounts. Each of these is plausible in its own right, and while each is independent of the others, they are compatible.

Let me end by summarizing how they fit together. The parity account says that one has a lesser-evil option to inflict harm when neither complying with the constraint against inflicting said harm, nor preventing the greater evil, is more important than the other.

The prerogative account says that sometimes, preventing the greater evil is *more important* than complying with the constraint, but the combined strength of (i) the constraint and (ii) the agent's prerogative-protected interest in complying with it, permit favouring the constraint. Lastly, the 'permissible moral mistakes' account says that when (i) the moral reason to prevent the greater evil defeats the moral reason to comply with the constraint, and (ii) the agent has no prerogative-protected interest at stake, the agent is nonetheless permitted to comply with the defeated constraint *just in case* the margin of defeat is sufficiently narrow.

Each of the three account thus demarcates a different candidate region of lesser-evil options, and defends its territory with (at least some) different premises. Since there is no reason to assume that all lesser-evil-options cases are homogenous, one can endorse, without tension, all three accounts, or any subset.

Someone who favours a wider, rather than narrower, range of lesser-evil options would do well to accept all three. Together, they support a broader range than any one of them does alone (or two do as pair).⁴³ Suppose, as I suggested earlier, the parity account supports a lesser-evil option to kill a person as a side effect of saving any number between five and fifteen lives. For the subset of persons with sufficiently strong prerogative-protected interests in complying with the constraint against killing, the prerogative account raises the upper limit of that range: such individuals might retain their option not to kill at the cost of up to twenty lives. The ‘permissible moral mistakes’ account raises the upper limit even higher: say, to twenty-five lives for those without the relevant prerogatives, and higher still for those who have them. What matters for the present purpose are not the numbers, which are merely placeholders: it is how the three accounts, cooperating, deliver an expanded set of lesser-evil options.

REFERENCES

Chang, Ruth. 2002. “The Possibility of Parity.” *Ethics* 112: 659-88

Chang, Ruth. 2016. “Parity, Imprecise Comparability, and the Repugnant Conclusion,” *Theoria* 82: 182-214.

Frowe, Helen. 2018. “Lesser-Evil Justifications for Harming: Why We Are Required to Turn the Trolley.” *The Philosophical Quarterly* 68: 460-80.

Frowe, Helen. 2021. “The Limited Use View of the Duty to Save.” In David Sobel, Peter Valentine (Eds.), *Oxford Studies in Political Philosophy Volume 7*. Oxford Scholarship Online. Doi: 10.1093/oso/9780192897480.003.0003

⁴³ The substance of this paragraph is indebted to an anonymous referee.

Gordon-Solmon, Kerah and Pummer, Theron. Forthcoming. "Lesser-Evil Justifications: A Reply to Frowe." *Law and Philosophy*.

Harman, Elizabeth. 2016. "Morally Permissible Moral Mistakes," *Ethics* 126: 366-93

Horton, Joe. 2020. "Aggregation, Risk, and *Reductio*." *Ethics* 130: 514-29.

Lazar, Seth and Graham, Peter. 2019 "Deontological Decision Theory and Lesser-Evil Options." *Synthese* <https://doi.org/10.1007/s11229-019-02496-z>.

McMahan, Jeff. 2009. "Intention, Permissibility, Terrorism, and War." *Philosophical Perspectives* 23: 345-72.

Quinn, Warren. 1989. "Actions, Intentions, and Consequences: The Doctrine of Double Effect." *Philosophy & Public Affairs* 18: 334-51.

Quong, Jonathan. 2015. "Proportionality, Liability, and Defensive Harm." *Philosophy & Public Affairs* 43: 144-73.

Quong, Jonathan. Manuscript. "The Permissibility of Lesser Evil"

Ramakrishnan, Ketan. 2016. "Treating People as Tools," *Philosophy & Public Affairs* 44: 133-165.

Walén, Alec and Wasserman, David. 2012. "Agents, Impartiality and the Priority of Claims Over Duties: Diagnosing Why Thomson Still Gets the Trolley Problem Wrong by Appeal to the "Mechanics of Claims."" *Journal of Moral Philosophy* 9: 545-71.