The ethics of harming
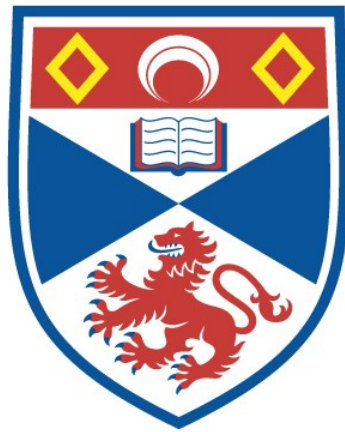
Joel Joseph

A thesis submitted for the degree of PhD
at the
University of St Andrews

2024

Candidate's declaration

I, Joel Joseph, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 40,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2020.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date   19/04/2024   Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date   19/04/2024   Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Joel Joseph, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date   19/04/2024   Signature of candidate

Date   19/04/2024   Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Joel Joseph, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date   19/04/2024   Signature of candidate

Abstract

Commonsense morality rejects an unqualified requirement to promote the overall good. *Moral constraints* prohibit you from bringing about the overall good. For example, you seem morally required to save five people over one. But you seem prohibited from hurling an innocent bystander onto a live grenade that is going to kill five innocent strangers, thereby killing the bystander while shielding the five from the blast. This dissertation provides a defence of moral constraints and explores their practical implications.

Contents

Acknowledgments

My greatest debt is to Theron Pummer. I have benefited tremendously from working under Theron's supervision over the past four years. Theron has not only provided extremely detailed and incisive feedback on all of my work on a regular basis, but he has been also a very encouraging and supportive mentor. Theron is a gifted philosopher and teacher. I could not have asked for a better supervisor.

I am also grateful to my secondary supervisor, Simon James Hope, and to the rest of the St Andrews Philosophy department. I also received generous funding from the *Philosophical Quarterly* Doctoral Scholarship these past four years for which I am very thankful.

I also owe a special mention to my past lecturers at UCL during my MA and at the University of Warwick during my undergraduate for nurturing my interest in philosophy over the years.

This dissertation would also have not been possible without support from my parents, my family and friends, my partner enduring endless trolley cases, and my dog, Plato.

Introduction

Commonsense morality rejects an unqualified requirement to promote the overall good. *Moral constraints* prohibit you from bringing about the overall good. For example, you seem morally required to save five people over one. But you seem prohibited from hurling an innocent bystander onto a live grenade that is going to kill five innocent strangers, thereby killing the bystander while shielding the five from the blast.

Although they seem plausible in certain cases, many argue that moral constraints have counterintuitive implications in other cases. Accordingly, many think we should reject them. This conclusion has considerable practical significance.

For example, it says that there is no sharp moral distinction between failing to donate to the global poor and killing them for financial gain, or between bombing an enemy munitions factory when this will kill nearby civilians and terror bombing them to secure an enemy surrender. So without moral constraints, it seems that we must radically revise many commonplace moral practices.

This dissertation is a collection of four independent essays on *the ethics of harming* from a deontological perspective.

Chapter I addresses the problem of secondary permissibility. You seem permitted to turn a trolley from five onto one, thereby killing her. And you seem prohibited from hurling one into a trolley to stop it from hitting five, thereby paralysing her. Yet if you can either turn or hurl, you are permitted to hurl and prohibited from turning. This paper offers a novel account of secondary permissibility. Whereas hurling when it is your only option of saving the five involves *using a morally objectionable factor*, hurling when you can also turn does not.

Chapter II addresses the problem of killing innocent threats. You are permitted to kill an innocent person to save yourself in some cases but not permitted to do so in others. What explains the moral asymmetry between these cases? This paper offers a novel account of your

permission to kill an innocent threat that provides a unified explanation of our intuitions in cases. Whereas killing an innocent bystander involves *harmfully treating them as a means*, killing an innocent threat does not.

Chapter III concerns the ethics of rescuing prey animals from predators in the wild. Many argue that you are prohibited from interfering with predation on a wide scale. This is because given our ignorance of the long term ecological effects, interfering with predation will probably make things worse overall. But even if it will certainly prove disastrous, this paper argues that you are still morally required to intervene. This is because interfering with predation is morally akin to certain cases in which you seem required to rescue despite this having harmful consequences overall.

Chapter IV responds to a recent set of challenges against the moral constraint against doing harm. Jason Hanna has recently argued that this constraint has counterintuitive implications in cases that involve *undoing one's past behaviour*. This paper argues that Hanna's challenges are unsuccessful.

Overall, I provide a defence of moral constraints and explore their practical implications.

Chapter I: The Problem of Secondary Permissibility

    I.   **Introduction**

The following set of cases raise a puzzle.

> *Turn*. A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by turning the trolley onto a side-track, where it will hit and kill an innocent bystander, *A*.

> *Hurl*. A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by hurling A into the trolley, thereby paralysing her but stopping the trolley.

> *TurnHurl*. A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by either (a) turning the trolley onto a side-track, where it will hit and kill A or (b) hurling A into the trolley, thereby paralysing her but stopping the trolley.

Here is the problem. You seem permitted to turn the trolley onto A in Turn and prohibited from hurling her into the trolley in Hurl. Yet in TurnHurl, you seem prohibited from turning the trolley onto A and permitted to hurl her into the trolley.

But how does turning go from permissible to impermissible, and hurling from impermissible to permissible, when both options are available in TurnHurl? How does hurling become 'secondarily' permissible when associated with the otherwise permissible option of turning? This is the *problem of secondary permissibility*.[1] In this paper, I provide a novel account of secondary permissibility.

In Section II, I argue that we should reject existing explanations of secondary permissibility. In Section III, I offer an alternative account of secondary permissibility. In particular, I argue

---

[1] Kamm (1996, 2007).

that saving the five involves *using a morally objectionable factor* in Hurl but not in TurnHurl. In Section IV, I defend my account of secondary permissibility against some objections.

II.   Against Existing Explanations of Secondary Permissibility

Existing explanations of secondary permissibility face two problems.

Here is the first problem. Existing explanations of secondary permissibility posit an *additional* factor to help explain the permission to hurl in TurnHurl that plays no additional explanatory role in Turn.

For example, some say that you are permitted to hurl in TurnHurl because A *contrastively consents* to being hurled *rather than* being turned onto.[2] But A does not contrastively consent to anything in Turn. Others says that you are permitted to hurl in TurnHurl because A is *morally susceptible* to being killed, or less, by you to save the five.[3] This is a *sui generis* moral relation that holds when you can save the five by proportionately killing A (i.e., by turning). It is true that A is morally susceptible to being killed, or less, by you to save the five in Turn. But positing this sui generis moral relation in this case makes no moral difference to your permission to turn. We could explain your permission to turn without reference to this factor.

It would be more parsimonious if we could accommodate the permission to turn in Turn and the permission to hurl TurnHurl without invoking any additional factors beyond what is minimally needed to distinguish morally between Turn and Hurl. Invoking an additional factor to help explain TurnHurl that does no additional explanatory work in Turn would also not capture the sense some might have that you are permitted to turn in Turn *on the same grounds* that you are permitted to hurl in TurnHurl. So it is worthwhile considering whether there is an

---

[2] Pummer (2023).

[3] Graham (2020).

alternative account of secondary permissibility that does not invoke any additional factors beyond what is minimally needed to distinguish morally between Turn and Hurl.

Here is the second problem. Consider

> *TurnHurl (DifferentPeople)*. A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by either (a) hurling A into the trolley, thereby paralysing her but stopping the trolley or (b) turning the trolley onto a side-track, where it will hit and kill a *different bystander*, *B*.

Some find it intuitively plausible that you are permitted to hurl in this case.[4] However, existing explanations of secondary permissibility cannot accommodate this intuition. This is because these views say that hurling A is permissible in TurnHurl only if A is the same individual that you can turn onto. For example, A can contrastively consent to being hurled *rather than* turned onto only if you can turn onto her. And A is morally susceptible to being killed, or less, by you to save the five only if you can proportionately kill her (i.e., by turning onto her). But in TurnHurl (DifferentPeople), there is no option to turn onto A. You can only turn onto B.

If the intuition that you are permitted to hurl in TurnHurl (DifferentPeople) is correct, this suggests that it is not the fact that A is the same individual that you can hurl or turn onto that helps explain the permission to hurl in TurnHurl.

One could argue that TurnHurl (DifferentPeople) does not count against existing explanations of secondary permissibility. This is because you might be permitted to hurl in TurnHurl (DifferentPeople) for a different reason than why you are permitted to hurl in TurnHurl.

But it would be more parsimonious if we could accommodate the permission to hurl in both TurnHurl and TurnHurl (DifferentPeople) by appealing to the same explanation in each case.

---

[4] For example, Unger (1992), Locke (2021).

Appealing to different explanations in TurnHurl and in TurnHurl (DifferentPeople) would also not capture the sense some might have that you are permitted to hurl *on the same grounds* in both cases. So it is worthwhile considering whether there is an alternative account of secondary permissibility that can also accommodate TurnHurl (DifferentPeople).

I have just raised two problems for existing explanations of secondary permissibility. In the next section, I will offer an alternative account of secondary permissibility that avoids these problems.

III.   The Principle of Secondary Permissibility

I will first outline my account of secondary permissibility, before arguing that we should accept it over its alternatives.

A. The Account

Let us build up to my account of secondary permissibility.

Many argue that you are prohibited from hurling in Hurl because you could not save the five *if A were absent*. This is because you could not hurl if A were absent, and there is no other option of saving the five aside from hurling. A is therefore useful to saving the five and so hurling involves harmfully *using her* to save them.

By contrast, this explanation does not prohibit you from turning in Turn. This is because you could save the five by turning even in A's absence in this case. Accordingly, A is not useful to saving the five and so turning does not involve harmfully using her in Turn. Many argue that this explains why you are permitted to turn in Turn yet prohibited from hurling in Hurl.[5]

It might seem that we could appeal to this explanation to accommodate TurnHurl. For just like in Turn, you could save the five even in A's absence in TurnHurl. This is because you

_____

[5] For example, Quinn (1989), Frowe (2014), Quong (2020).

could save the five by turning instead of hurling. Similar remarks apply to TurnHurl (DifferentPeople).

But not all accept this explanation of the moral difference between Turn and Hurl. So it would be better if our account of secondary permissibility did not rely on any particular explanation of the moral difference between Turn and Hurl.

Here is my account of secondary permissibility. Consider the factor that *does* prohibit you from hurling in Hurl. Rather than asking whether you could save the five *without A*, my account of secondary permissibility instead asks whether you could save the five *without this factor*.

Whereas you could not save the five without this factor in Hurl, you could save the five without it in TurnHurl. This is because hurling is your only option of saving the five in Hurl, but you could save the five by turning instead of hurling in TurnHurl. This means that the factor that prohibits you from hurling in Hurl is *useful* to saving the five in Hurl, but it is not similarly useful to saving them in TurnHurl. Accordingly, hurling in Hurl involves *using* this factor to save the five, whereas hurling in TurnHurl does not.

My account of secondary permissibility says that this difference between Hurl and TurnHurl is morally significant. On my view, the factor that prohibits you from hurling in Hurl contributes to *impermissibility* only by preventing you from *using* that factor. So since hurling in Hurl involves using this factor to save the five whereas hurling in TurnHurl does not, it prohibits you from hurling in Hurl but it does not prohibit you from hurling in TurnHurl.

Crucially, my account of secondary permissibility is neutral regarding the factor that prohibits you from hurling in Hurl. Potential candidates include the fact that hurling involves harming A as an intended means of saving the five,[6] the fact that hurling involves harmfully

---

[6] For example, Quinn (1989), Tadros (2011).

using A to save the five,[7] or the causal pathways through which the five are saved and A is harmed.[8]

My account of secondary permissibility has the same implications in Hurl and TurnHurl regardless of which factor we accept. For since you are permitted to turn in Turn, the factor under consideration is not involved in turning. And you can save the five by turning (and so without the relevant factor – *whatever it is*) in TurnHurl but not in Hurl. My account of secondary permissibility just says that the relevant factor does not have any *independent* prohibitive force. What it prohibits instead is merely the *use* of this factor.

But why are you prohibited from *using* the relevant factor? The factor that prohibits you from hurling in Hurl is *morally objectionable*, even *independently* of contributing to impermissibility. For example, some hold that by hurling A as an intended means, you thereby aim at evil. Yet since evil by its nature should be repelled, you thereby stand in a morally inappropriate relation to evil by hurling.[9] Others hold that by harmfully using A, you thereby treat her as though she were valuable merely as a means rather than valuable as an end in herself. You thus treat A morally inappropriately by hurling.[10] Still others hold that the causal pathway through which you save the five and harm A by hurling her puts A and the five in a morally inappropriate relation with each other.[11]

These sorts of explanations are standardly offered to account for why the relevant factor independently *prohibits* you from hurling. On my view, they instead provide a deeper rationale for why the relevant factor is independently *morally objectionable*. Since this factor is

---

[7] For example, Frowe (2014), Ramakrishnan (2016), Quong (2020).

[8] For example, Kamm (1996, 2007).

[9] For example, Nagel (1986).

[10] For example, Ramakrishnan (2016).

[11] For example, Kamm (2007).

independently morally objectionable, morality excludes it from your toolkit and so says that you cannot use it.

So just as some say that a *person* is not the kind of thing you are permitted to harmfully use, my view says that a *morally objectionable factor* is also not the kind of thing you are permitted to use.

We can now resolve the problem of secondary permissibility. I have just argued that the factor that prohibits you from hurling in Hurl does not prohibit you from hurling in TurnHurl. Accordingly, hurling A into the trolley in TurnHurl is relevantly like turning the trolley, thereby *paralysing her*.

The good of saving the five is sufficiently important to justify paralysing A by turning. This is evidenced by the fact that if it were your only option of saving the five you would be permitted to turn from five onto A if this would paralyse her. So hurling in TurnHurl is *proportionate*. And since hurling is the least harmful option of saving the five in TurnHurl, hurling is also *necessary*. Thus, you are permitted to hurl in TurnHurl.

But whereas turning is the least harmful option of saving the five in Turn, the additional harm imposed by turning rather than hurling in TurnHurl is *unnecessary*. This is because hurling achieves the same good of saving the five as turning, but imposes less morally-weighted harm than turning: it paralyses rather than kills. Thus, you are prohibited from turning in TurnHurl.

Accordingly, we can explain why you are permitted to hurl and prohibited from turning in TurnHurl, even though you are prohibited from hurling in Hurl and permitted to turn in Turn. We can therefore resolve the problem of secondary permissibility.

B. Avoiding the Two Problems

Why should we accept my account of secondary permissibility? Unlike existing explanations, my account of secondary permissibility avoids the two problems previously discussed.

On the one hand, my account of secondary permissibility does not invoke any additional factors beyond what is minimally needed to distinguish morally between Turn and Hurl. This is because my account simply makes a claim about *how* this factor contributes to impermissibility. But it does not introduce any new factors into the picture.

On the other hand, my account of secondary permissibility can capture the intuition some have that you are permitted to hurl in TurnHurl (DifferentPeople). After all, the fact that A is the same individual that you can hurl into the trolley or turn the trolley onto in TurnHurl plays no explanatory role on my view. What matters instead is whether saving the five involves using the factor that prohibits you from hurling in Hurl. And just like in TurnHurl, hurling does not involve using this factor in TurnHurl (DifferentPeople). For just like in TurnHurl, you could save the five without this factor by turning instead of hurling in TurnHurl (DifferentPeople).

Some might find it counterintuitive that you are permitted to hurl in TurnHurl (DifferentPeople). So if my account of secondary permissibility has this implication, some might view this as a reason to reject my account.

But if one feels that you are prohibited from hurling in TurnHurl (DifferentPeople), hold fire for the moment. I will soon argue that my account of secondary permissibility can also be developed in a way that prohibits you from hurling in TurnHurl (DifferentPeople) while still permitting you to hurl in TurnHurl. So unlike existing explanations, my account of secondary permissibility is sufficiently flexible to accommodate different intuitions about TurnHurl (DifferentPeople). This is still an advantage over existing explanations.

Overall, then, my account of secondary permissibility is more parsimonious and has greater explanatory power than existing explanations.[12] Accordingly, we should accept my account of secondary permissibility over its alternatives. Let us now consider some objections.

---

[12] While I lack space to explore these issue in detail, consider two further advantages of my account of secondary permissibility. Firstly, my explanation of TurnHurl (DifferentPeople) can also accommodate various cases raised

IV. Defending The Principle of Secondary Permissibility

Having provided my account of secondary permissibility, I will now defend it against four objections. These objections say that my account has counterintuitive implications in other versions of TurnHurl.

A. Impermissible Alternatives

Here is the first objection. Consider

> *PushHurl*. A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by either (a) hurling A into the trolley, thereby paralysing her but stopping the trolley or (b) pushing B into the trolley, thereby killing her but stopping the trolley.[13]

Intuitively, you are prohibited from hurling in this case. It might seem that my account of secondary permissibility can explain this intuition as it stands. This is because the factor that prohibits you from hurling in Hurl is involved whether you push or hurl. And since you can only save the five by either pushing or hurling, you cannot save the five without the relevant factor. Accordingly, hurling involves using this morally objectionable factor to save the five, and so it prohibits you from hurling.

---

by Peter Unger (1992) against the moral distinction between Turn and Hurl. Accordingly, my account provides an indirect defence of the moral significance of this distinction against Unger's challenge. Secondly, my account might also provide an indirect defence of the moral significance of *intending* harm. Many feel that you are permitted to turn in Turn even if this involves killing A as an intended means (say, because she is your enemy) (for example, Thomson (1999), Kamm (2007)). But you could save the five *without this intention* in this version of Turn, since you could turn without intending A's death. So even if one finds harmful intentions independently morally objectionable, they could hold that it nevertheless does not prohibit you from turning in this case by appealing to my explanation of secondary permissibility.

[13] Graham (2020: 3).

But one could argue that while the same *kind* of factor is involved whether you push or hurl, it is not literally the same factor that is involved in each option. Rather, the *particular factor* that is involved if you push is distinct from the particular factor that is involved if you hurl. This is because pushing and hurling are two distinct ways of saving the five. So you could save the five without the particular factor that is involved when you hurl since you could push instead. One could therefore argue that hurling does not involve using this factor and so this factor does not prohibit you from hurling in PushHurl on my view.

Nevertheless, we can accommodate PushHurl. It is true that you could *physically* save the five in PushHurl without hurling in Hurl. But *morally speaking* you could not save the five without hurling in PushHurl. This is because independently considered, pushing is *morally unavailable*: if it were your only option of saving the five, pushing would be morally impermissible.

It is plausible that the relevant baseline against which to assess your options of saving the five is not only limited to the options that are physically available to you but also to the options that are *morally available* to you. For we are considering whether it is *permissible* to save the five by hurling. So we must consider whether the relevant factor is useful to saving the five *within the confines of morality*. And even if you could physically save the five without hurling and so without the relevant factor in PushHurl, you could not *permissibly* save them without it. So this factor is still useful to saving the five when considering your other morally available options.

Crucially, you could still save the five in TurnHurl without the factor that prohibits you from hurling in Hurl on this view. This is because independently considered, turning is *morally available*: if it were your only option of saving the five, turning would be morally permissible (just like in Turn). We can therefore distinguish morally between TurnHurl and PushHurl on my account.

I will now consider two counters against this explanation of PushHurl.

Firstly, consider

> *PushHurl (ConditionalConsent).* The case is similar to PushHurl. You can save the five only by either (a) hurling A into the trolley, thereby paralysing her but stopping the trolley or (b) pushing B into the trolley, thereby killing her but stopping the trolley. Crucially, however, B conditionally consents to being pushed into the trolley if and only if this is your only option of saving the five.[14]

Intuitively, you are prohibited from hurling in this case. Yet one could object that my explanation of PushHurl says that you are permitted to do so. This is because you could still save the five in PushHurl (ConditionalConsent) without the factor that prohibits you from hurling in Hurl. This is because independently considered, pushing in PushHurl (ConditionalConsent) is *morally available*: if it were your only option of saving the five, pushing would be morally permissible. For since it is your only option of saving the five in this scenario, B would consent to being pushed. I take it that consensual pushing is permissible.

But pushing B *with his consent* and pushing B *without his consent* are two different options. To adequately assess whether you could save the five without the relevant factor, we must consider whether your actual alternative options of saving the five are morally available. Yet your actual alternative option of saving the five in PushHurl (ConditionalConsent) is pushing B *without his consent*. This is because you can save the five by either pushing or hurling in PushHurl (ConditionalConset), yet B would consent only if pushing were your only option of saving the five. And independently considered, pushing without B's consent is *morally unavailable*: if it were your only option of saving the five, pushing B without his consent would be morally impermissible. Accordingly, you could not save the five in PushHurl

---

[14] Graham (2020: 5).

(ConditionalConsent) without the factor that prohibits you from hurling in Hurl, and so hurling involves using this factor to save them in this case. It thus prohibits you from hurling in PushHurl (ConditionalConsent). So we can distinguish morally between PushHurl and PushHurl (ConditionalConsent).

Here is the second counter against my explanation of PushHurl. Consider

TurnHurl (Disproportionate). A trolley is heading towards five innocent people and will kill them if it hits them. You can save the five only by either (a) turning the trolley onto a side-track, where it will hit and kill twenty or (b) hurling one of the twenty, A, into the trolley, thereby paralysing her but stopping the trolley.[15]

Some might feel that even though you are prohibited from hurling in PushHurl, you are nonetheless permitted to hurl in TurnHurl (Disproportionate). Yet one could object that my explanation of PushHurl says that you are prohibited from hurling in TurnHurl (Disproportionate). This is because you could not save the five in this case without the factor that prohibits hurling in Hurl. This is because independently considered, turning is *morally unavailable*: even if it were your only option of saving the five, turning from five onto twenty would still be morally impermissible. Accordingly, hurling in TurnHurl (Disproportionate) involves using the morally objectionable factor and so it prohibits you from hurling.

However, there is a morally significant difference between PushHurl and TurnHurl (Disproportionate). It is plausible that when each option is considered independently, pushing is *harder to justify* than turning in TurnHurl (Disproportionate). This means that when each option is considered independently, *more good* is needed to justify pushing than is needed to justify turning. To illustrate: it might seem permissible to turn a trolley from one-hundred

---

[15] Øverland (2015: 291). What I say about this case also holds if A were not one of the twenty you could turn onto.

people onto twenty, thereby killing them, yet impermissible to save one-hundred from a trolley by pushing an innocent bystander into it, thereby stopping the trolley but paralysing her

So turning is *morally closer* to being morally available in TurnHurl (Disproportionate) than pushing is to being morally available in PushHurl: when each option is considered independently, *less additional good* must be achieved to render turning permissible in TurnHurl (Disproportionate) than is needed to render pushing permissible in PushHurl.

This means that although you could not save the five without the factor that prohibits you from hurling in Hurl in either case, your ability to save the five is even *more limited* in PushHurl than it is in TurnHurl (Disproportionate). And this in turn means that although this factor is useful to saving the five in both cases, it is even *more useful* to saving them in PushHurl than it is in TurnHurl (Disproportionate). Turning in TurnHurl (Disproportionate) is less 'closed off' than pushing in PushHurl.

Hurling therefore involves using this morally objectionable factor to a *greater degree* in PushHurl than it does in TurnHurl (Disproportionate). And it seems plausible that not only is it especially difficult to justify using this factor, but that the *greater the degree of use*, the *harder it is to justify* using it. Accordingly, hurling in PushHurl is harder to justify than hurling in TurnHurl.

This means that even if saving five lives is not important enough to justify hurling in PushHurl, it could be sufficient to justify hurling in TurnHurl (Disproportionate). This explains why hurling in TurnHurl (Disproportionate) could be permissible even when hurling is impermissible in PushHurl.

Overall, we can distinguish morally between TurnHurl and PushHurl while still getting the intuitively correct results in PushHurl (ConditionalConsent) and TurnHurl (Disproportionate). Accordingly, we can avoid this first objection.

B. Costly Alternatives

Here is the second objection. Consider

> *TurnHurl (CostlyTurn).* The case is similar to TurnHurl. While hurling A is pretty much
>
> costless to you, the lever that turns the trolley onto A is very heavy and so will kill you
>
> if you pull it. Even if it were your only option of saving the five, *you would not turn.*

Some might feel that you are morally prohibited from hurling in this case.[16] Yet one could object that my account of secondary permissibility says that you are permitted to do so. For just like in TurnHurl, it looks like you could save the five in this case without the factor that prohibits you from hurling in Hurl. This is because independently considered, turning is *morally available*: if it were your only option of saving the five, turning would be morally permissible. Accordingly, one could object that relevant factor is not useful to saving the five in TurnHurl (CostlyTurn). Hurling therefore does not involve using this factor and so it does not prohibit you from hurling in TurnHurl (CostlyTurn).

However, there are two ways we could capture the intuition that you are prohibited from hurling in TurnHurl (CostlyTurn).

Firstly, we could say that although you are physically strong enough to pull the lever, you might be *mentally* too weak to pull it. This is because you might be too weak to overcome the powerful survival impulse that stands against killing yourself by turning. This mental barrier could therefore act analogously to a mechanical barrier blocking the lever. TurnHurl (CostlyTurn) would then be relevantly like Hurl in which you physically cannot turn. This assessment of your mental strength could therefore be driving the intuition that you are prohibited from hurling. By contrast, since turning does not involve a personal cost that you must overcome in TurnHurl, this option is still physically available in this case.

---

[16] For example, Pummer (2023).

Here is a second way we could accommodate TurnHurl (CostlyTurn). Suppose that you are mentally capable of turning and so this option is physically available. Nonetheless, there is a different sense in which turning is unavailable in this case. This is because independently considered, turning is *practically unavailable*: if it were your only option of saving the five, *you would not take it*. Since you are the only person who can save the five, *given that you have independently excluded turning as an option for yourself*, hurling is your only remaining option of saving the five in TurnHurl (CostlyTurn). Accordingly, you could not save the five in TurnHurl (CostlyTurn) without the factor that prohibits you from hurling in Hurl, and so hurling involves using this factor to save them. It thus prohibits you from hurling in TurnHurl (CostlyTurn).

I have just offered two ways of accommodating the intuition that you are prohibited from hurling in TurnHurl (CostlyTurn). But on the other hand, it might seem permissible to hurl in TurnHurl (CostlyTurn) if the cost of turning were much *lower* than death, *even if you still would not turn* at this much lower cost if it were your only option of saving the five. Can we explain this judgement on my view too? I think so.

Consider the first explanation we offered of TurnHurl (CostlyTurn). Even if you are mentally incapable of turning when this will kill you, you are plausibly mentally strong enough to turn when the cost of turning is significantly lower. This is because even if you are mentally too weak to overcome the powerful survival impulse that stands against killing yourself by turning, you are plausibly mentally strong enough to overcome the much weaker impulse that stands against incurring a much lower cost by turning. So even if turning is physically unavailable in TurnHurl (CostlyTurn), it appears physically available when the cost of turning is much lower. In this respect, the case is relevantly like TurnHurl.

Consider the second explanation we offered of TurnHurl (CostlyTurn). Even if turning is practically unavailable regardless of the cost involved, it seems plausible that a *greater*

reduction in personal cost is needed for it to be the case that you will turn when the cost of turning is death than is needed for it to be the case that you will turn when the cost of turning is much lower than death. So turning is *practically closer* to being practically available when the cost of turning very low than when the cost of turning is death. Turning is less 'closed off' when the cost of turning is very low than when the cost of turning is death.[17]

This means that even if you could not save the five in TurnHurl (CostlyTurn) without the factor that prohibits you from hurling in Hurl, your ability to save the five is even *more limited* when the cost of turning is death than when the cost of turning is much lower than death. And this means that although the relevant factor is useful to saving the five in both versions of the case, the factor is *even more useful* to saving the five when the cost of turning is death than when the cost of turning is much lower than death. Hurling in TurnHurl (CostlyTurn) therefore involves using this morally objectionable factor to a *greater degree* when the cost of turning is death than when it is much lower, and so using this factor is correspondingly *harder to justify*.

So even if saving five lives is not important enough to justify hurling in in TurnHurl (CostlyTurn), it could be sufficient to justify hurling when the cost of turning is much lower than death. So on this second explanation of TurnHurl (CostlyTurn), we can also explain why hurling could be prohibited when the cost of turning is death yet permitted when the cost of turning is much lower.

---

[17] I am here making the (realistic) assumption that the degree of personal cost is correlated with whether you will turn if it were your only option of saving the five. But what if (unrealistically) one is 'maximally' unwilling to turn for any cost independently of the degree of cost? Since what matters morally is the degree to which an option is practically unavailable and so the extent to which you are unwilling to turn if it were your only option, we should treat this case like one in which the cost of turning is extremely high for a psychologically normal person.

On either of these two explanations, we can therefore distinguish morally between TurnHurl and TurnHurl (CostlyTurn) while still getting intuitively correct results in other versions of TurnHurl (CostlyTurn).

Which of these two explanations of TurnHurl (CostlyTurn) should we accept?

On the one hand, suppose one still finds hurling impermissible in TurnHurl (CostlyTurn) even when it is stipulated that A is mentally strong enough to turn. Since hurling is physically available in this case, this counts in favour of the second explanation which appeals to the practical unavailability of turning.

On the other hand, suppose one finds hurling impermissible in TurnHurl (CostlyTurn) only when it is stipulated that A is mentally too weak to turn. This counts in favour of the first explanation that appeals to physical unavailability of turning.

Since intuitions might vary, I leave it open which of these two explanations of TurnHurl (CostlyTurn) one should accept. Either way, we can avoid the second objection against my account of secondary permissibility.

C. Supererogatory Alternatives

Here is the third objection. Consider

> *JumpHurl*. A trolley is heading towards five innocent people and will kill them if it hits
> them. You can save the five only by either (a) jumping into the trolley, thereby killing
> yourself but stopping the trolley or (b) hurling A into the trolley, thereby paralysing her
> but stopping the trolley.

Intuitively, you are prohibited from hurling in this case. Yet one could object that my account of secondary permissibility says that you are permitted to do so. This is because you could save the five in this case without the factor that prohibits you from hurling in Hurl. This is because independently considered, jumping is *morally available*: if it were your only option of saving the five, jumping would be morally permissible. For I take it that you are permitted to sacrifice

your life for the five. One could therefore object that the relevant factor is not useful to saving the five in JumpHurl, and so it does not prohibit you from hurling in this case.

It might seem that we could accommodate JumpHurl in the same way we accommodated TurnHurl (CostlyTurn). On the one hand, we could say that since you are mentally too weak to sacrifice yourself by jumping, jumping is *physically unavailable*. On the other hand, we could say that since you would not jump if it were your only option of saving the five, jumping is *practically unavailable*. These are realistic assumptions about the case. On either explanation, you could not save the five without the relevant factor in JumpHurl, and so hurling involves using this morally objectionable factor to save them. It thus prohibits you from hurling in this case.

But this response could not accommodate the intuition that you are prohibited from hurling in JumpHurl *even when we stipulate that jumping is both physically and practically available*: that you physically can and actually would jump if it were your only option of saving the five. We need an alternative explanation of JumpHurl that avoids this troubling result.

Suppose then that if it were your only option of saving the five, you could and actually would jump in JumpHurl. I grant that the factor that prohibits you from hurling in Hurl is not useful to saving the five in this scenario, and so this factor does not prohibit you from hurling. Nonetheless, we can still explain why you are prohibited from hurling.

Since you would jump if it were your only option of saving the five, this means that *if you do not hurl* – in which case jumping is your only remaining option of saving the five – *then you will jump*. Accordingly, there is no moral justification for hurling: the five will be saved even if you do not hurl, yet there is a stronger moral reason against *hurling* than against *refraining from hurling*.

This is because we are supposing that if you refrain from hurling, you will jump. And while there is no (or at least relatively weak) moral reason against harming yourself, there is a very

strong moral reason against harming an innocent person. Accordingly, even if you will jump if you do not hurl, hurling is still impermissible because the harm imposed is *unnecessary*: hurling achieves the same good of saving the five as refraining from hurling (i.e., jumping), but imposes *greater* morally-weighted harm (i.e., you paralyse A rather than kill yourself).

Crucially, even if you will turn if you do not hurl in TurnHurl, there is still a moral justification for hurling: although the five will be saved if you do not hurl, there is a stronger moral reason against *refraining from hurling* than against *hurling*. For we are supposing that if you refrain from hurling, you will turn. And there is a stronger moral reason against imposing the greater harm of death upon A (by turning) than there is against imposing the smaller harm of paralysis upon A (by hurling). Accordingly, even if you will turn if you do not hurl, the harm imposed by hurling is still necessary in TurnHurl: hurling achieves the same good of saving the five as refraining from hurling (i.e., turning), but imposes *less* morally-weighted harm (i.e., you paralyse rather than kill). Similar remarks apply to TurnHurl (DifferentPeople).

Note that you are still permitted to refrain from jumping in JumpHurl and so it is permissible to do nothing in this case. It is just that *even if* in an act of supererogation you will jump if you do not hurl, you are still prohibited from hurling in JumpHurl.[18]

We can therefore distinguish morally between TurnHurl and JumpHurl. On the one hand, if jumping is either physically or practically unavailable, then hurling in JumpHurl involves using the factor that prohibits you from hurling in Hurl. On the other hand, if jumping is both physically and practically available, then although hurling does not involve using this factor,

---

[18] We could also explain JumpHurl by saying that since you are willing to jump to save the five, you cannot appeal to the cost involved in jumping to justify hurling (or doing nothing). See for example Horton (2017). But some might find it counterintuitive that you are morally required to jump in this scenario. My explanation of JumpHurl thus avoids this problem.

hurling is still impermissible. This is because it is unnecessary. Accordingly, we can avoid the third objection.

    D. TurnHurl (DifferentPeople)

Here is the final objection. So far, I have argued that we can accommodate the permission to hurl in both TurnHurl *and TurnHurl (DifferentPeople)* without generating counterintuitive implications in other cases.

    But some might feel that unlike in TurnHurl, you are prohibited from hurling in TurnHurl (DifferentPeople). So if my account of secondary permissibility says that you are permitted to hurl in TurnHurl (DifferentPeople), this might be viewed by some as a reason to reject my account. However, we can develop my account of secondary permissibility in a way that accommodates the intuition that you are permitted to hurl in TurnHurl yet prohibited from hurling in TurnHurl (DifferentPeople).

    To start, many feel that there is a morally significant difference between harming by *redistributing a pre-existing threat* and harming by *creating a new threat*. To illustrate: many feel that while you are permitted to turn a trolley from five onto one, you are nonetheless prohibited from exploding a trolley heading towards five when the blast will foreseeably kill an innocent bystander.

    Whereas turning the trolley involves killing the one by redistributing the pre-existing threat posed by the trolley from the five onto the one, exploding the trolley instead involves killing the one by creating the new threat posed by the grenade blast. So it seems that there is a moral barrier against creating a new threat that does not count against redistributing a pre-existing threat.[19]

---

[19] For example, Thomson (1986).

Return now to TurnHurl and TurnHurl (DifferentPeople). In TurnHurl, turning will kill A. And since turning would be permissible if it were your only option of saving the five, independently considered you can permissibly kill A by turning in this case. This means that although A is not under a pre-existing *physical* threat in TurnHurl, independently considered A is under a pre-existing *moral* threat: the threat of being *permissibly* killed. This is a threat in the sense that A is not shielded by morality prohibiting you from killing her.

So if the option to turn could render the option to hurl permissible in TurnHurl, this would involve redistributing the very same moral threat (i.e., the moral threat that A faces of being permissibly killed) from one option to another. This is because the permission to kill (and thus also paralyse) A *by hurling* is explained by appeal to the independent permissibility of killing A *by turning*.

By contrast, A is not under a pre-existing moral threat in TurnHurl (DifferentPeople). For although turning would be permissible if it were your only option of saving the five, turning will kill only *B*. So independently considered, you cannot permissibly kill A in this case. This means that if the option to turn could render the option to hurl permissible in this case, this would involve removing the moral shield that is protecting A by placing her under the moral threat of being permissibly killed for the very first time. It would thus involve *creating* a new moral threat.

Perhaps just as there is a moral barrier against you creating a new physical threat that does not count against you redistributing a pre-existing physical threat, there is also a barrier against morality creating a new moral threat that does not count against it redistributing a pre-existing moral threat. Accordingly, the option to turn could render it permissible to hurl in TurnHurl but not in TurnHurl (DifferentPeople).

We could capture this by saying that an option *1* is morally available *relative to* another option *2* only if the individual that will be harmed in option 2 can also be permissibility harmed

to at least the same degree in option 1 when option 1 is considered independently. For this ensures that option 1 can render option 2 permissible only by redistributing rather than creating a moral threat. This means that turning is morally available relative to hurling in TurnHurl but not in TurnHurl (DifferentPeople).

So although you could save the five in TurnHurl without the factor that prohibits you from hurling in Hurl, you could not save the five without this factor in TurnHurl (DifferentPeople) on this view. Hurling therefore involves using this factor to save the five in TurnHurl but not in TurnHurl (DifferentPeople). It thus prohibits you from hurling in TurnHurl (DifferentPeople) but not in TurnHurl. We can therefore explain why you are permitted to hurl in TurnHurl but not in TurnHurl (DifferentPeople) even on my account of secondary permissibility. Accordingly, we can avoid this final objection.

Since intuitions might vary about TurnHurl (DifferentPeople), I leave it open whether one should accept the amendment I have just offered. My account of secondary permissibility is sufficiently flexible to accommodate different intuitions about TurnHurl (DifferentPeople). This is an advantage over existing explanations of secondary permissibility.

Overall, then, my account of secondary permissibility can accommodate TurnHurl without generating counterintuitive implications in other cases. It thus avoids the four objections we have just considered in this section.

V.   Conclusion

In sum, I have provided a novel account of secondary permissibility. On my view, the factor that prohibits you from hurling in Hurl does not have any independent prohibitive force. It contributes to impermissibility only by preventing you from *using* this morally objectionable factor. This explains why you are permitted to hurl and prohibited from turning in TurnHurl, even though you are prohibited from hurling in Hurl and permitted to turn in Turn. We can therefore resolve the problem of secondary permissibility.

In Section II, I raised two problems for existing explanations of secondary permissibility. In Section III, I argued that my account avoids these problems. On the one hand, it does not posit any additional factors beyond what is minimally needed to distinguish morally between Turn and Hurl. And on the other hand, it can capture the intuition some have that you are permitted to hurl in TurnHurl (DifferentPeople). Accordingly, my account is more parsimonious and has greater explanatory power than existing explanation of secondary permissibility. Finally, in Section IV, I argued that my account of secondarily permissibility can also accommodate other potentially troubling versions of TurnHurl.

References

Frowe, Helen (2014). *Defensive Killing*. Oxford, UK: Oxford University Press.

Graham, Peter A. (2020). "Secondary Permissibility" and the Ethics of Harming. *Journal of Moral Philosophy* 18 (2):156-177.

Horton, Joe (2017). The All or Nothing Problem. *Journal of Philosophy* 114 (2):94-104.

Kamm, Frances (2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.

Locke D. I, Trolley: Self-Redirection and Hybrid Trolley Cases. *Utilitas*. 2021;33(4):458-473.

Nagel, Thomas (1986). *The View From Nowhere*. New York: Oxford University Press.

Øverland, Gerhard (2015). Why Kamm's Principle of Secondary Permissibility Cannot Save the Doctrine of Double Effect. *Journal of Applied Philosophy* 33 (3):286-296.

Pummer, Theron (2023). Contrastive Consent and Secondary Permissibility. *Philosophy and Phenomenological Research* 106 (3):677-691.

Quinn, Warren S. (1989). Actions, Intentions, and Consequences: The Doctrine of Double Effect. *Philosophy and Public Affairs* 18 (4): 334-351.

Quong, Jonathan (2020). *The Morality of Defensive Force*. Oxford University Press.

Ramakrishnan, Ketan H. (2016). Treating People as Tools. *Philosophy and Public Affairs* 44 (2): 133-165.

Tadros, Victor (2011). *The Ends of Harm: The Moral Foundations of Criminal Law*. New York: Oxford University Press.

Thomson, Judith Jarvis (1986). The Trolley Problem. In William Parent (ed.), *Rights, Restitution, and Risk: Essays in Moral Theory*. Cambridge, Mass.: Harvard University Press. Edited by William Parent.

Thomson, Judith Jarvis. "Physician-Assisted Suicide: Two Moral Arguments." *Ethics* 109, no. 3 (1999): 497–518.

Unger, Peter K. (1996). *Living high and letting die: our illusion of innocence*. New York: Oxford University Press.

Chapter II: Killing Innocent Threats

I. Introduction

You are permitted to kill an innocent person to save yourself in some cases but not permitted to do so in others. What explains the moral asymmetry between these cases?

Consider

> *Well.* You are stuck at the bottom of a well. A strong gust of wind has blown Albert down the well towards you. If he lands on you, the collision with Albert will kill you but leave him unharmed. You can save yourself only by vaporising Albert with your ray gun, thereby preventing him from hitting you.[1]

Most find it intuitively plausible that you are permitted to kill Albert in this case. Although he is not morally responsible for the threat to your life, it seems absurd that you must lie there and take it. What explains this permission?

Any plausible answer must avoid counterintuitive implications in the following case

> *Alcove.* You are stuck in a tunnel. Due to a mechanical fault, a trolley is hurtling towards you. If it hits you, the trolley will kill you. The only safe space you can avoid the trolley is a very tiny one-person alcove where Betty is stuck. You can save yourself only by vaporising her, thereby emptying the alcove.

Most find it intuitively plausible that you are prohibited from killing Betty in this case. Yet just like in Well, you can save yourself only by intentionally killing an innocent person who is not morally responsible for the threat to your life. If you are permitted to kill Albert in Well, why may you not also kill Betty in Alcove?

---

[1] Unless stated otherwise, I assume that no one is morally responsible for the fact that you must kill someone to survive in all cases considered here.

Against the intuitions of most, some say that there is no moral difference between Well and Alcove. And since you are clearly prohibited from killing Betty in Alcove, they conclude that you are also prohibited from killing Albert in Well.[2]

But it is hard to believe this conclusion. In my view, it is mistaken.

In Section II, I argue that existing explanations of the moral asymmetry between Well and Alcove do not get to the heart of the matter. This is because they have counterintuitive implications in cases that seem relevantly like Well or relevantly like Alcove. I offer an alternative account that provides a *unified* explanation of our intuitions in such cases. In particular, I argue that whereas killing Betty in Alcove involves *harmfully treating her as a means*, killing Albert in Well does not.

Although she is not useful to your survival, you nonetheless kill Betty on the basis of *how useful* she is to this end. You thus treat her as though she were valuable merely as a means of saving yourself. Yet whether we are permissibly killed should not hang on our utility to others. If morality said otherwise, it would be at odds with our value as ends in ourselves.

Accordingly, killing Betty is harder to justify than killing Albert. So even if saving your life justifies killing Albert in Well, it is insufficiently important to justify killing Betty in Alcove. In Section II, I argue that this explains why you are prohibited from killing Betty even if you are permitted to kill Albert.

In the remainder of the paper, I argue that your permission to kill Albert in Well is justified by appeal to his strict enforceable duty to bear the costs of your rescue. Accordingly, you are permitted to impose this cost on Albert to save yourself. However, the fact that he will kill you has no moral relevance on my view. The stringency of Albert's enforceable duty is rather explained by *how* he will incur the costs of your rescue.

---

[2] For example, McMahan (1994; 2002; 2009; 2018), Otsuka (1994; 2016), Tadros (2020).

In Section III, I defend this explanation of Well against two objections. The first objection says that even if killing Albert does not involve harmfully treating him as a means, it nevertheless involves intending his death and so is prohibited by the doctrine of double effect. The second objection says that my explanation of Well counterintuitively implies that you are permitted to kill an innocent bystander as a mere foreseen side-effect of saving yourself. In my view, both objections fail. Most notably, I argue that we can capture the intuitions driving each objection, but in a way that does not undermine my position.

In Section IV, I provide a positive argument for why you are permitted to kill Albert in Well. Whether this explanation is exclusively agent-neutral or must also invoke agent-relative considerations depends on whether the permission extends to an impartial third-party. I argue that my view can accommodate either position.

Overall, I offer a novel account of your permission to kill an innocent threat that provides a unified explanation of our intuitions in cases.

Before I go on, a short remark about practical relevance. Many soldiers fighting in wars today arguably qualify as innocent threats. Due to factors such as ignorance and duress, they may not be morally responsible for their threatening behaviour. Additionally, many morally responsible soldiers arguably pose objectively justified threats. Yet some hold that there is no liability-based justification for harming justified threats.[3] While I lack space to explore the issue in detail, the permission to kill Albert in Well advanced here could extend to defensively harming such innocent and justified threats in war.

---

[3] For example, McMahan (2005; 2009), Quong (2020).

II.   The Prohibition

In this section, I argue that we can distinguish morally between Well and Alcove. I first identify the factor that is stimulating our intuitions in these cases, before explaining why it is morally significant.

A. Moral Status

Why does Alcove seem morally different from Well? I will defend the following explanation:

Moral Status. *You are permitted to kill an innocent person if this helps save you by making a negative fact hold. By contrast, you are prohibited from killing an innocent person if this helps save you by making a positive fact hold.*[4]

When considering how a killing helps save you, I argue that we are looking for:

The fact that provides the least derivative explanation relevant to your survival, that is directly made true by the most proximate cause of the person's death.

At first glance, it may be unclear why my account includes all these specific features. But this will become clearer as we proceed to apply it to a range of cases.[5]

I first show how Moral Status captures our intuitions in Well and Alcove. I then argue that it is the best explanation of these intuitions.

To start, what are positive and negative facts? A positive fact is a fact that something *is* the case. To illustrate: the fact that *my car is in the parking lot* is a positive fact. It is a fact of where

---

[4] For simplicity, I assume throughout that a killing is necessary to save only your life, will not harm anyone else, and that you have not waived or forfeited your ordinary moral rights and permissions. I assume that when enough good is at stake you are nonetheless permitted to override the constraint against harming provided by Moral Status.

[5] Similarly, the name 'Moral Status' is tied to the rationale offered below.

my car *is*. A negative fact, by contrast, is a fact that something is *not* the case. The fact that *my car is not in my driveway* is a negative fact. It is a fact of where my car is *not*.[6]

Our awareness of facts can put us in a position to make assertions. But these assertions can often be phrased either positive or negatively.

For example, doesn't the fact that Al is alive tell us both that he is alive and that he's not dead; and doesn't the fact that Deb is dead tell us both that she is dead and that she's not alive? So which fact is negative and which fact is positive? Or is the idea supposed to be that the fact that Deb is dead and the fact that Deb is not alive, which sound like the same fact, are actually two different facts, the first positive and the second negative?

Since the distinction between positive and negative facts bears a lot of moral weight on Moral Status, one might worry that how we draw this distinction in certain cases might be arbitrary.

However, the fact that *Deb is dead* is reducible to the negative fact that *Deb is not alive* together with the positive fact that *Deb was earlier alive*. The fact that *Deb is dead* and the fact that *Deb is not alive* are two different facts, since trains and planes are not alive but they are also not dead. The positive fact that *Al is alive* and the negative fact that *Al is not dead* are also two different facts, since telephones and radios are not dead but they are also not alive.

Similar remarks apply to other examples. The fact that *Bill in blind* is reducible to the negative fact that *Bill cannot see* together with the positive fact that *Bill has sensory organs*. The fact that *Bill is blind* and the fact that *Bill cannot see* are two different facts, since tables and chairs cannot see but they are also not blind. Likewise, the fact that *Felicity failed the exam* is reducible to the negative fact that *Felecity did not pass the exam* together with the positive

---

[6] For example, Bennett (1995), Woollard (2015). It is enough for my purposes that we have an intuitive grasp of this distinction in cases, even if we are unaware of the deeper factors driving such judgements. I nonetheless take it to be a robust metaphysical distinction, not merely conventional.

fact that *Felicity sat the exam*. The fact that *Felicity failed the exam* and the fact that *Felecity did not pass the exam* are two different facts, since her pet dog did not pass the exam but they also did not fail it.

So even in controversial cases in which it might seem arbitrary or simply a matter of convention how the distinction between positive and negative facts could be drawn, we can draw this distinction in a principled manner.[7]

With this distinction in hand, we can now construct Moral Status in stages.

Firstly, *firing your ray gun*, for example, is *a* proximate cause of Albert's death in Well. However, it causes Albert's death only *by* causing *your ray to hit him* (a positive fact). If this were how killing Albert helps save you, Moral Status would mischaracterise Well. This shows that what matters is how the *most* proximate cause of the person's death helps save you. In Well, the most proximate cause of Albert's death is your ray hitting him. Similar remarks apply to killing Betty in Alcove.

Next, the fact that *Albert will not hit you* is negative. This negative fact is a causal consequence of your ray hitting Albert. Causation is a making relation; what matters is how the killing helps *make* you safe (i.e., how it is useful to your survival). Accordingly, killing Albert in Well helps save you by making a negative fact hold.

Alcove illustrates two additional features necessary for understanding how a killing helps save you. Here is the first. *Betty not being in the alcove* is also a negative fact that is a causal consequence of your ray hitting her, and so is made true by the most proximate cause of her death. If this were how killing Betty helps make you safe, Moral Status would mischaracterise

---

[7] Crucially, the concern I am addressing here is exclusively about whether how we draw the distinction between positive and negative facts is *arbitrary* in certain cases. I have argued that how we draw the distinction is not arbitrary or merely conventional, but I remain neutral on the deeper metaphysics of positive and negative facts. For a possible explanation of the metaphysics underlying this distinction, see for example Baker and Jago (2012).

Alcove. This shows that when considering how killing a person helps make you safe, what matters is the *least derivative* explanation of how it accomplishes this. For consider that Betty not being in the alcove is derivative on *you being able to access the alcove* as an explanation of how killing her helps make you safe. This is because Betty not being in the alcove is relevant to explaining your survival in virtue of the fact that you can access the alcove. Accordingly, your ability to access the alcove is a *less derivative* explanation in this respect than Betty not being inside of the alcove. More generally, for any fact *A* that is relevant to explaining your survival, we can ask whether A is relevant to explaining your survival in virtue of a different fact *B*. If it is, then A is *less derivative* than B with respect to explaining your survival.[8]

Here is the second additional feature. The fact that you can access the alcove is itself derivative on *the trolley not hitting you* as an explanation of how killing Betty helps make you safe. Accordingly, the trolley not hitting you is an *even less derivative* explanation in this respect than you being able to access the alcove. Yet the fact that the trolley will not hit you is *negative*. If this were how killing Betty helps make you safe, Moral Status would

---

[8] Could one instead say that killing Betty helps make you safe by making it such that there is *no obstacle to your moving out of the path of the trolley* (a negative fact), which is less derivative than the fact that you can access the alcove? No. The constraint I have set out on what contributes to a derivative explanation precludes saying this. This is because the fact that there is no obstacle to your moving out of the path of the trolley is still relevant to explaining your survival in virtue of a positive fact: that *you can avoid the trolley*. Accordingly, your ability to avoid the trolley is a *less derivative* explanation in this respect than there being no obstacle to your moving out of its path. Note that this worry about arbitrariness is distinct from the concern I address in the next footnote.

mischaracterise Alcove. This shows that what matters is the killing *per se*: how killing the person *itself* helps save you.

For consider that killing Betty does not *itself* make it such that the trolley will not hit you. After all, your ray hitting Betty does not *directly* make it such that the trolley will not hit you. This is instead mediated by intervening events that occur downstream from your ray hitting Betty. Rather, killing Betty *per se* helps save you by making it such that you can access the alcove (a positive fact). For your ray hitting Betty directly causes you to be able to access the alcove and so directly makes this fact hold.[9]

Accordingly, Well and Alcove reveal that when considering how a killing helps save you, we are looking for:

The fact that provides the least derivative explanation relevant to your survival, that is directly made true by the most proximate cause of the person's death.

So understood, killing Albert in Well helps save you by making a negative fact hold (i.e., that he will not hit you), whereas killing Betty in Alcove helps save you by making a positive fact hold (i.e., that you can access the alcove). Moral Status therefore gets the intuitively correct result that you may kill Albert but not Betty.

---

[9] One could respond that your ray hitting Betty directly causes Betty not to be in the alcove (a negative fact), which in turn makes it the case that you can access the alcove (a positive fact). However, it is not the case that Betty's not being in the alcove *makes* it such that you can access the alcove. For consider that whereas vaporising Betty *causes* you to be able to access the alcove and so *makes* this fact hold, Betty not being in the alcove is merely *non-causally* derivative on your being able to access the alcove as an explanation of how killing Betty helps save you. While it is plausible that there can be non-causal making relations (such as grounding), it does not seem plausible that *negative* facts (in particular, that Betty is not in the alcove) can non-causally *make* positive facts hold (in particular, that you can access the alcove).

In the remainder of this section, I argue that alternative explanations of the moral asymmetry between Well and Alcove have counterintuitive implications in cases that seem relevantly like Well or relevantly like Alcove. (I refer to these alternatives as: Harmful Use, Causal Pathways, and Victim's Role). Accordingly, they do not capture our sense that the deontic status of such killings are explained on largely the same grounds. By contrast, Moral Status provides a *unified* explanation of our intuitions in such cases. In subsection F, I provide a rationale for Moral Status.

B. Harmful Use

Consider, firstly, *Harmful Use*:

> *You are permitted to kill an innocent person to help save yourself if and only if this does not involve using them or their property.*[10]

To test whether killing a person to help save yourself involves using them or their property in the relevant sense, Harmful Use asks us to consider whether you would be safe in their absence. It is true that if Betty was absent in Alcove, you could safely access the alcove. But suppose we have a property claim over the space our body occupies. If Betty's space was inaccessible to you (for example, blocked by a barrier), you would be unable to access the alcove and so the trolley would kill you. Accordingly, Harmful Use prohibits you from killing Betty. In Well, by contrast, you would be safe in the absence of Albert and his property. For he would then not be heading towards you. Accordingly, Harmful Use permits you to kill Albert.

Harmful Use faces two problems. Firstly, it has counterintuitive implications in cases that seem relevantly like Well. Consider

> *Pedal.* Albert is pressing the pedal of an unowned trolley, causing it to accelerate towards you. If it hits you, the trolley will kill you. You can save yourself only by

---

[10] Quong (2009; 2020). I set aside further caveats irrelevant to my objections throughout.

vaporising Albert, thereby preventing him from pressing the pedal. The trolley will then slow to a natural halt before it reaches you.

You seem permitted to kill Albert in this case. Yet if Betty has a property claim over the space she occupies in Alcove, Albert must also have a property claim over the space he occupies in Pedal. This means that if Albert and his property were absent in this case, the trolley would still kill you. For even if Albert was absent, the pedal could not pass through his inaccessible space and so could not move from its currently depressed position. Accordingly, Harmful Use prohibits you from killing him.

Secondly, Harmful Use has counterintuitive implications in cases that seems relevantly like Alcove. Consider

> *Wall*. The case is similar to Alcove. However, Betty is lying on a remote button that is actively raising a wall blocking the alcove. You can save yourself only by vaporising Betty, thereby preventing her from pressing the button. The wall will then fall, unblocking the alcove.

You seem prohibited from killing Betty in this case. However, you would be safe in the absence of Betty and her property in Wall. For Betty would then not be pressing the button raising the wall, and so you could access the alcove. Accordingly, Harmful Use permits you to kill her.

What does Moral Status say about Pedal and Wall? In Pedal, vaporising Albert helps save you by making it such that the trolley will not hit you (a negative fact).[11] In Wall, by contrast, vaporising Betty helps save you by making it such that you can access the alcove (a positive fact). Unlike Harmful Use, Moral Status therefore correctly says that you may kill Albert in Pedal but not Betty in Wall.

---

[11] Since Albert not pressing the pedal is *non-causally* derivative on the trolley not hitting you (see footnote 9 for further clarification). For strictly speaking the pedal does not need to be upright (a positive fact) for the trolley to not hit you. It simply needs to *not* be pressed down.

C. Causal Pathways

Next, consider *Causal Pathways*:

> *You are permitted to kill an innocent person to help save yourself if their being killed is one and the same event as you being safe. By contrast, you are prohibited from killing an innocent person to help save yourself if their being killed is causally upstream you being safe.*[12]

In Well, Albert being killed is one and the same event as you being safe. For Albert being vaporised *just is* for him to not be falling towards you. In Alcove, by contrast, Betty being killed is causally upstream you being safe inside the alcove. Accordingly, Causal Pathways permits you to kill Albert in Well but prohibits you from killing Betty in Alcove.

Causal Pathways also accommodates Pedal and Wall. In Pedal, Albert being vaporised is one and the same event you being safe. After all, strictly speaking the pedal does not need to be upright for the trolley to not hit you. It simply needs to *not* be pressed down. And Albert being vaporised *just is* for him not to be pressing the pedal. In Wall, by contrast, Betty being killed is causally upstream you being safe inside the alcove.

Nevertheless, consider

> *Area Bomber*. A sentry gun is about to fire at you and kill you from a remote area. As you have only a bomb at hand, you can save yourself only by obliterating a broad area as a means of destroying the sentry gun contained within it. However, Betty is trapped inside the broad area and so will be killed if you destroy it.

By using a blunt weapon that destroys the entire area, your means to the end of destroying the sentry gun is to destroy the area containing it, which also happens to contain Betty. It is plausible that an area's being destroyed *constitutes* its occupant's being destroyed, just like a

---

[12] Kamm (2007), Otsuka (2016).

computer file's being erased constitutes its contained text's being erased.[13] Accordingly, since the area's being destroyed constitutes both the sentry gun and Betty being destroyed, Betty being killed is one and the same event as you being safe in this case. Causal Pathways therefore permits you to kill her. But you seem prohibited from killing Betty in Area Bomber on largely the same grounds as in Alcove.[14]

What does Moral Status say about Area Bomber? It is true that the most proximate cause of Betty's death is the explosion hitting her.[15] But this is not itself relevant to your survival. Rather, the most proximate cause of Betty's death that *also* helps save you is the explosion hitting the broad area. And the least derivative explanation of how it accomplishes this is by making a positive fact hold: that the explosion hit a specific region within this area (i.e., the sentry gun). For it is plausible that *positive* facts (in particular, hitting the broad area) can non-causally *make* positive facts hold (in particular, hitting a specific region within this area i.e., the sentry gun). Unlike Causal Pathways, Moral Status therefore correctly prohibits you from killing Betty in this case.

One could respond that Betty being killed in Area Bomber is part of a larger event (the area bombing) that contains another part (the destruction of the sentry gun) which is the same as your being safe. But just as a computer file's being erased constitutes the *entirety* of its contained text's being erased, the destruction of the area constitutes the destruction of the

---

[13] Fitzpatrick (2006).

[14] Nagel (1972), Kamm (2011).

[15] Although for simplicity I have focused on the most proximate *cause* of a person's death, causation is a making relation; what matters is what most proximately *makes* them dead. This is just like how what matters on Moral Status is the way a killing least derivatively helps *make* you safe. So while the explosion hitting Betty is merely *non-causally* related to her death in Area Bomber, it plays a functionally equivalent role to a cause on my account: it is what most proximately *makes* her dead.

entirety of its occupants, which includes *both Betty and the sentry gun*. Even if no presently existing bombs work like this, we could simply imagine dropping an object (for example, a large concrete block) on an area that will destroy the entirety of its occupants simultaneously. So to deny that the destruction of Betty and the sentry gun are a single event is no more plausible than denying that the erasure of the entire text contained within a deleted computer file is a single event. Accordingly, this response on behalf on Causal Pathways does not succeed.[16]

D. Victim's Role

Finally, consider *Victim's Role*:

> *You are permitted to kill an innocent person to help save yourself if and only if they are going to kill you.*[17]

Albert is going to kill you in Well. By contrast, Betty is merely obstructing your escape from the trolley that will kill you in Alcove. Accordingly, Victim's Role permits you to kill Albert in Well but prohibits you from killing Betty in Alcove. Similar remarks explain why Victim's Role also accommodates Pedal, Wall, and Area Bomber.

Despite its intuitive appeal, we should reject Victim's Role. Consider

> *Cave.* You are stuck inside a cave that is slowly filling with water. If you do nothing, you will soon drown to death. Betty is covering the only hole in the cave through which the water can be drained. You can save yourself only by vaporising Betty, thereby uncovering the hole.[18]

---

[16] There might be a way to rehabilitate Causal Pathways if one could somehow show that Betty being killed and the sentry gun being destroyed are different events, but I do not see how they could do this.

[17] For example, Thomson (1991), Kamm (1992), Tadros (2011), Frowe (2014), Doggett (2018).

[18] Foot (1967).

You seem prohibited from killing Betty in this case. But Victim's Role cannot accommodate this judgement without yielding counterintuitive results in

> *Broken Track.* A lethal trolley is heading towards you. Although a section of the track is missing, Albert is lodged between the two ends of the broken track. The trolley will therefore harmlessly travel over him instead of falling through the gap in the track. You can save yourself only by vaporising Albert such that the trolley falls through the gap before it hits you.

You seem permitted to kill Albert in this case. However, Albert and Betty play equivalent roles in Broken Track and Cave, respectively. For the interaction between Albert and Betty on the one hand, and the threat to your life on the other, is the same each case: they merely passively resist the threat from falling through a gap and it thereby travels towards you instead. So if Albert's role permits you to kill him in Broken Track (i.e., if he will kill you), the same must also be true of Betty in Cave.

By contrast, Moral Status can distinguish morally between these two cases. In Broken Track, vaporising Albert helps save you by making it such that the trolley will not hit you (a negative fact). This is because the trolley must merely *not* cross the gap in the track to avoid hitting you. And this is accomplished by killing Albert.[19] Accordingly, Moral Status permits you to kill him.

In Cave, by contrast, vaporising Betty helps save you by making it such that the water can access the hole (a positive fact). For the water needs to subsequently drain out of the hole to avoid drowning you. If it simply refrained from pushing up against Betty without actually

---

[19] Since Albert not bridging the gap is *non-causally* derivative on the trolley not crossing it. It is true that killing Albert also makes it such that the trolley can access the gap in the track (a positive fact). But this is not *how* killing Albert helps save you. For if the trolley simply refrained from crossing this gap without actually falling through it, the trolley would still not hit you.

draining, the water would still rise and drown you. Accordingly, Moral Status prohibits you from killing her.

E.   Defending Moral Status

I have just argued that unlike its alternatives, Moral Status provides a unified explanation of the cases considered so far. Yet one might object that Moral Status has counterintuitive implications in other cases that seem relevantly like Well or relevantly like Alcove, and so also fails to capture the essence of the asymmetry between them.

Firstly, consider:

> *Meteor*. A small meteor is falling towards you and will kill you if it hits you. The only safe space you can avoid the meteor is your very tiny one-person car. However, Albert is already in your car. You can save yourself only by vaporising him to empty it.[20]

Some feel that you are permitted to kill Albert in this case. Yet one could argue that killing him helps save you by making it such that you can access your car (a positive fact). If this were correct, Moral Status would mischaracterise Meteor.

But consider that when your bodily movement is restricted through its natural plane of motion (for example, as when you are confined within a tight cage), it intuitively seems that your body is *fixed* in this region, rather than merely unable to access the restricted area (for example, as when you cannot enter a locked vault). If the intuition that you are permitted to kill Albert in Meteor is correct, then Meteor illustrates that the relevant space against which to compare a limitation of your bodily movement includes not only its natural plane of motion but also *the space that you own*. Owning additional space is thus relevantly like expanding your body's natural plane of motion by becoming more flexible or growing longer limbs. Since it is

---

[20] Quong (2009).

your car, restricting your movement through this space is therefore relevantly like restricting your body's movement through its natural plane of motion.

Accordingly, the meteor will kill you because you are *fixed* within its trajectory. Killing Albert therefore helps save you by making it such that the meteor will *not* hit you (a negative fact).[21] Moral Status thus permits you to kill him. By contrast, since you do not own the obstructed alcove space, your body is not similarly restricted in Alcove. Accordingly, we cannot similarly argue that killing Betty in this case helps save you by making a negative fact hold. Killing her thus remains prohibited on Moral Status.

Next, consider:

> *Crush*. The case is similar to Alcove. Although Betty is inside the alcove, she is not physically obstructing your access. Nevertheless, moving into the alcove will fatally crush her.[22]

You seem prohibited from killing Betty in this case. Yet one could argue that killing her helps save you by making it such that the trolley will not hit you (a negative fact).[23] If this were correct, Moral Status would mischaracterise Crush.

---

[21] Since not being fixed in the meteor's path is *non-causally* derivative on the meteor not hitting you. It is true that killing Albert in Meteor also makes it such that you can access your car (a positive fact). But this is not *how* killing Albert helps save you. For you are unable to access your car *because* you are fixed in the meteor's path. And since being fixed in the meteor's path itself seems to explain why you will be hit *over and above* merely not occupying your car, your fixture within the meteor's path independently explains why it will hit you if you do not kill Albert. This is why killing Albert helps save you by making it such that the meteor will not hit you. It is true that *after* killing Albert, saving yourself involves accessing your car. Nevertheless, killing Albert helps save you by making it such that the meteor will not hit you *at the time of killing him*.

[22] Thomson (1991), Otsuka (2016).

[23] Since not being in the trolley's path is *non-causally* derivative on the trolley not hitting you.

But consider that you avoid the trolley *by* occupying the alcove. Although occupying the alcove is *non-causally* derivative on the trolley not hitting you as an explanation of how killing Betty helps save you, it is plausible that *positive* facts (in particular, your occupying the alcove) can non-causally *make* negative facts hold (in particular, your not being outside the alcove in the path of the trolley).

Accordingly, the most proximate cause of Betty's death in Crush that is *also* relevant to your survival is your moving into the alcove, and the least derivative explanation of how this helps save you is by making it such that you are inside of the alcove (a positive fact). After all, moving into the alcove *directly causes* you to be inside it, and so *directly makes* this fact hold. Crucially, it does not directly make it such that the trolley will not hit you. For we just saw that this is directly mediated by your occupying the alcove. Moral Status thus prohibits you from killing Betty in Crush.[24]

F. The Rationale

Having identified the factor driving our intuitions in Well and Alcove, I will now explain why it is morally significant.

We need a deeper explanation of why (a) we are looking for the fact that provides the least derivative explanation relevant to your survival, that is directly made true by the most

---

[24] Moral Status still explains why you may kill Albert in Well by *deflecting him to your side* though he will subsequently fatally hit a hard surface. It is true that just as you avoid the trolley by occupying the alcove in Crush, you avoid Albert by making him occupy your side in this case. But this does not show that killing him helps save you by making it such that he is occupying your side (a positive fact). For Albert's fatal impact is causally *downstream* from making him occupy this spot. This is relevant on Moral Status because it means that the most proximate cause of Albert's death that is also relevant to your survival is his occupying your side, and the least derivative explanation of how this helps save you is by making it such that he will not hit you (a negative fact).

proximate cause of the person's death, and why (b) it matters morally whether this fact is positive or negative.

Underlying (a) is the view that if an act can be decomposed into a sequence of distinct causal stages, none of which necessitates further intervention from the agent, then it is permissible to perform this act only if each distinct causal stage is permissible when considered individually.[25]

By performing a single act that brings about an automated sequence of distinct causal stages, you are effectively carrying out each distinct causal stage in turn. For example, even if you set off a row of dominoes only by toppling the first domino in the set, you effectively topple each individual domino in the row in turn.

If you were performing a sequence of distinct acts, each act in the sequence must be justified individually. Similarly, when you are performing a single act that brings about a sequence of distinct causal stages, each causal stage in this sequence must also be justified individually.

This explains why what matters morally is the *most proximate cause* of a person's death. For this is the particular causal stage that involves killing a person and so is the stage that stands in need of justification.

We are also considering what is made *directly* true by the most proximate cause of the person's death. For this ensures that the causal stage that involves killing the person is justified *individually*, by limiting the goods that we can invoke to justify killing only to those goods that obtain within the same causal stage.

Finally, the good at stake is your own survival. So of the restricted goods that we can appeal to, the good that *ultimately* justifies the harm within a particular causal stage is the fact that provides the *least derivative explanation* relevant to your survival.

But why does it matter whether (b) this fact is positive or negative?

---

[25] See also Frick (2015), Gordon-Solmon (2019).

Suppose that saving your life justifies overriding the constraint against doing lethal harm to an innocent person. Although imposing this harm upon Betty in Alcove will help secure this good, it does so by making a positive fact hold. This positive fact is itself *useful* to your survival. After all, being able to access the alcove presents an opportunity that you can take advantage of to avoid the trolley. This is like how being able to see presents an opportunity that you can take advantage of to read a book or watch a film, for example.

This means that Betty will be killed because she is *not* useful: it is as if you consider whether, like your ability to access the alcove, Betty herself presents an opportunity that you can take advantage of to save yourself, and then kill her because she does not. By killing Betty on the basis on *how useful* she is to your survival, you thus treat her as though she were valuable (and so not be killed) only insofar as she is useful to this end.[26]

It might seem that you are not killing Betty on the basis of how useful *she* is. You are killing her on the basis of how useful accessing to alcove is to your survival, which is something she impedes. Accordingly, you are killing Betty because she is an obstacle to your taking a means (accessing the alcove) to your survival.

It is true that you are killing Betty because she is an obstacle to your accessing the alcove. But killing Betty on the basis of how useful accessing to alcove is to your survival *also* involves killing Betty on the basis of how useful *she* is. For the grounds on which you favour your access to the alcove over Betty is that the alcove is more useful to your survival than Betty is: if only she were more useful to your survival, you would not decide against her. It is her inability to meet this standard of usefulness that counts towards killing her over forgoing access to the alcove.

---

[26] See also Kamm (1993).

Accordingly, although killing Betty does not involve harmfully *using* her as a means to your survival, it nonetheless involves harmfully *treating* her as such. Yet whether we are permissibly killed should not hang on our utility to others. If morality said otherwise, it would be at odds with our value as *ends in ourselves*. Morality would thus not adequately reflect our moral status.[27]

This means that killing Betty in Alcove is harder to justify than killing Albert in Well. For killing Albert helps save you by making a negative fact hold. Yet this negative fact is not *itself* useful to saving yourself. Although Albert not hitting you is necessary for your survival, it does not itself present an opportunity that you can wield to save yourself. After all, a negative fact tells us only that something is *not* the case, rather than saying that there is something tangible obtaining that we can take advantage of. This is like how not having a splinter in your eye is necessary to see clearly, though it would be odd to say that this is itself useful to your vision (unlike the room being brightly lit, for example). Unlike killing Betty, you thus do not kill Albert on the basis of how useful he is to your survival, and so do not harmfully treat him as a means to this end. This explains why you are prohibited from killing Betty in Alcove even if you are permitted to kill Albert in Well.

In saying that killing Betty involves treating her as a means, I am not making a claim about your intentions or attitudes towards her. Rather, my point is that you act *as if* Betty was valuable merely as a means to your survival rather than as an end in herself. For even if you do not have this intention or attitude towards Betty, your act still has this objective moral feature since the permissibility of your act depends on the premise that Betty is valuable merely as a means. In Alcove, for example, the only thing that could make killing Betty permissible would be if, like your ability to access the alcove, Betty was valuable merely as a means to your survival. Since

---

[27] Hence the name 'Moral Status'.

she is not, you wrongly treat her as a mere means, regardless of your intentions or attitudes towards her.

G. Extending Moral Status

To round off, compare the following pair of cases

> *Liquid I.* To save your life, you must melt Betty into a liquid that you need to drink to live.

> *Liquid II.* To save your life, you must melt Betty so you can access a liquid you need to drink to live.

These two cases seem morally equivalent: you seem prohibited from killing Betty in both cases on largely the same grounds.[28] However, many would say that you are prohibited from killing Betty in Liquid I because it involves harmfully *using her* as a means of saving yourself.[29] But unlike in Liquid I, killing Betty in Liquid II does not involve harmfully using her as a means. This shows that this factor does not capture the essence of why killing Betty in Liquid I is impermissible.

By contrast, Moral Status provides a *unified* explanation of the prohibition in both cases. For in both cases melting Betty helps save you by making it such that you can access the liquid (a positive fact). Accordingly, Moral Status prohibits you from killing Betty in both cases on the same grounds. The fact that Betty in particular is useful to your survival in Liquid I thus

---

[28] Thomson (1991). Though it is enough for my purposes to have shown that Liquid 2 is worse than Well, even if one finds Liquid 1 worse for a different reason. For one could supplement Moral Status with the explanation that follows.

[29] For example, Quinn (1989), Frowe (2014), Ramakrishnan (2016).

has no explanatory role on this account. This further raises the plausibility of Moral Status as an explanation of the moral difference between Well and Alcove.[30]

I have just morally distinguished Well from Alcove. We must now explain why Moral Status permits you to kill Albert in Well. In the remainder of this paper, I argue that this permission is justified by appeal to Albert's enforceable duty to bear the costs of your rescue. I first defend this explanation against two objections in Section III. This will lay the groundwork for a positive argument for my view in Section IV.

III.   Eliminative Harming without Intentions

In this section, I address two objections against my explanation of Well.

A. The Intuitive Appeal of DDE

Here is the first objection. Consider the following pair of cases

> *Roughshod.* You are driving to the hospital for an emergency life-saving operation. However, Victim is lying in the only road that will get you there in time. Although Victim is not physically obstructing your path, she is too heavy to be moved aside. You can save yourself only by driving over Victim en route to the hospital, thereby killing her.

> *Obstruction.* The case is similar to Roughshod. However, you cannot simply drive over Victim on your way to the hospital. This is because her presence in the road is

---

[30] What if you must hurl Betty into a trolley to prevent it from killing you? Since the trolley's activity is independent of your act, this act is not wholly responsible for the most proximate cause of Betty's death (i.e., the trolley hitting her) and so we must exclude it from our explanation of how your killing Betty helps save you (i.e., by making a negative fact hold: that the trolley will not hit you). The most proximate cause of Betty's death for which you are wholly responsible is instead hurling her at the trolley, which helps save you by making it such that she hits the trolley (a positive fact). Moral Status thus prohibits killing her.

physically obstructing your path. You can save yourself only by blowing Victim to smithereens, thereby clearing the road ahead.

It seems impermissible to kill Victim in either case.[31] However, many find it intuitively plausible that killing Victim in Obstruction is *harder* to justify than in Roughshod. To my knowledge, the only discussed explanation of the moral difference between these two cases says that when other things are equal, intentionally harming an innocent person as a means is harder to justify than foreseeably but unintentionally harming them as a side-effect. This is the *Doctrine of Double Effect* (DDE).[32]

Whereas Roughshod involves killing Victim as a foreseen side-effect of saving yourself, Obstruction involves killing her as an intended means to this end.[33] Although it does not involve *opportunistically* intending her death, killing Victim in Obstruction nonetheless involves *eliminatively* intending her death.[34] Accordingly, DDE prohibits killing Victim in Obstruction but not in Roughshod. It thus captures the intuition that killing in Obstruction is worse than killing in Roughshod.

Many deontologists reject DDE.[35] Although they argue that we can still explain why opportunistic harming is worse than harming as a side-effect, deontologists have not offered

---

[31] As implied by Moral Status: Roughshod and Obstruction are structurally similar to Crush and Alcove, respectively.

[32] For example, Quinn (1989), McMahan (2009), Tadros (2015).

[33] Some hold that DDE prohibits killing Victim in Roughshod: you arguably intend *something* for Victim – namely, to run over her – which is harmful (Quinn 1989). However, I set this aside as DDE could not then distinguish Roughshod from Obstruction.

[34] Quinn (1989). Killing Victim to harvest her organs to cure your illness would involve opportunistically harming her: if she was absent, you could not save yourself. By contrast, if Victim was absent in Obstruction, you could still save yourself because the road ahead would be clear and so you must merely eliminate her.

[35] For example, Thomson (1991), Kamm (2007), Quong (2020).

an alternative explanation to DDE of why killing Victim in Obstruction is worse than in Roughshod. Accordingly, it seems that they are left with the counterintuitive implication that there is no moral difference between these two cases. It thus looks like they must endorse at least the eliminative aspect of DDE to avoid this troubling result.

Why is this problematic for my explanation of Well? Consider

> *Divert.* A trolley is heading towards you and will kill you if it hits you. You can save yourself only by diverting the trolley onto a side-track where it will hit and kill an innocent bystander.

As it stands, Moral Status permits you to kill the bystander in this case. For the most proximate cause of the bystander's death that is *also* relevant to your survival is the trolley occupying the side-track, and the least derivative explanation of how this helps save you is by making it such that the trolley will not hit you (a negative fact).[36]

Intuitions vary regarding the permissibility of killing the bystander in Divert. Suppose you are *permitted* to kill the bystander in this case.[37] Nevertheless, whereas the bystander will be killed as a foreseen side-effect of saving yourself in Divert, Albert will be killed as an eliminatively intended means of saving yourself in Well.[38] So if DDE is sufficiently stringent, the moral wedge it drives between Well and Divert could render it impermissible for you to kill Albert in Well, even if you are permitted to kill the bystander in Divert.[39]

In what follows, I argue that the factor that correctly distinguishes Roughshod from Obstruction does not count against killing Albert in Well.

B. Against DDE

---

[36] This parallels the reasoning in footnote 24.

[37] Quong (2009), Kamm (2015). I consider what to say if one does not share this intuition shortly.

[38] *Eliminatively* since you would be safe in his absence.

[39] For example, Kagan (1989), Otsuka (1994), McMahan (2018), Tadros (2020).

To start, DDE does not get to the heart of the intuitive moral difference between Roughshod and Obstruction. Consider

> *Boulder.* The case is similar to Obstruction. However, this time a massive boulder blocking your path. Victim is lying out of your way to the side of the road next to the boulder. You can save yourself only by blowing the boulder to smithereens (along with Victim who will be caught in the blast), thereby clearing the road ahead.

Intuitively, this case is relevantly like Obstruction. However, DDE cannot accommodate this intuition. This is because Boulder involves killing Victim as a side-effect of blowing up the boulder to save yourself. In this respect, Boulder is relevantly like Roughshod.[40]

Accordingly, even if it can distinguish morally between Roughshod and Obstruction, DDE cannot distinguish Roughshod from Boulder. This suggests that it is not the fact that you kill Victim as an intended means in Obstruction but as a side-effect in Roughshod that explains the intuitive difference between these two cases.

I will next offer an alternative explanation of the moral difference between Roughshod and Obstruction that also gets the intuitively correct result in Boulder.

C. The Alternative to DDE

Each of us is required to incur certain costs in the course of rescuing a person. I take it that such duties are *enforceable*. This means that even if you are unable to perform a rescue yourself (say, because you are paralysed), a third-party is nonetheless permitted to make you bear the cost that you would otherwise be required to incur in the course of performing the rescue.[41]

---

[40] It is true that unlike Roughshod and akin to Divert, Boulder does not involve intending *anything* for Victim. Nevertheless, Moral Status still prohibits killing her in Boulder. In these respects, Boulder is relevantly like driving over an acid pool en route to the hospital that will fatally splash Victim at your side (Kamm 2007: 132), which still seems akin to Roughshod.

[41] For example, Tadros (2011).

I will now introduce a new factor that has not yet been considered by philosophers that can alter the strength of one's duty to rescue. Crucially, I will argue that this factor also explains why Victim's enforceable duty is *stronger* in Roughshod than in Obstruction, and so why you are permitted to impose *greater* harm on Victim in Roughshod than in Obstruction.

To illustrate the factor that I have in mind, consider the following pair of rescue cases

> *Rescue I.* A trolley is heading towards Victim and will kill her if it hits her. You can save Victim only by gently pushing her out of the way. However, because Victim is very heavy, you will injure yourself to some degree while pushing her.
>
> *Rescue II.* The case is similar to Rescue I. But since Victim is very light, pushing her will itself be pretty much costless to you. Unfortunately, Victim is just on the other side of a large boulder blocking your path. You must therefore first roll the boulder out of the way, thereby clearing the path ahead, after which you can proceed to Victim unimpeded. But because the boulder is very heavy, you will injure yourself to some degree while moving it.

Intuitively, you are required to incur greater costs to help Victim in Rescue I than you are required to incur to help Victim in Rescue II. For example, while you might be required to move Victim out of the path of the trolley at the expense of breaking your arms in Rescue I, it does not seem intuitively plausible that you must merely clear your path to Victim by rolling the boulder out of the way at the expense of breaking your arms in Rescue II. (One can adjust the costs to fit one's own intuitions: the thought is just that there is some cost C such that you are required to help Victim at cost C in Rescue I but not required to help Victim at cost C in Rescue II.)

Drawing on this intuitive difference between Rescue I and Rescue II, it seems that the strength of one's duty to rescue a person is affected by *the way in which* the costs of the rescue arise to you.

In Rescue I, the costs of helping Victim arise from *moving her out of the path of the trolley*. It is intuitively plausible that moving Victim away from the trolley counts as part of *rescuing her*. By contrast, in Rescue II, the costs of helping Victim arise from *removing an obstacle* that is preventing you from moving Victim to safety. Although doing so is necessary to later move Victim away from the trolley, removing the obstacle does not itself seem to count as part of rescuing Victim. Rather, it seems that when you reach the boulder, you must suspend the rescue while you take care of the obstacle to clear your path, after which you can resume the rescue and proceed to move Victim to safety.

Our intuitive judgements about whether something counts as part of rescuing Victim seem to be driven by whether it is a *causal means* of eliminating the threat of the trolley to Victim. For example, moving Victim away from the trolley is a causal means of eliminating the threat to her (i.e., it *causes* her to be out of its path), and so this is why it seems to count as part of rescuing Victim. Similarly, running to Victim's location, for example, also intuitively seems to count as part of the rescue since it *causes* you to be where you can subsequently *cause* Victim to be away from the trolley.

By contrast, removing the boulder obstructing your path is *not* a causal means of eliminating the threat to Victim (i.e., it does not *cause* her to be away from the trolley). Rather, removing the boulder is a means of *eliminating a potential defeater of* the causal means of eliminating the threat to Victim, and so it is *non-causally* related to eliminating this threat. This is why removing the boulder does not seem to count as part of rescuing Victim.[42]

---

[42] Consider the removal of bone obstructing a surgeon's access to a defective heart. This might seem part of rescuing the patient, even though it is non-causally related to eliminating the threat (i.e., acting on the defective heart itself). However, the surgery that eliminates the threat is a *complex act*: it involves acting both on the heart *and* the patient to access it.

This means that in Rescue I, the costs of helping Victim arise from *doing good*, whereas in Rescue II, they arise from *removing a barrier* to a beneficial sequence. This shows that doing good is *better* than removing a barrier to a beneficial sequence, in the sense that the costs that you are required to bear to do good are greater than those you are required to bear to remove a barrier to a beneficial sequence.[43] This mirrors a distinction that seems plausible with respect to harm.

For just as doing good seems better than removing a barrier to a beneficial sequence, doing *harm* seems *worse* than removing a barrier to a *harmful* sequence in this cost-requiring sense. For imagine that you can save yourself only by releasing a gas that will cure you but leak into the next room and poison an innocent bystander. Intuitively, this seems harder to justify than if the gas will simply destroy a pill that would otherwise be administered to cure an innocent bystander's illness.[44]

Since the distinction between doing good (harm) and removing a barrier to a beneficial (harmful) sequence seems morally significant, perhaps this is partly because it is an instance of a more general distinction between *doing* and *removing a barrier*. Perhaps actively causing a sequence that produces an outcome more closely ties you to it, thereby making you more responsible for it than if the sequence was already in train independently of the act under consideration.[45]

---

[43] One might object that it does not seem plausible that there is stronger reason to do good than to remove a barrier to a beneficial sequence, in the *conflict-of-duty* sense of 'stronger reason'. For suppose you can save one person by doing good or another by removing a barrier to a beneficial sequence. You seem required to flip a coin (see also MacAskill & Mogensen 2021: 5). However, we could explain this by saying that the distinction is insufficiently strong to defeat the requiring reason to flip a coin present in conflict-of-duty cases.

[44] While removing an *operative* barrier that is stabilising a sequence might seem morally equivalent to doing (McMahan 1993, Hanna 2015), the barriers we are considering here are all *inoperative*.

[45] McMahan (2002: 461) suggests a similar point with respect to doing and *allowing*.

Let us take stock. I have just offered an account of the intuitive difference between Rescue I and Rescue II. I argued that whereas the costs of helping Victim in Rescue I arise from the rescue itself, the costs of helping Victim in Rescue II do not. This factor captures the intuitive judgement that you are required to incur greater costs to help Victim in Rescue I than in Rescue II.

Return now to Roughshod and Obstruction. In Roughshod, Victim incurs the costs of your rescue in a structurally similar manner to how the costs of Victim's rescue arise to you in Rescue I. This is because Victim is killed by you driving to the hospital, and it is intuitively plausible that driving to the hospital counts as part of rescuing yourself. For driving to the hospital is a *causal means* of reaching the hospital and thereby undergoing the procedure that eliminates your life-threatening condition.

In Obstruction, by contrast, Victim incurs the costs of your rescue in a structurally similar manner to how the costs of Victim's rescue arise to you in Rescue II. This is because Victim is killed by you removing an obstacle (i.e., the obstacle that she poses), and removing this obstacle does not intuitively seem to count as part of rescuing yourself. Rather, removing the obstacle that Victim poses is a means of *eliminating a potential defeater of* you driving to the hospital and thereby undergoing the life-saving procedure, and so it is *non-causally* related to eliminating the threat you face.

We are now in a position to explain the moral difference between Roughshod and Obstruction. For just as you are required to bear greater costs to help Victim in Rescue I than in Rescue II, so too is Victim required to bear greater costs in the course of your rescue in Roughshod than in Obstruction, and for precisely the same reason.

Accordingly, Victim is under a *stronger enforceable duty* to bear the costs of your rescue in Roughshod than in Obstruction. You are therefore permitted to impose *greater* harm on Victim in Roughshod than in Obstruction. This explains why killing Victim in Obstruction is harder

to justify than in Roughshod.[46] Importantly, the fact that Victim is killed as an intended means in Obstruction but as a side-effect in Roughshod is morally insignificant on my account. So this is an alternative explanation to DDE.

It is true that Victim is not under an *all-things-considered* enforceable duty to bear such costs of your rescue in either Roughshod or Obstruction. For imposing this cost on Victim also involves killing her, which is impermissible in both cases. Nevertheless, Victim is under a *pro tanto* enforceable duty to bear such costs: her duty to bear the costs of your rescue still *contributes* to the proposed moral justification for harming her. The fact that Victim's duty is stronger in Roughshod than in Obstruction supports a corresponding difference in the justifiability of killing Victim.[47]

I have just offered an alternative explanation to DDE of the moral difference between Roughshod and Obstruction. This explanation is extensionally superior to DDE.

For unlike DDE, my account gets the intuitively correct result that Boulder is relevantly like Obstruction. This is because Victim is killed by you removing an obstacle (i.e., the boulder) to eliminating the threat to your life in Boulder, and so does not incur these costs as part of your rescue itself. So just as you are permitted to impose greater harm on Victim in Roughshod than in Obstruction because of her stronger enforceable duty, you are also permitted to impose

---

[46] One might object that the difference in justifiability between Obstruction and Roughshod seems *greater* than the difference in cost you must bear in Rescue I and Rescue II. However, it is plausible that DDA's strength increases *disproportionately* with the magnitude of non-liable harm imposed. Accordingly, Obstruction is disproportionately worse than Roughshod.

[47] Footnote 62 offers an alternative defence of my account that does not appeal to enforceable duties and so bypasses this concern altogether.

greater harm on Victim in Roughshod than in Boulder, and for precisely the same reason. Accordingly, we should reject DDE in favour of my account.[48]

D. Self-Defence

Return finally to Well. What does my account of the moral difference between Roughshod and Obstruction say about this case? Consider that Albert incurs the costs of your rescue in a structurally similar manner to how the costs of Victim's rescue arise to you in Rescue I rather than in Rescue II. This is because killing Albert is a causal means of eliminating the threat to your life: namely, *Albert himself*. In this respect, Well is relevantly like Roughshod rather than Obstruction. So unlike DDE, my account of the moral difference between Roughshod and Obstruction does not say that killing Albert in Well is worse than killing the bystander in Divert.[49]

---

[48] One might object that Obstruction still seems slightly worse than Boulder. However, we can explain this by saying that intending *unjustified harm* (i.e., "aiming at *evil*" (Nagel (1986: 181)) is worse than foreseeing it. To illustrate: Obstruction does not seem worse than Boulder if the number of lives at stake in Obstruction would justify killing Victim in Boulder. Yet since I argue that killing Albert in Well is justified independently of your intentions, this view does count against killing him.

[49] Consider two further respects in which Well might appear worse than Divert. Firstly, whereas you kill Albert by *creating a threat* (i.e., your ray), you kill the bystander by *redistributing a pre-existing threat* (i.e., the trolley). Yet some feel that creating a threat is *worse* than redistributing a pre-existing threat (Thomson 1986). For example, even if you seem permitted to divert the trolley onto the bystander, you might seem prohibited from exploding the trolley foreseeing the bystander will be caught in the blast (Otsuka 1994, McMahan 2018). However, we can explain this because both Well and Divert involve killing by *affecting* a pre-existing threat (i.e., by *acting on* Albert and the trolley, respectively). And the property of redistributing a pre-existing threat seems subsumed under this broader moral category. To illustrate: suppose you must destroy a lethal life-support machine hurtling towards you that is also sustaining a bystander, thereby killing her by removing this support. This case seems relevantly like Divert. But you do not kill the bystander by redistributing a pre-existing threat in this case. For the machine does not kill her. Nevertheless, you kill her by *affecting* a pre-existing threat (i.e., by acting on the machine with

E. Mere Side-Effect

Having addressed the first objection to my explanation of Well, we can now answer the second. Suppose you are *prohibited* from killing the bystander in Divert.[50] How can we explain this if Moral Status permits you to kill Albert in Well?

Consider that whereas Albert incurs the costs of your rescue as part of the causal means of eliminating the threat to your life, the bystander is hit by the trolley causally downstream from it being diverted away from you, and so incurs these costs as an *after-effect* of your threat having already been eliminated. So unlike Albert, the bystander will not incur the costs of your rescue from the rescue *itself*. In this respect, Divert is relevantly like Obstruction: just as removing a barrier to a beneficial sequence does not count as part of the rescue itself, neither does the causally downstream effects of the rescue.

Accordingly, Albert is under a stronger enforceable duty to bear the costs of your rescue in Well than the bystander is under in Divert. You are therefore permitted to impose greater harm on Albert than on the bystander. This explains why killing the bystander in Divert could be

---

your weapon). Secondly, suppose *ten* innocent threats are falling towards you in Well. You might seem permitted to kill all ten, even though you seem prohibited from diverting a trolley away from yourself onto ten bystanders (McMahan 1994). However, we can explain this by saying that since each of the ten individually threaten your life, a sequence of pairwise comparisons between yourself and each threat permits you to kill each of them. Alternatively, if each individually contribute to a *single threat*, we could say that what matters is the overall threat they *contribute to*. Accordingly, the case is relevantly like facing ten individual life-threats. By contrast, since we cannot divide the threat of the trolley into ten separate threats in this way, we cannot similarly justify diverting the trolley away from yourself onto ten bystanders.

[50] Frowe (2014), Doggett (2018), Tadros (2020).

impermissible even if killing Albert in Well is permissible. We must therefore amend the permissive arm of Moral Status to accommodate this factor.[51]

Since intuitions vary about Divert, I ultimately leave it open whether killing in this case is *sufficiently worse* than killing Albert in Well to render killing the bystander impermissible. I have argued that Moral Status can accommodate either position.

Overall, both objections against my explanation of Well fail. We can capture the intuitions driving each objection (i.e., that Obstruction is worse than Roughshod, and that you are prohibited from killing the bystander in Divert) but in a way that does not undermine my position.[52]

IV.   The Permission

I shall now provide a positive argument for why Moral Status permits you to kill Albert in Well. The previous section laid the groundwork for this explanation.

A. Agent-Neutral

There is a moral presumption against killing Albert in Well. For there is a moral asymmetry between doing and allowing harm, such that saving an innocent person does not usually justify

---

[51] Moral Status still explains why you may kill Albert by harmlessly deflecting him to your side foreseeing he will subsequently fatally hit a hard surface *causally downstream* from your rescue (discussed in footnote 24). For by deflecting him, Albert is thereby still *affected by* (and so incurs) what will result in his death *as part of your rescue itself*. If we amend Rescue I to reflect this factor (for example, imagine that while affecting yourself as a means of moving Victim is itself costless, it will result in exposing yourself to the oncoming trolley that will subsequently hit and injure you to some degree), it makes no intuitive difference to the costs you are required to bear to help Victim. By contrast, since the bystander in Divert is only affected by what will result in her death when the trolley actually hits her, this explanation does not similarly apply to killing her.

[52] If one does not share my intuitions in Boulder or Rescue I and Rescue II, footnote 62 offers alternative ways of framing my responses to these two objections.

killing another. Why does Moral Status nevertheless permit you to override this presumption to save yourself in Well?

Consider that there is a moral asymmetry between yourself and Albert. As we have already seen, when you kill Albert, he will incur the costs of your rescue as part of the rescue itself. But what about the costs *you* are required to bear for *Albert's sake*? When considering how the factor under consideration affects the costs one is required to bear to *avoid harming*, we should view *refraining from harming* as the relevant analogue to aiding a person. So understood, consider how you will incur the costs of Albert's 'rescue' (from the harm you could impose on him).

If you allow Albert's body to hit you instead of killing him, you will incur the costs of Albert's rescue as an *after-effect* of his rescue, and so not as part of his rescue itself. For you incur the costs of refraining from killing Albert only when his body actually hits you, which is causally downstream from your abstaining to kill him (i.e., the last opportunity to vaporise him). You therefore incur the costs of Albert's rescue in a structurally similar manner to how you incur the costs of helping Victim in Rescue II.

Accordingly, Albert's enforceable duty to bear the costs of your rescue is *stronger* than your enforceable duty to bear the costs of Albert's rescue. It is true that just like an ordinary bystander, Albert is not under an enforceable duty to incur the full cost of *death* to save you. But it means that when we subtract the harm each is under an enforceable duty to bear to rescue the other, Albert being killed is a *lesser-evil* relative to your life being saved.

Suppose that an impartial third-party is *permitted* to kill Albert on your behalf.[53] This shows that the disparity between the strength of yours and Albert's enforceable duties in Well is great

---

[53] I consider what to say if one does not share this intuition shortly.

enough to generate a *lesser-evil justification* for killing Albert. For since a lesser-evil justification is agent-neutral, it is available to yourself and impartial third-parties alike.

Crucially, we cannot similarly justify killing the bystander in Divert. For you and the bystander are symmetrically placed with respect to how you will incur the costs of each other's rescue: just as you will be hit as an after-effect of the trolley not being diverted (i.e., the last opportunity to divert it), the bystander will be hit as an after-effect of diverting the trolley. Accordingly, there is no difference in the cost each is required to bear to rescue the other, and so there is no lesser-evil justification for killing the bystander. This explains why an impartial third-party is prohibited from killing the bystander in Divert, even if they are permitted to kill Albert in Well.[54]

The fact that Albert will kill you in Well thus plays no explanatory role on my view. As previously argued against Victim's Role, we cannot appeal to this factor to explain Well. The difference between Well and Divert is instead explained by *how* Albert and the bystander will incur the costs of your rescue.[55]

---

[54] Similar remarks apply to Victor Tadros's (2020: 195) *Flip.* We could still explain why you may save *yourself* in Divert and Flip by invoking the agent-relative value of your life (an option I explore shortly). We can also accommodate Quong's (2020: 63-64) *Broken Arms*. Briefly, this is because the constraint against doing harm prohibits imposing *pro tanto* harm (to illustrate: though it will leave you no worse-off overall, saving your right hand does not justify overriding the constraint against imposing the pro tanto harm of cutting off your left hand against your will). Yet even if the *overall* non-liable harm imposed is equivalent in Broken Arms and Well (i.e., <death minus broken arms>), the *pro tanto* harm imposed in Broken Arms (i.e., <*death*> as a means of preventing <broken arms>) is *greater* than in Well (i.e., <*death minus broken arms*>). This explains why you could be prohibited from killing the bystander in Broken Arms, even if there is a lesser-evil justification for killing Albert in Well.

[55] It might seem that if he had the option, Albert would be required to incur greater costs than a bystander to prevent Albert hitting you (for example, Frowe 2014). But we could say that unlike the bystander, Albert would

B. Agent-Relative

Now suppose an impartial third-party is *prohibited* from killing Albert in Well. This shows that your permission to kill Albert is instead partly justified by an *agent-relative prerogative*: a permission to act from a partial perspective, thus favouring your own life disproportionately to its impartial significance.[56] For positing this prerogative is now necessary to adequately capture our intuitions in Well.

Importantly, we must still invoke Albert's enforceable duty to bear the costs of your rescue on this account. For we must still explain why the factor that correctly distinguishes Roughshod from Obstruction does not count against killing Albert, and why you could nonetheless be prohibited from killing the bystander to save yourself in Divert. It is just that the asymmetry between you and Albert considered previously is too weak to generate a lesser-evil justification for killing him.

Helen Frowe (2021) has recently argued that we can explain the permission to forgo saving others at extraordinary costs to oneself without invoking agent-relative prerogatives. On Frowe's view, this permission is instead explained by the agent-neutral fact that we have only a limited claim that others use themselves as a means for our sake. This means that if an impartial third-party is prohibited from killing Albert in Well, you are similarly prohibited. For Frowe's view denies that you have an agent-relative prerogative to favour yourself over Albert.

---

now be morally responsible for killing you. Accordingly, unlike in Well, he is not an *innocent* threat in this case. He has therefore forfeited his right not to be harmed to a certain degree, and so has also forfeited his *prerogative* to avoid such harm. For it is plausible that one's prerogative is closely tied to one's right not to be harmed, such that you retrain the former only if you retain the latter (Quong 2009). This is why Albert's prerogative is weaker than a mere bystander's in this case. It thus does not show that innocent Albert's enforceable duty in Well is explained by the fact that he will kill you.

[56] For example, Scheffler (1982), Davis (1984), Quong (2009; 2020).

And since refraining from doing harm does not involve using oneself as a means, it also denies an extensionally similar permission to kill him justified by agent-neutral considerations.

Against Frowe, consider the following pair of cases

> *Hammer I.* A lethal trolley is heading towards five people, and another is heading towards Victim. You can stop either trolley only by throwing Victim's random hammer at it. You cannot stop both.

> *Hammer II.* The case is similar to Hammer I. However, Victim can either use her hammer to save herself, or leave it be knowing you will then use it to save the five instead. She cannot do both.

You seem permitted to save the five using Victim's hammer in Hammer I.[57] Nevertheless, Victim seems permitted to use it for herself in Hammer II.[58] By contrast, Frowe's view implies that Victim is permitted to use her hammer to save herself in Hammer II only if you are prohibited from using it to save the five in Hammer I. Accordingly, we must reject Frowe's account.[59]

We can explain these judgements only by invoking agent-relative prerogatives. For even if you are permitted to save the five in Hammer I, Victim nonetheless has an agent-relative

---

[57] See also Bowen (2021). It might nonetheless seem impermissible for you to save the five in Hammer I if it was not Victim's *random* hammer but instead bought specifically for her protection.

[58] I take it that since she *owns* the hammer, Victim removing the barrier it could provide to the five in Hammer II is morally equivalent to *letting them die*. If Victim did not own the hammer, her taking it might seem morally worse than letting die and so impermissible (Woollard 2015).

[59] Similar remarks apply to justifications of *prerogatives* that appeal to one's right against being harmfully used (for example, Muñoz 2020) which seemingly exclude prerogatives to do harm.

prerogative to favour herself in Hammer II. Accordingly, our appeal to agent-relative prerogatives in Well remains intact.[60]

In the end, I leave it open whether a third-party is permitted to kill Albert in Well, and so whether your permission to kill Albert is justified by agent-neutral or also agent-relative considerations. Since intuitions vary regarding third-party permission, I have argued that Moral Status can accommodate either position.[61]

Crucially, though, neither view permits you to kill Betty in Alcove. The agent-neutral (or agent-relative) importance of your life is enough to override the constraint against killing Albert. But it is insufficient to override the combined force of the constraints against killing

---

[60] Frowe also argues that agent-relative prerogatives are either too strong or too weak to accommodate our intuitions in cases: for example, they either permit refraining from saving a leg at the cost of breaking your arm or do not permit refraining from saving a life at the cost of your leg – assuming the ratio between the magnitude of harm prevented and agent-relative cost to you in each case is equivalent. However, we can explain this by saying that the strength of one's prerogative increases *disproportionately* with the magnitude of cost incurred. While Frowe also objects to existing arguments for prerogatives to do harm in particular, she does not address the conditional argument defended here. For a powerful defence of prerogatives to do harm against objections preceding Frowe, see Quong (2020: 85-90).

[61] Does Moral Status permit Albert shooting *you* to prevent you from vaporising Albert? I think so. Since you are the threat to him, Albert shooting you structurally mirrors you vaporising Albert in the manner described previously. Accordingly, Albert shooting you is a lesser-evil *relative to you vaporising him*. This is the case, even though you vaporising Albert is a lesser-evil *relative to his body hitting you*. If we viewed the case as simply *your life versus Albert's*, we would blur the subtle distinction that generates a lesser-evil on my account. You and Albert thus have equivalent permissions to kill each other in Well on the agent-neutral view. Similarly, on the agent-relative view.

Betty *and* against harmfully treating her as a means. This is evidenced by the intuitive permissibility of killing Albert in Well but not Betty in Alcove.[62]

V.   Conclusion

In sum, I have argued that Moral Status explains your permission to kill an innocent threat. For given the manner in which he will incur the costs of your rescue, Albert is under a strict enforceable duty to bear such costs in Well. In Section II, I argued that unlike existing explanations, Moral Status gets to the heart of the moral asymmetry between Well and Alcove. Accordingly, you are prohibited from killing Betty in Alcove because it involves harmfully treating her as a means, whereas killing Albert in Well does not. In Section III, I considered two objections against my explanation of Well. I argued that we can capture the intuitions driving each objection, but in a way that does not undermine my position. For the factor that correctly distinguishes Roughshod from Obstruction does not count against killing Albert, but it does count against killing the bystander in Divert. This is because one has a stronger enforceable duty to bear the costs of a rescue as part of the rescue itself. In Section IV, I provided a positive argument for why Moral Status permits you to kill Albert in Well. The exclusively agent-neutral account extends this permission to an impartial third-party, while the appeal to agent-relative considerations limits this permission only to yourself.

---

[62] What if one does not share my intuitions in Boulder or Rescue I and Rescue II? Regarding Boulder, we could alternatively say that intending *unjustified harm* is worse than foreseeing it (as suggested in footnote 48). This factor would count against Obstruction, but not Roughshod, Boulder, or Well. Regarding Rescue I and Rescue II, we could alternatively say that killing outside of a rescue itself is worse than killing as part of a rescue itself as a matter of *mode of agency* only. For the latter killings have the right-making property of doing good that the former killings lack. This factor would count against Obstruction, Boulder, and Divert, but not Roughshod or Well. Although this route would preclude my agent-neutral explanation of Well, the agent-relative explanation would remain available.

References

Baker, Stephen, and Jago, Mark (2012). "Being Positive About Negative Facts." *Philosophy and Phenomenological Research* 85, no. 1: 117–38.

Bennett, Jonathan (1995). *The Act Itself*. New York: Oxford University Press.

Bowen, Joseph (2021). Quong, Jonathan. The Morality of Defensive Force. Oxford: Oxford University Press, 2020. *Ethics* 131 (3): 625-630.

Davis, Nancy (1984). Abortion and Self-Defense. *Philosophy and Public Affairs* 13 (3):175-207.

Doggett, Tyler (2018). Killing Innocent People. *Noûs* 52 (3): 645-666.

Fitzpatrick, William J. (2006). The Intend/Foresee Distinction and the Problem of "Closeness". *Philosophical Studies* 128 (3): 585-617.

Foot, Philippa (1967). The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review* 5: 5-15.

Frick, Johann (2015). Contractualism and Social Risk. *Philosophy and Public Affairs* 43 (3): 175-223.

Frick, Johann (2015). Contractualism and Social Risk. *Philosophy and Public Affairs* 43 (3):175-223.

Frowe, Helen (2014). *Defensive Killing*. Oxford, UK: Oxford University Press.

Frowe, Helen (2021). The Limited Use View of the Duty to Save. In Peter Vallentyne, Steven Wall & David Sobel (eds.), *Oxford Studies in Political Philosophy Vol. 7*. New York, NY, USA: pp. 66-99.

Gordon-Solmon, Kerah (2019). Should Contractualists Decompose? *Philosophy and Public Affairs* 47 (3): 259-287.

Gordon-Solmon, Kerah (2019). Should Contractualists Decompose? *Philosophy and Public Affairs* 47 (3):259-287.

Hanna, Jason (2015). Enabling Harm, Doing Harm, and Undoing One's Own Behavior. *Ethics* 126 (1): 68-90.

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

Kamm, F. M. (ed.) (2011). *Ethics for Enemies: Terror, Torture, and War*. New York: Oxford University Press.

Kamm, Frances (2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.

Kamm, Frances Myrna (1992). *Creation and Abortion: A Study in Moral and Legal Philosophy*. New York: Oxford University Press.

Kamm, Frances Myrna (ed.) (2015). *The Trolley Problem Mysteries*. New York: OUP USA.

M. Kamm, F. (1993). *Morality, Mortality: Volume 1*. Oxford University Press USA.

McMahan, Jeff (1993). Killing, Letting Die, and Withdrawing Aid. *Ethics* 103 (2): 250-279.

McMahan, Jeff (1994). Self-defense and the Problem of the Innocent Attacker. *Ethics* 104 (2): 252-290.

McMahan, Jeff (2002). *The Ethics of Killing: Problems at the Margins of Life*. New York, US: OUP USA

McMahan, Jeff (2005). The Basis of Moral Liability to Defensive Killing. *Philosophical Issues* 15 (1): 386–405.

McMahan, Jeff (2009). Intention, Permissibility, Terrorism, and War. *Philosophical Perspectives* 23 (1): 345-372.

McMahan, Jeff (2009). *Killing in War*. New York: Oxford University Press.

McMahan, Jeff (2018). Nonresponsible Killers. *Journal of Moral Philosophy* 15 (6): 651-682.

Mogensen, Andreas & MacAskill, William (2021). The Paralysis Argument. *Philosophers' Imprint* 21 (15).

Muñoz, Daniel (2020). From Rights to Prerogatives. *Philosophy and Phenomenological Research* 102 (3): 608-623.

Nagel, Thomas (1972). War and Massacre. *Philosophy and Public Affairs* 1 (2): 123-144.

Nagel, Thomas (1986). *The View From Nowhere*. New York: Oxford University Press.

Otsuka, Michael (1994). Killing the Innocent in Self-Defense. *Philosophy and Public Affairs* 23 (1): 74-94.

Otsuka, Michael (2016). The Moral Responsibility Account of Liability to Defensive Killing. In Christian Coons & Michael Weber (eds.), *The Ethics of Self-Defense*. Oxford, U.K.: Oxford University Press.

Quinn, Warren S. (1989). Actions, Intentions, and Consequences: The Doctrine of Double Effect. *Philosophy and Public Affairs* 18 (4): 334-351.

Quong, Jonathan (2009). Killing in Self-Defense. *Ethics* 119 (3): 507-537.

Quong, Jonathan (2020). *The Morality of Defensive Force*. Oxford University Press.

Ramakrishnan, Ketan H. (2016). Treating People as Tools. *Philosophy and Public Affairs* 44 (2): 133-165.

Scheffler, Samuel (1982). *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. New York: Oxford University Press.

Tadros, Victor (2011). *The Ends of Harm: The Moral Foundations of Criminal Law*. New York: Oxford University Press.

Tadros, Victor (2015). Wrongful Intentions without Closeness. *Philosophy and Public Affairs* 43 (1): 52-74.

Tadros, Victor (2020). *To Do, to Die, to Reason Why: Individual Ethics in War*. Oxford University Press.

Thomson, Judith Jarvis (1986). The Trolley Problem. In William Parent (ed.), *Rights, Restitution, and Risk: Essays in Moral Theory*. Cambridge, Mass.: Harvard University Press. Edited by William Parent.

Thomson, Judith Jarvis (1991). Self-defense. *Philosophy and Public Affairs* 20 (4): 283-310.

Woollard, Fiona (2015). *Doing and Allowing Harm*. Oxford, GB: Oxford University Press.

Chapter III: What Is the Harm of Preventing Predation?

I. Introduction

Predation in the wild causes immense animal suffering and death. If you can easily rescue prey animals from predators on a wide scale, are you morally required to do so?

Many think not. This is because they argue that given our current ignorance of the long term ecological effects, preventing predation on a wide scale will probably make things worse overall. For example, unless predators are given an alternative food source, they might starve to death. The herbivore population could also expand beyond the ability of the habitat to sustain them. So instead of being killed by predators, many more herbivores might die from starvation and disease. Many therefore hold that at least in practice, you are prohibited from interfering with predation on a wide scale.[1]

This challenge is strengthened by acknowledging that even if the benefits will outweigh the harms, interfering with predation still seems impermissible. For whereas failing to intervene involves *allowing* prey animals to be killed by predators, interfering with predation appears to involve *doing* harm to other wild animals. Yet doing harm seems much harder to justify than allowing harm to occur. So saving many prey might still be insufficient to justify harming even fewer wild animals.[2]

While not all share this pessimistic outlook on the expected consequences of interfering with predation, most accept the *moral* judgement that if such empirical claims are correct, then interfering with predation is morally impermissible.

---

[1] For example, Donaldson and Kymlicka (2011), McMahan (2016).

[2] For example, Johannsen (2020).

But even if it will certainly prove disastrous, in my view you are nonetheless morally required to intervene. This is because interfering with predation is morally akin to certain cases in which you seem required to rescue despite this making things worse overall.

In Section II, I argue that rescuing prey animals from predators is relevantly like rescuing a person who will otherwise be used as a means of saving a greater number. In Section III, I argue that rescuing prey animals from predators is relevantly like rescuing an identified victim at the expense of a greater number of merely statistical victims.

Overall, then, we must not wait until our understanding of the ecosystem advances. We are morally required to interfere with predation immediately.[3]

II.   The Causal Structure of Predation

In this section, I will argue that interfering with predation is relevantly like rescuing a person who will otherwise be used as a means of saving a greater number.

A. Harmful Use

Suppose that a nearby zebra is about to be killed by a lion. You can easily save her by firing your gun into the air, thereby scaring the lion away. But suppose you know that due to its long term ecological effects, rescuing the zebra will somehow lead to the deaths of five other wild animals.

Many hold that you are prohibited from rescuing the zebra from the lion in this scenario. This is because it seems plausible that when other things are equal, you are morally required to maximise the good. For example, if you can save either one person or five, you are morally

---

[3] I assume throughout that other things are equal. I therefore bracket out more principled moral objections against interfering with predation on a wide scale. For example, Donaldson and Kymlicka (2011). I also set aside the personal cost involved that might render intervening supererogatory as opposed to morally required.

required to save the five over the one. Similarly, you seem prohibited from rescuing the zebra from the lion if this will bring about more deaths overall.

Additionally, many hold that whereas failing to intervene involves foreseeably *allowing* the zebra to be killed by the lion, rescuing the zebra involves foreseeably *killing* the five other wild animals. Yet doing harm seems much harder to justify than merely allowing harm to occur. Saving five lives seems to justify letting one person die. But it does not seem to justify releasing a gas that will cure five people of a deadly disease yet leak into the next room and fatally poison an innocent bystander. So even if the benefits of interfering with predation outweigh the harms imposed, it still seems impermissible to intervene. So saving just *one* zebra seems clearly insufficient to justify *killing five* other wild animals.

But in my view, this characterisation of the case is mistaken. It is instead relevantly like:

> *Guinea Pig*. Victim is suffering from a life-threatening disease. You can easily treat her disease by administering a cure. But if you withhold treatment, a group of scientists can observe the fatal progression of her disease. They can thereby gain the medical knowledge necessary to save five other people from a similar condition.[4]

Although it will make things worse overall, you are intuitively required to save Victim in this case. Here is a plausible explanation of why.

By treating Victim, you thereby withdraw aid from the five. This is because you prevent Victim's body from being used to cure them. If Victim were to withdraw this aid from the five herself (say, by self-administering her own medication), she would prevent herself from saving the five and so would merely allow them to die. Although in Guinea Pig you withdraw this aid instead of Victim, she would presumably want to be saved. So by saving her you withdraw this aid *on her behalf*. This explains why withdrawing aid from the five seems morally equivalent

---

[4] For example, Quinn (1989).

to merely letting them die: in this respect, it is just as if Victim were to withdraw this aid herself.[5]

But whereas saving Victim involves letting the five die as a *foreseen side-effect*, failing to save Victim involves harmfully *using her as a means* of saving the five. This is because Victim's body will be used to observe the fatal progression of the disease to cure the five only if you let her die. And since harmfully using someone seems much harder to justify than foreseeing harm as a side-effect, saving the five does not justify letting Victim die in this case. Overall, I take it that this explains why you are morally required to save Victim in Guinea Pig.[6]

This explanation also implies that you are morally required to rescue the zebra from the lion.

We are considering the harms of *preventing predation*. This means that rescuing the zebra from the lion will only harm the five other wild animals *merely because* the lion will not kill the zebra. But if the five will die merely because predation will not occur, this can only be because the five face an independent threat that predation would somehow eliminate. So even if we lack more detailed information about the case, we have an insight into its general structure. In particular, we know that rescuing the zebra from the lion will harm the five only by *withdrawing aid from them* (i.e., aid in the form of the lion killing the zebra).

I will now argue that the zebra is useful in bringing it about that the lion kills her. So just as Victim is useful to saving the five in Guinea Pig, the zebra is also useful to saving the five other wild animals.

Consider the *way in which* the lion will kill the zebra.

Firstly, the lion will make use of various parts of the zebra's body to help him kill her. For example, the lion might mount the zebra's back, latch onto her tail, pin her torso to the ground, and so on, all as a means killing her. Even if we cannot know in advance precisely how the lion

---

[5] For example, McMahan (1993), Woollard (2015).

[6] For example, Quinn (1989), Ramakrishnan (2016), Quong (2020).

will kill the zebra, we know that the lion will adopt these general sorts of methods to kill her. This is one way in which the zebra is useful in bringing it about that the lion will kill her.

Here is a second. The lion will track the zebra to kill her by using his knowledge of where the zebra is located in his environment. In particular, the lion has a *representation of the zebra in his mind*. This mental representation is itself provided by the zebra. This is because the zebra emits odours that the lion smells, makes sounds that the lions hears, and reflects light into the lion's eyes such that the lion can see her. These sensory inputs are transformed in the lion's brain to produce a mental representation of the zebra. This mental representation is then used by the lion to navigate his surroundings to target and kill the zebra. This is a second way in which the zebra is useful in bringing it about that the lion will kill her.

Crucially, these two points hold for predation more generally. After all, many other predators will also kill their prey by taking advantage of various parts of their body. For example, frogs latch onto an insect's body with their tongue to catch them in the air, and bears catch fish by pinning a fish's body to the bottom of a lake. Perhaps this is not true of all predators. A whale might simply open its mouth so that a fish will unknowingly swim inside. But virtually all predators use their senses to target and kill their prey. So the prey animal will still be useful in providing a representation of themselves in the predator's mind.

I have just argued that the zebra is useful in bringing it about that the lion kills her. We can now explain why this case is relevantly like Guinea Pig.

On the one hand, we know that the aid you withdraw from the five when you rescue the zebra is provided by predation (i.e., the lion killing the zebra). Since the zebra is useful in bringing it about that the lion kills her, this means that the zebra provides the aid that you withdraw from the five when you save her. And since the zebra would presumably want to be saved from the lion, you withdraw this aid on the zebra's behalf by saving her. Accordingly,

withdrawing this aid from the five is morally equivalent to merely *allowing* them to die. In this respect, saving the zebra is relevantly like saving Victim in Guinea Pig.

It is true that given our current ignorance of the long term ecological effects, we do not know for certain whether rescuing the zebra will be in her interest overall. Nevertheless, it is in the zebra's *expected* interest to be rescued. This is because she will almost certainly die painfully now if you do not intervene, and so her chances of suffering an even worse fate later on are inevitably lower than this.[7]

On the other hand, the very same considerations we have just invoked also show that failing to rescue the zebra involves harmfully *using her* as a means of saving the five. For we have just seen that the five will die if you rescue the zebra precisely because the zebra will not then be useful as a means of saving them. By contrast, the five will die as a merely foreseen side-effect of withdrawing aid from them if you rescue the zebra. In this respect, saving the zebra is also relevantly like saving Victim in Guinea Pig.[8]

---

[7] See also footnote 23.

[8] Consider two ways in which saving the zebra from the lion might appear morally different from saving Victim in Guinea Pig in this respect. Firstly, the way in which the zebra will be useful to saving the five other wild animals is *itself harmless* to her. This is because having parts of her body held onto by the lion and producing a mental representation of herself in the lion's mind is *itself* harmless to her. By contrast, the way in which Victim will be useful to saving the five in Guinea Pig is itself harmful to her. This is because the progression of Victim's disease is useful to developing a cure for the five and is also what will kill her. However, suppose that in Guinea Pig you only have enough time to either treat Victim's disease or develop a cure for the five by performing a battery of harmless medical experiments on Victim that are completely unrelated to her disease. The way in which Victim will be useful to saving the five in this version of the case is itself harmless to her. Nonetheless, you still seem morally required to save her. Secondly, whereas the five could not be saved if Victim were absent in Guinea Pig, perhaps the five other wild animals could be saved even if the zebra were absent. This is because the zebra herself might be what kills them (for example, by eating their food), and predation would have saved the five by eliminating her. However, suppose that Victim is immovably lying on a pill that could otherwise be used to save

In sum, whereas failing to rescue the zebra involves harmfully using her as a means of saving the five, rescuing the zebra involves merely foreseeing the deaths of the five. So the explanation of why you are morally required to rescue Victim in Guinea Pig also implies that you are morally required to rescue the zebra from the lion despite this having harmful consequences overall.

B. Carnivorous Prey

We have been considering a case in which you can rescue a *herbivorous* prey animal from a predator. But suppose you can instead rescue a *carnivorous* prey animal from a predator. Suppose you can rescue a nearby lion that is about to be killed by a hyena, though due to its long-term ecological effects this will also lead to the deaths of five other wild animals.

It might seem that explanation I have just offered for why you are morally required to rescue the zebra from the lion also says that you are morally required to rescue the lion from the hyena. But there is a potentially relevant difference between these two cases.

When you rescue the zebra from the lion, the continued survival of the zebra does not depend on any other wild animals being harmfully used. This is because the zebra is a herbivore and so will nourish herself only by eating plants. But if you rescue the lion from the hyena, the continued survival of the lion will depend on other wild animals being harmfully used. This is because the lion is a carnivore and so will need to kill other wild animals as a means of nourishing himself. Why might this difference matter morally?

I have argued that whereas failing to rescue the zebra involves harmfully using her as a means of saving the five, rescuing the zebra involves merely foreseeing the deaths of the five. Yet it looks like this asymmetry might not hold between the harms of rescuing the lion from

the five in Guinea Pig. The five could be saved if Victim were absent in this case since you could then access the pill. Nonetheless, you still seem morally required to save her. For these cases see Ramakrishnan (2016).

the hyena and the harms of letting him die: if you let the lion die, this will involve harmfully using him as a means of saving the five other wild animals. But if you save him, the five might also be harmfully used as a means of saving the lion. For they might be five animals that the lion will prey upon to stay alive.

So it seems that we cannot appeal to this consideration to distinguish morally between saving the lion from the hyena and letting him die. And since it seems plausible that when other things are equal you are morally required to maximise the good, it looks like you are morally prohibited from rescuing the lion from the hyena (and carnivorous prey animals from predators more generally) if this will make things worse overall.

Before addressing this concern, note that it does not touch my earlier argument that you are morally required to rescue herbivorous prey animals from predators. Even if this point shows that you are prohibited from rescuing carnivorous prey, you are still required to selectively rescue herbivorous prey.

Nonetheless, I think this worry is unfounded. Consider:

> *Guinea Pig**. Victim is suffering from *two* life-threatening diseases: disease *A* and disease *B*. If left untreated, disease A will kill Victim before disease B takes effect. You can easily treat disease A by administering a cure. But if you withhold treatment, a group of scientists can observe the fatal progression of disease A. They can thereby gain the medical knowledge necessary to save five other people from a similar condition. In this respect, the case is just like Guinea Pig. But consider the following addition: if you treat Victim's disease A and thereby allow the five to die, the group of

scientists will then study the progression of their condition. They can thereby gain the medical knowledge necessary to save Victim from disease B.[9]

It seems impermissible to let Victim die from disease A so that the five can be cured in this case. But unlike in Guinea Pig, the continued survival of Victim depends on the five being harmfully used as a means of saving her in this case. For the five will be used as a means of treating Victim's disease B only if you let them die by treating Victim's disease A.

Rescuing the lion from the hyena when he will subsequently prey upon five other wild animals is relevantly like rescuing Victim in Guinea Pig*. This is because disease A is relevantly like the hyena that will kill the lion, and disease B is relevantly like the threat of starvation that will subsequently kill the lion unless he eats his prey: if you do not rescue the lion from the hyena, this will involve harmfully uses him as a means of saving the five. But if you do rescue him from the hyena, the five will subsequently be harmfully used as a means of saving him from starvation. These two cases are therefore structurally equivalent. So any explanation that accommodates Guinea Pig* will also imply that you are morally required to rescue the lion from the hyena.

One might nonetheless wonder why you are morally to rescue Victim even in Guinea Pig*. Here is a potential explanation. It is plausible that when securing a certain good involves harmfully using a person, this good is *excluded* (or *significantly discounted*) from our moral consideration.[10] So since securing the good of the five being saved in Guinea Pig involves harmfully using Victim as a means of saving them, this good is excluded (or significantly discounted) as a consideration in favour of letting her die. Failing to treat Victim is therefore

---

[9] For a similar case, see *Six Heart Attacks* in Ramakrishnan (2016: 151). Six Heart Attacks differs from Guinea Pig* since the single person in Six Heart Attacks is not useful to saving the five. So my explanation of Guinea Pig* is consistent with but does not imply that you are permitted to save the one in Six Heart Attacks.

[10] See also Ramakrishnan (2016), Tadros (2018).

morally equivalent to letting her die when there is no (or almost no) good at stake. This explains why you are morally required to rescue Victim in Guinea Pig.

We can now explain Guinea Pig*. In this case, the good of the five being saved also involves harmfully using Victim as a means of saving them by letting Victim die from disease A. The good of the five being saved is therefore excluded (or significantly discounted) as a consideration in favour of letting her die. It is true that securing the good of Victim being saved from disease B involves the five being harmfully used the five as a means of saving her. But since the good of the five being saved has already been excluded (or significantly discounted) as a consideration in favour of letting Victim die, saving Victim is morally equivalent to the five being *harmlessly* used as a means of saving Victim from disease B. Yet there does not seem to be a moral constraint against *harmlessly* using a person as a means. Accordingly, the good of Victim being saved from disease B is *not* excluded (or significantly diminished) as a consideration in favour of treating her. This explains why you are morally required to rescue Victim in Guinea Pig*: just like in Guinea Pig, failing to treat Victim is morally equivalent to letting her die when there is no (or almost no) good at stake. Similar remarks explain why you are still morally required to rescue the lion from the hyena even when the lion will subsequently prey upon five other wild animals as a means of staying alive.

That completes my argument that rescuing the zebra from the lion (or the lion from the hyena) is morally equivalent to saving Victim in Guinea Pig (or Guinea Pig*).

We have been considering cases in which you can rescue a *single* prey animal from a predator when this will make things worse overall. But merely changing the numbers involved so that we can instead rescue many prey animals instead of just one makes no moral difference to the case. This is because what matters morally is the basic causal structure of the case. And this remains the same regardless of whether we are rescuing just one prey animal from a predator or one-thousand.

III. Long-Term Harm

I have just argued that interfering with predation is relevantly like saving Victim in Guinea Pig. But let us set this argument aside for the moment. Suppose that failing to rescue the zebra from the lion really does involve merely *foreseeably* letting her die, and that rescuing the zebra really does involve *doing harm* by leading to the deaths of five other wild animals. After all, this is how many naturally view the case.

In this section, I will argue that you are still morally required to rescue the zebra from the lion. This is because it is relevantly like rescuing an identified victim at the expense of a greater number of merely statistical victims.[11]

A. Identified and Statistical Victims

To start, consider

> *Death versus Paracetamol*. A villain has kidnapped Victim and ten billion other children. He will either (a) kill Victim or (b) give each of the ten billion children a paracetamol pill carrying (say) a one-in-a-billion chance of death. You must choose which.[12]

Intuitively, you are morally required to choose option (b) over (a) in this case, thereby sparing Victim while imposing a one-in-a-billion chance of death upon each the ten billion children. This is the case, even though it is statistically certain that choosing option (b) over (a) will result in ten deaths.

Whereas we know that Victim will be spared of death if we choose option (b), we cannot know in advance which particular ten children will be spared of death if we choose option (a). Choosing option (b) will therefore spare an *identified* victim, whereas choosing option (a) will

---

[11] See also Kagan (2019).

[12] See also Horton (2020).

spare ten merely *statistical* victims. Our duty towards identified victims appears significantly stronger than our duty towards statistical victims.

We have been considering a case in which you can rescue a nearby zebra who is about to be killed by a lion. Accordingly, you know which particular zebra will be saved by your intervention. But rescuing the zebra will lead to the deaths of five other wild animals only due to its *long-term* ecological effects. So we cannot know in advance which particular five wild animals will die (of the many potential five animals in the wild that could die) if you rescue the zebra. The zebra that you can save is therefore an identified victim, whereas the five other wild animals that you will harm if you rescue the zebra are merely statistical victims.

In my view, you are morally required to rescue the zebra from the lion even if this involves *killing* the five for the same kind of reason you are morally required to choose option (b) over (a) in Death versus Paracetamol.

I will first outline how many explain the moral distinction between identified and statistical victims in Death versus Paracetamol but argue that it is mistaken. I will then offer an more plausible explanation of this moral distinction, before drawing out its implications for saving the zebra.

B. The Ex Ante Distinction between Identified and Statistical Victims

Here is how many explain cases like Death versus Paracetamol.[13] We could accommodate Death versus Paracetamol by focusing on the *strength of the individual complaint* that could be made against your act. We could say that other things being equal, you are morally required to choose the option that *minimises the strongest individual complaint*, rather than the option that *maximises the sum of benefits minus harms*.[14]

---

[13] For example, Frick (2015).

[14] For example, Scanlon (1998).

From an *ex ante* perspective, a person's complaint against an act *A* is based on the *prospect* A gives them. If you choose option (a), the strongest individual complaint that will be generated is a complaint against a one-in-a-billion chance of death. This is because the individual risk of death you will impose upon each child if you choose option (b) is one-in-a-billion. By contrast, if you choose option (a), the strongest individual complaint that will be generated is an undiscounted complaint against death. This is because you will impose certainty of death upon Victim if you choose option (a).

Note that although choosing option (b) will generate *ten billion* complaints against a one-in-a-billion chance of death, we cannot aggregate these many complaints to outweigh Victim's complaint against death that will be generated if you choose option (a). Briefly stated, this is because it seems plausible that you can aggregate complaints only when the competing complaint in question are sufficiently close in strength. A complaint is *relevant* when it is sufficiently close in strength to the strongest competing complaint.[15]

Since a complaint against a one-in-a-billion chance of death is not relevant to a complaint against death, these many complaints cannot be aggregated to outweigh Victim's complaint against death. Accordingly, you are morally required to choose option (b) over (a) in Death versus Paracetamol. This is the ex ante explanation of the moral distinction between identified and statistical victims.

---

[15] For example, Voorhoeve (2014), Tadros (2019). To illustrate: while there is some number of people you should save from permanent paralysis over a single person from death, there is no number of people you should save from a temporary, minor headache over a single person from death. This is because a complaint against permanent paralysis is relevant to a complaint against death, and so we can aggregate many complaints against permanent paralysis to outweigh a single complaint against death. But a complaint against a temporary, minor headache is not relevant to a complaint against death, and so we cannot aggregate many complaints against a temporary, minor headache to outweigh a single complaint against death.

Similarly, if you rescue the zebra from the lion, the strongest individual complaint that will be generated is a complaint against a very small chance of death. This is because the individual risk of death you will impose upon each wild animal if you rescue the zebra is very small. After all, you cannot know in advance which particular five wild animals will die (of the many potential five that could die) if you rescue the zebra.

But if you refrain from rescuing the zebra, the strongest individual complaint that will be generated is a near undiscounted complaint against death. This is because the zebra will almost certainly be killed by the lion if you do not intervene. Since a complaint against a very small chance of death is not relevant to a complaint against death, we cannot aggregate these many complaints to outweigh the zebra's complaint against death.

Accordingly, rescuing the zebra from the lion when this kill five other wild animals is morally equivalent to rescuing her when this will impose on*ly a very small risk of death upon a single wild animal*. This is what the ex ante explanation says. And you would be morally required to rescue the zebra under these circumstances.

However, we should reject the ex ante explanation of the moral distinction between identified and statical victims. For although it can accommodate Death versus Paracetamol, it has very counterintuitive implications in

> *Death versus Poison*. A villain has kidnapped Victim and ten billion other children. He will either (a) kill Victim or (b) give each of the ten billion children a pill from a jar of ten-billion pills, ten of which are poisoned and so will kill whoever ingests it (the rest are completely harmless). You must choose which.[16]

Intuitively, you are morally required to choose option (a) over (b) in this case, thereby sparing ten children while imposing death upon Victim. But the ex ante explanation says that Death

---

[16] For example, Horton (2017; 2020).

versus Poison is morally equivalent to Death versus Paracetamol. This is because choosing option (b) will impose the same individual risk of death upon each of the ten billion child in both cases. Accordingly, the ex ante explanation has the very counterintuitive implication that you are morally required to choose option (b) over (a) in Death versus Poison. Accordingly, we should reject the ex ante explanation.

In the next sub-section, I shall develop an alternative explanation of the moral distinction between identified and statistical victims that avoids this problem. We can then return to the case of rescuing the zebra from the lion.

C. The Ex Post Distinction between Identified and Statistical Victims

From an *ex post* perspective, a person's complaint against an act *A* is based on the *actual* results of A. Furthermore, if there is (say, a ninety-percent chance) of A causing some outcome *O*, the ex post view says that we should posit a full, undiscounted complaint against O and then discount this complaint by ten-percent to reflect the chance of O occuring.[17]

I will now argue that we can distinguish morally between Death versus Paracetamol and Death versus Poison on the ex post view.

It might seem that the ex post view says that you are morally required to choose option (a) over (b) in both cases. This is because in both cases it looks like choosing option (a) will generate a single ex post complaint against death, whereas choosing option (b) will generate ten ex post complaints against death. And ten complaints against death can be aggregated to outweigh a single complaint against death.

But there is a crucial difference between Death versus Paracetamol and Death versus Poison that we can appeal to on the ex post view. Whereas the risks involved in choosing option (b) in

---

[17] For example, Otsuka (2015). See also Reibetanz (1998).

Death versus Paracetamol are probabilistically *independent*, the risks involved in choosing option (b) in Death versus Poison are probabilistically *dependent*.

This is because in Death versus Poison, the chance that *someone* will be killed if a particular child is given a pill depends on *how many other* pills are given from the jar of pills. As the number of pills given from the jar increases, so increases the likelihood that giving a pill will result in *someone* having been given a poisoned pill. When the final pill from the jar is given out, this ensures that *all ten poisoned pills* have been distributed.

In Death versus Paracetamol, by contrast, the chance that someone will be killed if a particular child is given paracetamol is independent of how many other paracetamols are given. Each paracetamol has its own probabilistically independent natural lottery.

I suggest that we evaluate each gamble involved in giving out a pill in each case *individually*. It would be odd to hold that while you are permitted to sequentially hand out a pill one by one, you are nonetheless prohibited from cutting to the chase by giving them out all at once like in the cases under consideration.

Yet this is precisely what is implied on the ex post view if we evaluate the set of gambles *as a whole* rather than individually in Death versus Paracetamol. This is because when the set of gambles is evaluated as a whole, choosing option (b) in Death versus Paracetamol generates ten ex post claims against death. And since ten complaints against death can be aggregated to outweigh the single complaint against death that will be generated if you choose option (a), the ex post view says that you should choose option (a) in this case. But now consider what the ex post view says when each gamble is evaluated individually (akin to a case of sequential choices).

When each gamble is considered individually in Death versus Poison, we know that one of these gambles will generate ten ex post complaints against death. For one of these gambles will ensure that *all ten poisoned pills* have been distributed. But when each gamble is considered

individually in Death versus Paracetamol, we know that each gamble involved in choosing option (b) will generate only an ex post complaint against death *discounted by one-in-a-billion*. The lottery is reset after each gamble.

Since ten complaints can be aggregated to outweigh a single complaint against death, you are morally required to choose option (a) over (b) in Death versus Poison. By contrast, since a complaint against death discounted by one-in-a-billion is not relevant to a complaint against death, you cannot aggregate such complaints to outweigh a complaint against death. Accordingly, you are morally required to choose option (b) over (a) in Death versus Paracetamol. These are the intuitively correct results in both cases.

So unlike the ex ante explanation, the ex post explanation of Death versus Paracetamol can also accommodate Death versus Poison. We should therefore accept this ex post explanation of the moral distinction between identified and statistical victims.

D. Long-Term Harm

With the ex post explanation of the moral distinction between identified and statistical victims in hand, we can now return to our original aim. What does this view say about rescuing the zebra from the lion?

If you refrain from rescuing the zebra, this will generate a near undiscounted ex post complaint against death. This is because the zebra will almost certainly be killed by the lion if you do not intervene.

On the other hand, we are supposing that rescuing the zebra from the lion will lead to the deaths of five other wild animals only due to its *long term ecological effects*. So rescuing the zebra from the lion will kill the five only by bringing about a *very long sequence of minor disruptions* within the ecosystem.

The probability of any given ecological disruption itself killing a wild animal is *independent* of how many other disruptions within the sequence kills a wild animal. Each disruption has its

own probabilistically independent natural lottery with respect to killing a wild animal. And since each disruption to the ecosystem within this sequence is very minor, the probability that any given disruption will itself kill a wild animal is *very low*. It is just that because there will be so many of these ecological disruptions if you rescue the zebra, *overall* this is very likely to result in the deaths of five other wild animals in the long run. In this respect, rescuing the zebra is relevantly similar to choosing option (b) in Death versus Paracetamol (as opposed to in Death versus Poison).

This means that when each gamble is considered individually, each gamble produced by rescuing the zebra will generate only an ex post complaint against death discounted for a *very small probability of occuring*. Since a complaint against death discounted for a very small probability of occuring is not relevant to a complaint against death, we cannot aggregate such complaints to outweigh a complaint against death.

Accordingly, even if rescuing the zebra from the lion involves *killing* five other wild animals, this is morally equivalent to rescuing her when this involves imposing only a *very small risk of death upon a single wild animal*. And you would be morally required to rescue the zebra under these circumstances.[18]

---

[18] While I lack space to explore this issue in detail, I believe that similar remarks provide a solution to William MacAskill and Andreas Mogensen's (2021) 'Paralysis Argument'. Briefly, their argument says that the constraint against doing harm implies that we must do as little as possible (e.g., sit motionlessly for as long as possible). This is because when we consider the long-term consequences of your usual everyday actions (e.g., going to the supermarket), you have no greater reason to expect that they will involve benefiting as opposed to doing harm to others. By contrast, if you sit motionlessly, you will at worst merely allow harm to occur. MacAskill and Mogensen then conclude that the only other permissible alternative on non-consequentialism is to spend all of your time improving the long-term future (e.g., by trying to reduce the risk of human extinction). But your usual everyday actions will harm others in the long run only by bringing about a very long sequence of very low risk, probabilistically independent events. So each event in the sequence generates only an ex post complaint against

E. Sequence of Rescues

Let us take stock. Up to this point, I have provided an ex post explanation of the moral distinction between identified and statistical lives. I argued that we should accept this explanation of Death versus Paracetamol because unlike the ex ante view, it can also accommodate Death versus Poison. I then argued that on this ex post view, you are morally required to rescue the zebra from the lion even if this involves killing five other wild animals due to its long-term ecological consequences. Let us now consider an objection.

We have been considering cases in which you can rescue a *single* prey animal from a predator when this will make things worse overall. It might seem that merely changing the numbers involved so that we can instead rescue many prey animals instead of just one makes no moral difference to the case. This is because it might seem that what matters morally is simply the *relative strength* of the complaints at stake. And this remains the same regardless of whether we are rescuing just one prey animal from a predator or one-thousand.

Yet one could argue that when interfering with predation on a *wide scale*, we should evaluate the entire sequence of individual rescues as a *single unit*. This is because each individual rescue within the sequence is unified under the common goal of rescuing prey animals from predators.

Yet even if I am correct that each individual rescue carries (what is morally equivalent to) only very small risk of killing a wild animal on the ex post view, the sequence *itself* is very likely to kill some wild animals by repeatedly imposing this risk. So if we evaluate this sequence as a whole, implementing it will generate some complaints against death from the ex post perspective.

---

death discounted for a very small probability of occuring. Accordingly, going to the supermarket when this is likely to kill many others in the long run is morally equivalent to going to the supermarket when this carries only a very small risk of killing someone. And you are permitted to perform such actions.

This might seem unproblematic for my argument. This is because the *relative* strength of the individual complaints at stake remains the same regardless of whether we evaluate the sequence as a whole or instead evaluate each rescue within the sequence individually. And since we have already seen that you are required to carry out a rescue when it is considered individually, it might seem that you are similarly required to carry out the sequence of individual rescues even when it is evaluated as a whole.

But the problem is that it seems plausible that the strength of the constraint against doing harm increases *disproportionately* with the strength of the individual complaint at stake. For example, it seems disproportionately harder to justify killing a person than to break their finger.

And so even though the *relative* strength of the individual complaints at stake remains the same either way, the individual complaints generated by interfering with predation are *stronger* when we evaluate the sequence of rescues as a whole than when each rescue is evaluated individually. This is because the sequence itself will generate ex post complaints against death, whereas I have argued that each rescue considered individually will generate only ex post complaints against death discounted for a very small probability of occuring. Accordingly, carrying out the sequence is *disproportionately harder to justify* when it is evaluated as a whole.

I have argued that morally speaking each rescue carries only a very small risk of killing a wild animal, and so morally speaking the sequence itself will still save many *more* prey animals than the number of wild animals it will kill in expectation. But you seem prohibited from doing harm even when this will be beneficial overall.

To illustrate, consider:

> *Gas*. Five people are suffering from a life-threatening disease. You can easily treat their disease by releasing a gas into the room that will cure them. However, the gas will also leak into the next room and poison Victim, thereby killing her.[19]

You seem prohibited from saving the five in this case. Given the moral asymmetry between doing and allowing harm, saving the five seems insufficiently important to justify killing Victim. Likewise, one could argue that even if morally speaking interfering with predation on a wide scale will save many more wild animals than it will kill in expectation, it is similarly prohibited if the harm imposed is *great enough*.

The upshot is that if the objection under consideration is correct that we must evaluate the sequence of rescues as a whole, then we cannot plausibly say that interfering with predation on a wide scale is morally permitted (let alone required) on the ex post view.[20]

Let us grant (for the sake of the argument) that we must evaluate the sequence involved in a wide scale rescue as a whole. Nonetheless, we can still avoid this objection. Compare Gas with

> *Gas**. Five people are suffering from a life-threatening disease. You can easily treat their disease by releasing a gas into the room that will cure them. Although the gas itself is completely harmless to Victim, it will nonetheless move some poisonous germs that are already present in the air towards her, thereby infecting and killing Victim.[21]

---

[19] Foot (1967)

[20] Note that this objection does not arise on the ex ante view. This is because whether we evaluate the entire sequence of rescues as a single unit or instead evaluate each rescue individually, the individual risk imposed on each wild animal remains the same.

[21] Kamm (2015).

Unlike in Gas, you are intuitively permitted to save the five in this case. Here is why. Whereas in Gas you kill Victim by *creating a threat*, in Gas* you kill her with an *independently existing threat*. This is because in Gas you create the gas that kills Victim by initially releasing it into the atmosphere. By contrast, you do not create the poisonous germs that kill Victim in Gas*. Although you kill Victim by *moving* these poisonous germs towards her, the germs that kill her are themselves independently present in environment.[22]

Perhaps being morally responsible for the existence of the threat that will kill your victim ties you more closely to the harm that you impose upon her than when you are not responsible for the threat itself but instead merely for the fact that it will kill someone. This could explain why killing by creating a threat is harder to justify than killing with an independently existing threat.

Return now to the objection under consideration. The *way in which* interfering with predation will kill other wild animals is relevantly like Gas* as opposed to Gas. To see why, consider that we prevent predation by affecting either the predators or the prey animals. So it is only through our initial effects on either of them that preventing predation will eventually kill other wild animals. The chain of harmful long-term ecological effects is first mediated through either the predators or the prey animals.

---

[22] See Kamm (2015) who also draws this distinction. However, Kamm does not consider that this distinction can also accommodate other cases in which you seem permitted to save five when this involves killing one. To illustrate: you seem permitted to turn a lethal trolley from five onto one, and also permitted to move five people away from a lethal trolley heading towards them when they will land on and kill a single person. Like in Gas* and unlike Gas, these cases also involve killing with an *independently existing threat* (i.e., you kill by merely moving the trolley, or by merely moving the five). So while I cannot fully defend this claim here, it looks like we can accommodate of all of these cases by appealing to a single distinction.

So unlike in Gas and akin to Gas*, interfering with predation involves killing with an *independently existing threat* rather than by creating a new threat. This is because you kill the other wild animals *through your effects on either the predators or the prey animals*. And just like the poisonous germs in Gas*, the predators or prey animals that will subsequently cause the deaths of other wild animals are themselves *independently present in the environment*.

Accordingly, even if the objection under consideration is correct that interfering with predation on a wide scale is disproportionately harder to justify than an individual rescue on the ex post view I am defending here, you are still required (or at least permitted) to intervene.[23] This is because the case is relevantly like Gas* as opposed to Gas.

F. Pre-Emptive Rescue

That completes my argument that rescuing interfering with predation is morally equivalent to rescuing an identified victim at the expense of a greater number of merely statistical victims in cases like Death versus Paracetamol. So even if my argument from the previous section is mistaken and interfering with predation really does involve *killing* (as opposed to merely allowing) many more wild animals than it will save, you are still morally required to intervene.

My argument in this section appealed to an ex post explanation of the moral distinction between identified and statistical victims. But suppose we accept the ex ante explanation anyway. We have already seen that the ex ante explanation also says that you are morally required to rescue the zebra from the lion even at the expense of the five other wild animals.

---

[23] Some hold that you are *required* to save the five in cases like Gas*. For example, Frowe (2018). Others hold that intervening in such cases involves a *supererogatory* cost. For example, Kamm (2015). But recall we are setting aside the personal costs involved in interfering with predation (see footnote 3). Still others hold that you are *merely permitted* to intervene. For example, Gordon-Solmon and Pummer (2022). I leave it open which of these views we should accept and so whether the objection under consideration succeeds at showing that you are merely permitted (as opposed to morally required) to intervene.

To close off, let me address a potential difficulty with showing that you are required to interfere with predation in certain cases on the ex ante explanation of the moral distinction between identified and statistical victims.

We have been considering a case in which the zebra that you can rescue is an identified victim. But one could object that interfering with predation on a wide scale might involve rescuing zebras who are merely statistical victims. This is because it would be inefficient to rescue one zebra at a time by repeatedly intervening only when they are about to be killed by a lion.

We could instead rescue many zebras at a time. For example, we could kill lions when they are asleep, sterilise them to prevent them even more lions from coming into existence, or in the future genetically engineer lions from carnivores into herbivores. By doing so, we could thereby pre-emptively rescue many zebras from being killed at a time.

Yet unlike in the case we have been considering so far, we cannot know in advance which particular zebra will be saved (of the very many potential zebras that could be saved) by such interventions. For example, we cannot know in advance which particular zebra will be killed by the offspring of the lion we could now sterilise. Accordingly, it looks like we cannot invoke the distinction between identified and statistical victims in this scenario on the ex ante explanation.

Note that this problem does not arise on the ex post explanation. This is because even if we cannot know in advance which particular zebra will be saved by our intervention, we know that *some* zebra will be saved. Crucially, the pre-emptive intervention will not save zebra through a sequence of very low risk, probabilistically independent gambles. It will instead save some zebra by eliminating the threat *here and now*. Accordingly, failing to pre-emptively intervene will still generate an undiscounted ex post complaint against death.

But we can still avoid the problem even on the ex ante view.

It is true that if you pre-emptively interfere with predation, this will only *slightly* reduce the individual risk of death each zebra faces *given the evidence that is currently available to you*. This is because you cannot know in advance which particular zebra will be saved by your intervention.

Nevertheless, you do know that your intervention will *significantly* reduce the individual risk of death that a zebra faces *given the evidence that is available to this zebra* (or a rationale person occupying her position). This is because when a zebra is about to be killed a lion, the lion will be actively chasing her and so the zebra (or a rationale person occupying her position) will know that she is very likely going to be killed. Yet it is plausible that on the ex ante view we should evaluate the zebra's complaint against your failing to rescue her based on the prospect this gives her *from her perspective* (or from the perspective of a rationale person occupying her position).

To see why, consider

> *Death versus Pill*. A villain has kidnapped Victim and ten billion other children. He will either (a) kill Victim or (b) give each of the ten billion children a pill that will kill them if they have a certain genotype. You must choose which. Crucially, you know that only ten of the children have the deadly genotype but do not know who. You nonetheless know that *each child knows this fact about themselves*.

Intuitively, you are morally required to choose option (a) over (b) in this case, thereby sparing ten children while imposing death upon Victim. It is true that *given to the evidence currently available to you*, choosing option (b) will impose only a one-in-a-billion chance of death upon each of the ten billion children. In this respect, Death versus Pill is just like Death versus Paracetamol. So it might seem that the ex ante view must say that you are morally required to choose option (b) over (a) in Death versus Pill, just like in Death versus Paracetamol.

But unlike in Death versus Paracetamol, you know that *relative to the evidence available to the ten billion children*, choosing option (b) in Death versus Pill will generate ten ex ante complaints against death. This is because you know that ten of these children are aware of the fact that given their particular genotype they will die if they are given a pill.

The ex ante view can therefore accommodate Death versus Pill by evaluating each person's complaint against your act based on the prospect this gives them *from their perspective*. On this view, since ten complaints against death can be aggregated to outweigh a single complaint against death, the ex ante view can accommodate the judgement that you are morally required to choose option (a) over (b) in Death versus Pill.

Similarly, even if you cannot know in advance which particular zebra will be saved by your pre-emptive intervention, you know that whoever this zebra is, *she will know this fact about herself*. For at the time your intervention would otherwise save her, the lion will be chasing her and so she will know she is almost certain to die. Failing to intervene will therefore generate a near undiscounted complaint against death. In this respect, the statistical zebra you can save is still *morally* like an identified victim.

By contrast, you know at the time your intervention would otherwise save the zebra (i.e., when she is being hunted by a lion), the five other wild animals that will die as a result of this *do not yet know this fact about themselves*. So intervening only slightly increases the individual risk of death for each wild animal. So the strongest individual complaint that intervening will generate is a complaint against a very small chance of death. In this respect, the five are still relevantly like statistical victims. Since a complaint against a very small chance of death is not relevant to a complaint against death, we cannot aggregate these many complaints to outweigh the zebra's complaint against death.

Accordingly, we can still explain why you are morally required to *pre-emptively* interfere with predation and thus rescue merely statistical as opposed to identified prey animals, *even if*

we accept the ex ante explanation of the moral distinction between identified and statistical victims.[24]

Overall, then, even if interfering with predation involves *killing* many more wild animals than it will save, you are still morally required to intervene. This holds on either the ex ante or the ex post explanation of the moral distinction between identified and statistical victims.

IV.   Conclusion

In sum, many argue that you are prohibited from interfering with predation on a wide scale. This is because they argue that given our current ignorance of the long term ecological effects, preventing predation on a wide scale will probably make things worse overall.

But even if it will certainly prove disastrous, I have argued that you are nonetheless morally required to intervene. This is because interfering with predation is morally akin to certain cases in which you seem required to rescue despite this making things worse overall.

In Section II, I argued that rescuing prey animals from predators is relevantly like rescuing a person who will otherwise be used as a means of saving a greater number (like Victim in Guinea Pig). In Section III, I argued that rescuing prey animals from predators is relevantly like rescuing an identified victim at the expense of a greater number of merely statistical victims (like Victim in Death versus Paracetamol).

So we must not wait until our understanding of the ecosystem advances. We are morally required to interfere with predation immediately.

---

[24] My remarks in this sub-section also addresses the objection some might have that pre-emptively interfering with predation is *against* the overall expected interest of prey animals and so undermines the very reason for intervening in the first place. For example, McMahan (2016). This objection would also count against pre-emptively interfering with predation on my argument from the previous section if we accept the ex ante view.

References

Donaldson, Sue & Kymlicka, Will (2011). *Zoopolis: A Political Theory of Animal Rights*. New York: Oxford University Press. Edited by Will Kymlicka.

Frowe, Helen (2014). *Defensive Killing*. Oxford, UK: Oxford University Press.

Frowe, Helen (2018). Lesser-Evil Justifications for Harming: Why We're Required to Turn the Trolley. *Philosophical Quarterly* 68 (272):460-480.

Gordon-Solmon, Kerah & Pummer, Theron (2022). Lesser-Evil Justifications: A Reply to Frowe. *Law and Philosophy* 41:639–646.

Horton, Joe (2017). Aggregation, Complaints, and Risk. *Philosophy and Public Affairs* 45 (1):54-81.

Horton, Joe (2020). Aggregation, Risk, and Reductio. *Ethics* 130 (4):514-529.

Horton, Joe (2020). Aggregation, Risk, and Reductio. *Ethics* 130 (4):514-529.

Johannsen, Kyle (2021). *Wild Animal Ethics: The Moral and Political Problem of Wild Animal Suffering*. New York, NY, USA: Routledge.

Kagan, Shelly (2019). *How to Count Animals, More or Less*. Oxford: Oxford University Press.

Kamm, Frances Myrna (ed.) (2015). *The Trolley Problem Mysteries*. New York: Oup USA.

McMahan, Jeff (1993). Killing, Letting Die, and Withdrawing Aid. *Ethics* 103 (2): 250-279.

McMahan, Jeff (2016). "The Moral Problem of Predation" in Chignell, Andrew; Cuneo, Terence & Halteman, Matthew C. (eds.). Philosophy Comes to Dinner: Arguments on the Ethics of Eating. Routledge.

Mogensen, Andreas & MacAskill, William (2021). The Paralysis Argument. *Philosophers' Imprint* 21 (15).

Otsuka, M., 2015, "Risking Life and Limb: How to Discount Harms by their Improbability," *Identified versus Statistical Lives: An Interdisciplinary Perspective*, *Identified*

versus Statistical Lives: An Interdisciplinary Perspective, Cohen, G., N. Daniels, and N. Eyal (eds.), New York: Oxford University Press: pp. 77–93.

Quinn, Warren S. (1989). Actions, Intentions, and Consequences: The Doctrine of Double Effect. *Philosophy and Public Affairs* 18 (4): 334-351.

Quong, Jonathan (2020). *The Morality of Defensive Force*. Oxford University Press.

Ramakrishnan, Ketan H. (2016). Treating People as Tools. *Philosophy and Public Affairs* 44 (2): 133-165.

Ramakrishnan, Ketan H. (2016). Treating People as Tools. *Philosophy and Public Affairs* 44 (2): 133-165.

Reibetanz, Sophia (1998). Contractualism and aggregation. *Ethics* 108 (2):296-311.

Scanlon, Thomas (1998). *What we owe to each other*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Tadros, Victor (2019). Localized Restricted Aggregation, in David Sobel, Peter Vallentyne, and Steven Wall (eds), *Oxford Studies in Political Philosophy Volume 5*, Oxford Studies in Political Philosophy (Oxford).

Tadros, Victor, 'Dimensions of Intentions: Ways of Killing in War', in Seth Lazar, and Helen Frowe (eds), *The Oxford Handbook of Ethics of War*, Oxford Handbooks.

Voorhoeve, Alex (2014). How Should We Aggregate Competing Claims. *Ethics* 125 (1):64-87.

Woollard, Fiona (2015). *Doing and Allowing Harm*. Oxford, GB: Oxford University Press.

Chapter IV: The Moral Status of Undoing One's Past Behaviour: A Reply to Hanna

I.   Introduction

Consider the following moral principle:

> *The Doctrine of Doing and Allowing* (DDA). At least when other things are equal, doing harm to others is harder to justify than merely allowing harm to befall them.[1]

DDA is a central aspect of common-sense morality. For example, DDA can capture the intuitive judgement that while saving five lives justifies merely letting one person die, it does not justify driving over and thereby killing someone trapped on the road enroute to rescuing the five.[2] Despite its widespread intuitive appeal, Jason Hanna has recently argued that DDA appears to have implausible implications in cases that involve *undoing one's past behaviour*.[3]

To avoid these problematic implications, Hanna contends that proponents of DDA (henceforth 'deontologists') must resolve three challenges. To a first approximation, Hanna's *First Challenge* is to explain why *withdrawing aid that one provided in the past* can be permissible when doing harm would be prohibited; his *Second Challenge* is to explain why *allowing oneself to have done harm* can be prohibited when merely allowing harm would be permissible; his *Third Challenge* is to explain why one is prohibited from *allowing oneself to*

---

[1] Some clarificatory points about DDA. Firstly, that doing harm is *harder to justify than* merely allowing harm means that the duty to aid can be overridden by costs to the agent or to third-parties that would be insufficient to override the duty not to harm, other things being equal. Secondly, these duties are *pro tanto*. This means that the duty not to do harm can be overridden when the stakes are sufficiently high. I omit the "other things being equal" clause hereafter. I also assume throughout that death would be a harm to potential victims, and that DDA's stringency is such that it typically prohibits killing one instead of letting five die.

[2] Foot (1984).

[3] Hanna (2014, 2015). See also Persson (2013), Unruh (2021).

have done harm to fewer rather than a greater number of people in certain cases. Hanna argues that deontologists have not satisfactorily met these challenges.

In this article, I defend DDA against all three challenges. In Section II, I argue that Hanna's First Challenge fails because implementing a plan that involves providing and then later withdrawing aid is morally equivalent to merely allowing harm. In Section III, I argue that Hanna's Second Challenge fails because <initiating a harmful sequence and then failing to terminate it> is itself morally equivalent to doing harm. In Section IV, I argue that Hanna's Third Challenge fails on similar grounds as his Second Challenge.

II.   Hanna's First Challenge

In this section, I address Hanna's First Challenge to DDA.

A. Rockslide and Double Rockslide

To start, consider:

> *Rockslide*: While Agent and his friend are hiking in the mountains, Agent's friend is bitten by a rattlesnake. There is a dose of antidote back at their campsite. Agent can save his friend's life only by loading the friend into Agent's nearby truck and then driving promptly to the campsite to retrieve the antidote. Unfortunately, in driving the truck, Agent would trigger a rockslide. This rockslide would remove a boulder presently interposed between Victim and an unoccupied car headed toward Victim. In short, if Agent acts to save his friend's life, he will remove an obstacle to a lethal threat confronting Victim.[4]

---

[4] Hanna (2015): 79.

If Agent saves his friend in Rockslide, he will remove a barrier (i.e., the boulder that is protecting Victim) to a pre-existing threat that will kill Victim (i.e., the unoccupied car headed toward her). Removing a barrier to a pre-existing threat is sometimes called *enabling* harm.[5]

Intuitively, it is morally impermissible for Agent to enable Victim's death in this case. Why? I will assume throughout that this is because enabling Victim's death in Rockslide is morally equivalent to *killing* her.[6] After all, the barrier that Agent withdraws from Victim would remain *even in his absence*. We thus cannot plausibly say that Agent prevents himself from saving Victim and so merely allows her to die by removing this barrier, as we could if the barrier required further contributions from Agent.

By contrast, if Agent does not save his friend, he merely lets them die. Since Agent faces a choice between killing Victim and letting his friend die, and since DDA says that doing harm is harder to justify than merely allowing harm, DDA prohibits Agent from saving his friend in this case.

Next, consider:

> *Double Rockslide*. As before, Agent's friend is bitten by a rattlesnake. Agent can reach the campsite in time to give his friend the antidote only by driving his truck down two steep inclines. When Agent starts down the first incline, he will trigger a rockslide that interposes a large boulder between Victim and an unoccupied car that is presently heading unobstructed toward Victim. When Agent starts down the second incline, he

---

[5] Foot (1967).

[6] Although Hanna notes that this assumption is not necessary for his challenge to work (he only needs Agent's behaviour to seem *impermissible* in Rockslide), it will simplify our discussion. Hanna (2015): 79. For accounts of enabling harm that have this implication, see Hanna (2015), Woollard (2015): 62-79.

will trigger a second rockslide that removes the boulder interposed by the first rockslide, with the result that Victim is re-exposed to the original threat from the car.[7] Intuitively, it is morally permissible for Agent to save his friend in this case. This is so, despite the fact that by descending the second incline, Agent will enable Victim's death by removing a barrier that requires no further contributions from him. After all, once Agent has descended the first incline the boulder will remain in place even if Agent were to suddenly disappear.

Yet we have just seen in Rockslide that it seems impermissible for Agent to save his friend when this involves enabling Victim's death by removing a barrier that requires no further contributions from him. So, why is Agent permitted to save his friend in Double Rockslide (which involves withdrawing aid that he earlier provided) but not in Rockslide (which involves withdrawing aid that he did not earlier provide)? Hanna's First Challenge is to explain why this is the case.[8]

Suppose that deontologists are unable to resolve Hanna's First Challenge. They must then reject either the judgement that (a) Agent may save his friend in Double Rockslide, or the judgement that (b) Agent is prohibited from saving his friend in Rockslide. Either way, DDA will have implausible implications.[9] In what follows, I respond to Hanna's First Challenge.

B. Reply to Hanna's First Challenge

Having outlined Hanna's First Challenge, I will now begin my critique. To start, consider:

---

[7] Hanna (2015): 80.

[8] Hanna (2015): 82.

[9] Alternatively, deontologists could accept what Shelly Kagan (1989: 14) calls a 'dangling distinction' between Agent's behaviour in Rockslide and Double Rockslide. This is a distinction that accommodates our intuitions but lacks a deeper explanation of its moral significance. Yet while the moral distinction between these two kinds of behaviours may be a brute fact about morality, it is preferable to take on fewer brute moral facts, rather than more.

Sled. The case is similar to Double Rockslide. However, Agent has an additional option: at the top of the first incline, he can also jump onto a sled so that it slides downhill uninterrupted. Agent cannot stop the sled once he has jumped onto it, though it comes to a natural halt at the bottom of the second incline. At this point, Agent can administer the antidote to his friend. Importantly, the sled will initiate the same rockslides as the truck would have done during its descents. Agent jumps onto the sled and saves his friend.

By jumping onto the sled in this case, Agent performs a *single* act that has two relevant effects: interposing the boulder (when the sled descends the first incline), and then later displacing the boulder (when the sled descends the second incline). Intuitively, Agent's behaviour in Sled is morally permissible. It is not difficult for deontologists to explain why this is the case. This is because jumping onto the sled is morally equivalent to merely *letting* Victim die.[10] After all, this sequence of effects leaves everything as it would have been with respect to Victim and the boulder had Agent not descended any inclines at all, yet this would uncontroversially count as merely letting Victim die. And Agent is permitted to save his friend over Victim (who is a stranger).[11]

We have just seen why Agent may save his friend when this involves descending both inclines via a *single* act. Now go back to Double Rockslide. The only difference between Agent's options in Sled and in Double Rockslide is that he cannot descend both inclines by performing a single act in Double Rockslide. Instead, he can descend both inclines only by performing *multiple* acts that together produce the same sequence of effects as in Sled: Agent

---

[10] Many prominent accounts of DDA have this implication. See Kagan (1989: 83-111), Foot (1984), Quinn (1989a). Bennett (1995: 85-120), Kamm (2006: 17-21), Woollard (2015: 21-61).

[11] Deontologists typically endorse agent-relative prerogatives that permit Agent to accord disproportionate (when considered impartially) weight to his friend's life.

can descend the first incline (thereby interposing the boulder), followed by descending the second incline (thereby displacing the boulder).

Nonetheless, Agent has a *plan* to save his friend in Double Rockslide. This plan involves descending both inclines. Rather than the individual acts that constitute the sequence involved in this plan, it is Agent's plan *itself* that appears to be the relevant unit of evaluation in Double Rockslide.[12] By evaluating Agent's plan as a single unit of deliberation rather than the individual acts that constitute the sequence involved in this plan, his plan is treated as though it were a single act.

On this view, since descending both inclines via a *single act* in Sled is morally equivalent to merely *letting* Victim die, then so is descending both inclines via multiple acts as part of a *single plan* in Double Rockslide. And since Agent is permitted to save his friend over Victim, this explains why Agent may implement his plan in Double Rockslide (and *a fortiori*, why he may perform the individual acts that constitute the sequence involved in this plan, including descending the second incline).[13]

---

[12] What constitutes a 'plan'? I only have space to note that my use of the term tracks its everyday usage. Thus, it seems compatible with many formulations of plans. See Bratman (1987). Although my view distinguishes morally between an agent's plan and the individual acts involved in implementing this plan, this is consistent with rejecting views such as the Doctrine of Double Effect that distinguish morally between planning (or intending) harm and merely foreseeing it. See Quinn (1989b), Tadros (2015). The evaluative role of a plan could instead be replaced by a 'complex act' without affecting my overall argument. For example, just as 'building a house' is a complex act that involves layering bricks and hammering nails, 'saving Agent's friend' in Double Rockslide is a complex act that involves descending each incline and administering the antidote.

[13] Jeff McMahan (2006: 642-645) argues against evaluating plans as a single unit. For example, he thinks that even if one could permissibly attempt to deter an enemy attack via a *single act* by irreversibly programming a nuclear weapon that will automatically nuke the enemy city if they attack our own, one could not permissibly bring about this same sequence of effects via *multiple acts as part of a single deterrence plan* (that is, threaten the

Let us take stock. I have just offered an explanation of why Agent may save his friend in Double Rockslide. Recall, Hanna's First Challenge is to explain why this case is morally different from Rockslide. We have been assuming that DDA prohibits Agent from saving his friend in Rockslide. Does my account of Double Rockslide say otherwise? It does not.

To see why, consider that implementing Agent's plan in Double Rockslide has the effect of *interposing and then later displacing* the barrier (i.e., the boulder) that would otherwise save Victim. When we evaluate Agent's plan, it is *this* aspect of the plan that makes its implementation morally equivalent to merely letting Victim die. After all, it is this particular sequence of effects that leaves everything as it would have been with respect to Victim and the boulder had Agent not descended any inclines at all, yet this would uncontroversially count as merely letting Victim die.[14]

---

same form of retaliation, and then manually nuke the enemy city if they defy this threat). Yet, McMahan argues, if we evaluate the nuclear deterrence plan as a single unit, then we must conclude that it is morally equivalent to the single deterrence act, and so could also be permissible. In response to McMahan's objection, even if we evaluate the nuclear deterrence plan as a single unit, implementing this plan could not be permissible. This is because the harm that results from nuking the enemy city as part of this plan is an *unnecessary* effect of deterrence. For one could implement an alternative plan that would be an equally effective deterrent yet does not involve nuking anyone: insincerely threaten the same form of retaliation, and then *refrain* from fulfilling this threat if it is defied. By contrast, since the single deterrence act is presumably one's only means of deterrence, any harms that result from performing this act are a *necessary* consequence of deterrence. Thus, McMahan's objection against evaluating plans as a single unit is unsuccessful (I owe this point to [REDACTED]). Accordingly, it does not undermine my account of Double Rockslide. While McMahan's objection shows that we cannot infer the permissibility of a plan from the permissibility of a single act *just because* they produce the same sequence of effects, my account of Double Rockslide rests instead on the claim that *other things being equal*, a single act and a plan have the same DDA status.

[14] For various explanations of why this is so, see the DDA accounts cited in footnote 10. What should we say about the DDA status of the *individual acts* involved in implementing Agent's plan in Double Rockslide? On the

By contrast, Agent did not interpose the boulder that is currently protecting Victim in Rockslide. So, the question of whether implementing his plan to save his friend has the effect of interposing and then later displacing the boulder does not even arise in Rockslide. This means that we cannot argue that Agent merely lets Victim die in Rockslide by evaluating his plan, as we argued in response to Double Rockslide.[15] Thus, my account is consistent with DDA prohibiting Agent from saving his friend in Rockslide. Having outlined my reply to Hanna's First Challenge, let us now consider an objection.

---

one hand, we could say that no individual act within this plan has any DDA status. On the other hand, we could say that the individual acts within this plan have a DDA status, but that this status is irrelevant to the individual act's permissibility. For example, we could say that enabling Victim's death by descending the second incline is *conceptually* equivalent to *killing* her, but that this killing is permissible because it is part of a plan that is itself permissible. Either way, Agent's plan in Double Rockslide is permissible, and so are the individual acts involved in implementing it.

[15] What if Agent must unexpectedly modify his plan to save his friend in Double Rockslide when the boulder is *already protecting* Victim (i.e., after descending the first incline but before descending the second)? If Agent still counts as descending both inclines as part of the *same plan*, then my view will still imply that he merely lets Victim die by implementing this plan, and so is still permitted to proceed. For example, suppose that one's individual acts are part of a single plan when they are each performed for the same *end-goal*. On this account, Agent's first-incline descent and his subsequent, unanticipated acts nonetheless occur within the same plan if they are each performed for Agent's end-goal of *saving his friend*. By contrast, if we cannot plausibly claim that Agent descends both inclines as part of a single plan when he adjusts it midstream, then my view will not morally distinguish this case from Rockslide. However, this implication is unproblematic for my view. For under these circumstances, it no longer seems permissible for Agent to save his friend (for further elaboration, see 'Pipe Sealer I' below).

C. Defending the Reply

Although he does not endorse it, Hanna considers whether the following moral principle resolves his First Challenge:

> *Undoing Principle for Past Aid* (UP-Past Aid): All else being equal, an agent's choice to undo past behavior by which he provided aid is easier to justify than an agent's choice to do harm in other, more standard ways (e.g., by initiating a threat or interfering with aid provided by others).[16]

What does UP-Past Aid say about Double Rockslide? By descending the second incline, Agent displaces the boulder that is currently protecting Victim. Agent thereby undoes his past behaviour of descending the first incline by which he provided this aid to Victim. Accordingly, UP-Past Aid says that enabling Victim's death in Double Rockslide is easier to justify than in Rockslide, and so Agent may save his friend in Double Rockslide but not in Rockslide. One may therefore object that I have not shown why we should accept my reply to Hanna's First Challenge over UP-Past Aid.

In response to this objection, there are two respects in which my reply to Hanna's First Challenge is superior to UP-Past Aid. Firstly, Hanna thinks that it is unclear how deontologists could explain *why* it is the case that an agent's choice to undo past behaviour by which he provided aid is easier to justify than standard instances of doing harm. In Hanna's view, UP-Past Aid lacks a clear rationale.[17]

By contrast, there is a clear rationale for my reply to Hanna's First Challenge. Agent's plan to save his friend in Double Rockslide is morally significant because although Agent performs distinct bodily movements by performing the individual acts that constitute the sequence

---

[16] Hanna (2015: 85).

[17] Hanna (2015: 88-89).

involved in his plan, his *agency* is unified across the entire plan. So to adequately assess Agent's conduct *qua* agent, we must evaluate his overall plan as a single unit. The implementation of Agent's plan is morally equivalent to merely letting Victim die. And deontologists accept that the distinction between doing and merely allowing harm is morally significant.[18]

Secondly, in addition to Hanna's concern about its lack of rationale, UP-Past Aid has implausible implications that he does not mention. Consider:

> *Pipe Sealer I*. An earthquake cracks a pipe at a factory, releasing poisonous chemicals into the water supply. Before a dangerous amount is released, Agent seals the pipe. But a year later, he realises that the pipe contains an antidote that is required to cure his friend's fatal disease. Agent therefore returns and removes the seal to obtain the antidote. As a result, Agent saves his friend but Victim dies from drinking contaminated water.[19]

Just like in Double Rockslide, Agent undoes his past aid when he removes the seal in this case. Nonetheless, doing so seems morally impermissible. This is at odds with UP-Past Aid. For if UP-Past Aid permits Agent to save his friend when this involves enabling Victim's death in Double Rockslide, then it must also permit this in Pipe Sealer I.[20] We should therefore find a

---

[18] Appealing to DDA here is not question-begging. My point is simply that the rationale underlying my reply to Hanna's First Challenge is acceptable to deontologists on grounds that they already accept.

[19] This case is based on McMahan's (1993: 256) 'Pipe Sealer' case. I assume that just like the boulder in Rockslide and Double Rockslide, the seal is unowned.

[20] If UP-Past Aid is to resolve Hanna's First Challenge, it should not matter morally that the seal is already preventing the gas from being released when Agent removes it. For consider that in Double Rockslide, whether the boulder is already preventing the car from hitting Victim when Agent eventually descends the second incline does not seem to affect whether he may descend it.

different moral principle that not only distinguishes morally between Rockslide and Double Rockslide, but also has plausible results in Pipe Sealer I.

Unlike UP-Past Aid, my view has plausible implications in Pipe Sealer I. For although he earlier sealed the cracked pipe in this case, Agent was ignorant about the presence of the antidote when doing so. Thus, Agent's plan to save his friend in Pipe Sealer I does not involve *adding and then later removing* the barrier (i.e., the seal) that would otherwise save Victim. In this respect, the case is akin to Rockslide. This means that just like in Rockslide, we cannot argue that Agent merely lets Victim die in Pipe Sealer I by evaluating his plan, as we argued response to Double Rockslide. Thus, my view is consistent with claiming that saving Agent's friend in Pipe Sealer I involves killing Victim, and so is prohibited by DDA.[21]

To bring this point into focus, consider:

> *Pipe Sealer II.* The case is similar to Pipe Sealer I. However, Agent knows that the chemicals within the cracked pipe must interact with the surface of the seal for a year to synthesise the antidote required to cure his already sick friend. Agent therefore seals the pipe in order to produce the antidote with the intent of later removing it. A year later, he returns and remove the seal to access the finished product. As a result, Agent saves his friend but Victim dies from drinking contaminated water.

Unlike in Pipe Sealer I, Agent seems permitted to save his friend in this case. My view explains this judgement. This is because Agent's plan to save his friend in Pipe Sealer II involves adding and then later removing the barrier (i.e., the seal) that would otherwise save Victim. And as we saw before in Double Rockslide, when we evaluate Agent's plan in Pipe Sealer II, this aspect of the plan makes its implementation morally akin to merely letting Victim die. Since Agent may save his friend over Victim, this explains why Agent may implement his plan in Pipe

---

[21] For accounts of enabling harm that have this implication, see Hanna (2015: 250-279), Woollard (2015: 62-79).

Sealer II (and *a fortiori*, why he may perform the individual acts involved in implementing it, including removing the seal).[22]

To summarise. We should reject UP-Past Aid in favour of my reply to Hanna's First Challenge for two reasons. Firstly, unlike UP-Past Aid, my reply has a clear rationale. Secondly, unlike UP-Past Aid, my reply has plausible implications in Pipe Sealer I.

Before concluding, let me offer a final reason to accept my account of Double Rockslide. In my view, Double Rockslide is simply an instance of a more general set of cases in which an act that would be impermissible when considered individually is permissible because it is part of a permissible plan. The fact that my general explanation of Double Rockslide can accommodate structurally similar cases further raises its plausibility.

To illustrate, consider

> Painting vs Person. You can easily save either a famous paining or a person.[23]

Even if you are an art lover, you are morally required to save the person over the painting.

But now consider

> Painting vs Person (Burning Building). A famous painting and a person are in different rooms of a burning building. The personal cost of entering the building is sufficiently high that you are permitted to refrain from entering and saving either. But once you are inside you can easily save either the famous painting or the person.[24]

---

[22] Woollard (2015: 73) argues that removing an unowned barrier is akin to merely allowing harm if the agent removing the barrier is *already using it*. However, Woollard's view does not imply that removing the seal in Pipe Sealer II is morally akin to merely letting Victim die when this act is evaluated individually. For once the antidote has been produced, Agent is no longer *using* the seal. At this point, the seal is merely an obstacle preventing Agent from accessing the antidote contained within the pipe. Thus, we cannot appeal to this aspect of Woollard's view to accommodate our judgement that Agent is permitted to remove the seal in this case.

[23] For example, Kagan (1989), McMahan (2018), Kamm (2023).

[24] For example, Kagan (1989), McMahan (2018), Kamm (2023).

Suppose you enter the burning building *in order to save the painting*. Intuitively, you are permitted enter the burning building and save the painting over the person in this case. But it might seem puzzling that whereas you are permitted to save the painting over the person in Painting vs Person (Burning Building), you are prohibited from doing in Painting vs Person. For once you are inside the burning building you can easily save either the famous paining or the person, just like in Painting vs Person. This is just like how it seems puzzling that you seem permitted to descend the second incline and thereby enable Victim's death in Double Rockslide, even though this individual act seems morally indistinguishable from impermissibly enabling Victim's death in Rockslide.

We can explain Painting vs Person (Burning Building) on similar grounds to Double Rockslide. For since you enter the burning building *to save the painting*, you have a *plan* to rescue the painting. This plan involves both entering the burning building and saving the painting. Rather than the individual acts that constitute the sequence involved in this plan, it is your plan *itself* that appears to be the relevant unit of evaluation. By evaluating your plan as a single unit of deliberation rather than the individual acts that constitute the sequence involved in this plan, your plan is treated as though it were a single act.

On this view, morally speaking there is no stage at which you can easily save either the painting or the person in Painting vs Person (Burning Building). Your options are instead relevantly like the following set of options: (a) press button A at a very high personal cost that will save the famous painting (b) press button B at a very high personal cost that will save the person or (c) do nothing.[25]

You are permitted to refrain from pressing button B. This is because you are permitted to forgo saving a person at a very high cost to yourself. You are also permitted to refrain from

---

[25] I set aside the option of entering the burning building and saving neither the person nor the painting.

pressing the button A for similar reasons. Nonetheless, you are also permitted to press button A because you are permitted to incur a very high cost to save the painting.[26]

Entering the burning building *in order to* save the painting is relevantly like pressing button A, and entering the burning building *in order to* save the person is relevantly like pressing button B. Since you are permitted to refrain from pressing button B, you are permitted to refrain from entering the burning building to save the person. And since you are permitted to press button A, you are permitted to implement the plan to save the painting (and *a fortiori*, permitted to perform the individual acts that constitute the sequence involved in this plan, including saving the painting over the person once inside the burning building).[27]

---

[26] Some might find my assessment of this button case counterintuitive. They might argue that if pressing a button were costless to you, you would be morally required to press button B over button A, thereby saving the person over the painting. So can how can it be permissible to press button A rather than button B when there is cost involved in each option? My response is that the mere fact that there is *no difference in cost* between pressing either button does not explain why you must press button B over button A when there is no cost involved. It is instead that there is no cost involved in pressing button B and so you cannot appeal to this cost to justify refraining from pressing button B. And since the painting is insufficiently valuable, you also cannot appeal the painting to justify refraining from pressing button B. By contrast, when pressing button B and thereby saving the person is very costly, you can appeal to this cost to justify refraining from pressing it, even if it is no more costly than pressing button A and thereby saving the painting.

[27] Suppose that five people are in one room of a burning building and a single person is in another room of the same building. For example, Pummer (2016), McMahan (2018). The cost of entering the building is sufficiently high that you are permitted to refrain from entering and saving either. But once you are inside you can easily save either the five or the one. One might argue that you seem prohibited from saving the one in this case, but that my explanation of Painting vs Person (Burning Building) seems to imply that you are permitted to do so. My response is that there is a morally significant difference between these two cases. In the case in which you can save either the five or the one, you have no special concern for the one and so would presumably save the five if this was your only option of saving anyone from the burning building. This means that there is no difference in *cost to you*

Accordingly, I have provided a *unified* explanation of structurally alike cases such as Double Rockslide and Painting vs Person (Burning Building). This further raises the plausibility of my account of Double Rockslide.

Altogether, then, deontologists can resolve Hanna's First Challenge by evaluating Agent's plan to save his friend in Double Rockslide. Let us now turn to Hanna's Second Challenge to DDA.

III.   Hanna's Second Challenge

I will outline Hanna's Second Challenge, before arguing that it is unsuccessful.

A. Poisoner I and Poisoner II

To start, consider:

> *Poisoner I*. Earlier this morning, Agent deposited a dose of lethal poison into a teapot from which one person drinks tea at the same time each afternoon. As this time

---

between *saving the one* and *refraining from saving the one*: if you do not save the one, you will save the five and so will incur the high cost of entering the burning building anyway. But there is a difference in the *good* that will be produced between saving the one and refraining from saving the one: if you do not save the one, you will save the five. You are thus required to refrain from saving the one because refraining from saving the one involves saving five lives instead of just one at no additional cost to yourself. Crucially, you are still permitted to do nothing since you are permitted to refrain from saving the five at a high cost. It is just that if in an act of supererogation you would save the five if it were your only option of saving anyone, you are prohibited from saving the one when you can save either. By contrast, as an art lover you have a special concern for the painting that you do not have for the person. Accordingly, you presumably would *not* save the single person at a high cost if this were your only option of saving anything even though you would incur this cost for the painting. This means that there is no difference in the good that will be produced between *saving the painting* and *refraining from saving the painting*: if you do not save the painting, you will save *no one*. You are thus not required to refrain from saving the painting because refraining from saving the painting does not involve producing a greater good at no additional cost to yourself.

approaches, Agent sets off to warn his potential victim. On the way, Agent encounters five other people drowning in a shallow pond. There is insufficient time to save all six people: Agent can rescue the five would-be drowning victims or warn the one would-be poisoning victim, but he cannot do both.[28]

By failing to warn the one would-be poisoning victim, Agent will now allow her to die. However, Agent will not *merely* allow her die. For if he does not warn the one would-be poisoning victim, Agent will have killed her in virtue of having previously poisoned her teapot. So, we can say that Agent will now *allow himself to have killed* the one would-be poisoning victim if he does not warn her.

Other things being equal, it seems plausible that you are permitted to save the greater number. So, if allowing oneself to have done harm is morally equivalent to merely allowing harm, Agent would also be permitted to save the five would-be drowning victims over the one would-be poisoning victim. Yet Agent seems morally required to warn the one would-be poisoning victim instead.

Thus, while saving five lives seems to justify merely letting one die, it does not seem to justify allowing oneself to have killed someone. This suggests that allowing oneself to have done harm is *harder* to justify than merely allowing harm.

Hanna considers whether deontologists could accommodate Poisoner I by appealing to:

> *THE PAST MATTERS* (TPM). If an agent will have killed (or, more generally, done harm to) a victim provided that he does not now assist that victim, then his refraining from providing such assistance when he could do so has the same moral status as a killing (or act of harm-doing).[29]

---

[28] Hanna (2014: 678-679).

[29] Hanna (2014: 680).

If TPM is correct, then allowing oneself to have killed the one would-be poisoning victim would be morally equivalent to *killing* her. And since DDA says that doing harm is harder to justify than merely allowing harm, TPM would explain why failing to warn the one would-be poisoning victim is (a) likewise harder to justify than merely letting her die and (b) impermissible, given that DDA typically prohibits killing one instead of merely letting five die. So, if deontologists can appeal to TPM, then they can accommodate the judgement that Agent is morally required to save the one would-be poisoning victim in Poisoner I.

Nonetheless, Hanna contends that TPM has implausible implications. Consider:

> *Poisoner II*. Earlier this morning, Agent deposited some lethal poison into a teapot from which *five* people share tea at the same time each afternoon. As this time approaches, Agent sets off to warn her potential victims. Unfortunately, Agent can reach the five only by driving down a narrow mountain path, which is blocked by the body of an innocent, unconscious person. If Agent takes the time to move this person, she will be too late to save the five.[30]

In Poisoner II, Agent will now allow herself to have killed the five would-be poisoning victims if she does not warn them. TPM therefore says that failing to warn the five would-poisoning victims is morally equivalent to killing them.

Why is this problematic for deontologists? Hanna claims that most deontologists agree that "an agent who confronts a choice between killing one and killing five is permitted (or perhaps even required) to minimize harm by killing the one".[31] So, if TPM is correct, deontologists should similarly agree that Agent would be at least permitted to warn the five would-be poisoning victims in Poisoner II, even though this involves killing the unconscious person. Intuitively, however, Agent may not do this.

---

[30] Hanna (2014: 681).

[31] Hanna (2014: 681).

Hanna argues that Poisoner II therefore shows that deontologists must accept that allowing oneself to have done harm is *easier* to justify than doing harm, and so accept that these two kinds of acts are morally *in*equivalent. This is at odds with TPM. Accordingly, Hanna argues that deontologists must reject TPM, and so cannot use it to accommodate Poisoner I.[32]

Hanna's Second Challenge is to accommodate Poisoner I without generating counterintuitive implications in Poisoner II.[33] Suppose that deontologists are unable to resolve Hanna's Second Challenge. They must then reject either the judgement that (a) Agent is prohibited from saving the five in Poisoner I, or the judgement that (b) Agent is prohibited from saving the five in Poisoner II. Either way, DDA will have implausible implications.[34] In what follows, I respond to Hanna's Second Challenge.

B. Reply to Hanna's Second Challenge

Having outlined Hanna's Second Challenge, I will now begin my critique.

In Poisoner I, Agent poisoning the teapot does not have an *independent* DDA-status. This is because whether Agent poisoning the teapot kills the one would-be poisoning victim depends on whether agent subsequently warns her. This means that <poisoning the teapot and not warning the would-be poisoning victim> *as a whole* is morally equivalent to killing the one would-be poisoning victim, whereas <poisoning the teapot and warning the would-be poisoning victim> is morally equivalent to refraining from killing her. Since DDA prohibits saving five when this involves killing one, it prohibits saving the five would-be drowning victims when this involves <poisoning the teapot and not warning the would-be poisoning

---

[32] Hanna (2014: 682).

[33] Hanna (2014: 680).

[34] Again, deontologists could accept a 'dangling distinction' between, on the one hand, allowing oneself to have done harm, and on the other hand, doing harm and merely allowing harm. But as noted earlier, it is preferable to take on fewer brute moral facts, rather than more.

victim>. So *given that* Agent *has poisoned the teapot*, he is morally required to warn the one would-be poisoning victim over saving the five would-be drowning victims in Poisoner I.

What about Poisoner II? Recall Hanna's claim that most deontologists agree that "an agent who confronts a choice between killing one and killing five is permitted (or perhaps even required) to minimize harm by killing the one".[35] Is this correct? In defense of this claim, Hanna appears to have the following case in mind:

> *Trolley.* Driver's fatal trolley is headed toward five people. If Driver does nothing, the trolley will kill the five. However, Driver can redirect the trolley away from the five and onto a sidetrack where it will kill only one person.[36]

Having started the trolley, Driver will kill the five if he does nothing in this case. Driver must therefore choose between killing five and killing one in Trolley.[37] In this respect, my view says that Driver's options in Trolley are akin to those faced by Agent in Poisoner II. Intuitively, Driver is permitted to kill the one in Trolley. Yet contrary to Hanna's contention, this case does not show that Agent would also be permitted to kill the one in Poisoner II on my explanation of Poisoner I. This is because even if my view is correct, there is still a morally relevant difference between Agent's options in Poisoner II and Driver's options in Trolley. To see why, consider:

---

[35] Hanna (2014: 681).

[36] Thomson (2008: 360). Additionally, consider:

> *Crossroads.* A trolley is at a crossroads. Driver must send it toward either one or five people, thereby killing them (the trolley will otherwise explode and kill one-zillion people).

Intuitively, Driver may kill the one (Kamm (2006: 141-142)). Although Hanna never explicitly mentions these cases, I infer that this is what he had in mind from his references to Thomson and Kamm in Hanna (2014: 681, footnote 13).

[37] Arguably, Driver will only *allow himself to have killed* the five in Trolley. Nonetheless, Crossroads clearly involves either killing one or *killing* five. My remarks below regarding Trolley apply equally to Crossroads.

Two Trolleys. As in Trolley, except that Driver can avoid killing the five only by pressing a button on his trolley's control panel. This will start up a second trolley that will fatally run over one person enroute to derailing the trolley headed toward the five.[38]

In this case, Driver must also choose between killing five and killing one. In this respect, Driver's options in Two Trolleys are morally akin to those he faced in Trolley. So, if Driver is permitted to kill the one in Trolley *because* he faces a choice between killing five and killing one, Driver would also be permitted to kill the one in Two Trolleys. Intuitively, however, he may not do this. This shows that Driver's permission to kill the one in Trolley is not fully explained by the fact that his alternative involves killing five. I will now argue that killing the one in Poisoner II is akin to doing so in Two Trolleys, and so is prohibited for the same reason.

Let us start by considering why Driver is permitted to kill the one in Trolley but not in Two Trolleys. There is widespread disagreement among deontologists regarding the details of this explanation. Yet for the purposes of showing that killing the one in Poisoner II is relevantly akin to doing so in Two Trolleys, we need not decide among competing accounts. It is enough to note that virtually any way of morally distinguishing Trolley from Two Trolleys will also prohibit Agent from killing the one in Poisoner II. And this is enough to show that my view does not imply that Agent is permitted to kill the one in Poisoner II.

Some deontologists morally distinguish Trolley from Two Trolleys by appealing to the fact that killing the one in the former case involves *redistributing* a pre-existing threat (i.e., Driver's trolley), whereas in the latter case it involves *creating* a new threat (i.e., the second trolley).[39] Others appeal to the fact that in Two Trolleys, killing the one involves *using the space* that the

---

[38] Kamm (2015: 61).

[39] Thomson (1985). Since Driver must send the trolley *somewhere* in Crossroads, killing the one in this case still involves *distributing* (as opposed to creating) a threat.

one is in (i.e., with the second trolley), whereas in Trolley it does not.[40] Still others appeal to differences between the causal pathways through which the one is killed and the five are saved in each case. Whereas in Two Trolleys killing the one is causally *upstream* relative to the survival of the five, it is causally *downstream* in Trolley.[41] These are among the most prominent deontological explanations of why Driver may kill the one in Trolley but not in Two Trolleys.

Return now to Poisoner II. Killing the one in this case is akin to doing so in Two Trolleys in all of the above respects. For killing the one in Poisoner II also involves creating a new threat (i.e., Agent's car), using the space that the one is in (i.e., with Agent's car), and is causally upstream relative to the survival of the five. Accordingly, the most prominent explanations of why Driver is prohibited from killing the one in Two Trolleys also prohibit Agent from killing the one in Poisoner II. By appealing to these distinctions, deontologists who believe that Driver may kill the one in Trolley can explain why Agent is nonetheless prohibited from killing the one in Poisoner II, even if <poisoning the teapot and failing to warn the five would-be poisoning victims> is morally equivalent to killing the five would-be poisoning victims.[42] Thus, Hanna's Second Challenge fails.[43] Let us now finally turn to Hanna's Third Challenge to DDA.

IV.   Hanna's Third Challenge

---

[40] Quong (2009). For reasons I cannot discuss here, it matters that the one also has a moral claim to this space on Quong's view.

[41] Kamm (2006).

[42] Unruh (2021: 379) concedes Hanna's objection against TPM (and so against extensionally similar explanations such as my own). In this respect, her response to Hanna's Second Challenge differs from mine.

[43] While I lack space to show this, deontologists can similarly appeal to the moral distinctions examined above between Crossroads/Trolley and Two Trolleys to undermine Hanna's objection against the view that *allowing oneself to do harm in the future* is morally equivalent to doing harm. See Hanna (2014: 686-689).

I will outline Hanna's Third Challenge, before arguing that it is unsuccessful.

A. Button

Hanna present his Third Challenge in stages. He first asks us to consider

> *Button I*. Last week, Agent initiated a threat that is now about to kill five people. The five can be saved if, and only if, they now receive a certain drug. By pressing a button (Button I), Agent is able to initiate a chemical reaction that will produce this drug. Unfortunately, this reaction would release fumes into an adjacent room occupied by an innocent person who cannot be moved. Agent thus has a choice. If he refrains from pressing Button I, the five will die as a result of the threat he initiated last week. If he presses Button I, the five will be saved, but the one will die from the dangerous fumes.[44]

Button I is relevantly like Poisoner II. Intuitively, Agent is morally prohibited from pressing Button I. But now consider

> *Button II*. As before, but Agent was not sure whether there was anyone in the adjacent room, so he sent an assistant to check. While the assistant was gone, Agent grew impatient, so he negligently pressed Button I, thereby initiating the reaction. The assistant now returns and tells Agent that the adjacent room is occupied by one person who cannot be moved and who will certainly be killed by the dangerous fumes if nothing is done. Fortunately for this person, Agent now discovers a second button (Button II). By pressing Button II, Agent can terminate the chemical reaction before it produces the drug or lethal fumes. (Perhaps pressing Button II will send a reaction-inhibiting substance to the reaction site.) Agent thus has a choice. If he refrains from pressing Button II, the reaction will run to completion, the one will die from the fumes, and the five will be saved by the drug. If he presses Button II, the reaction will be

---

[44] Hanna (2014: 689-690).

terminated, the one will survive, and the five will die from the threat Agent created last week.[45]

Intuitively, Agent is morally required to press Button II in this case. Hanna's Third Challenge is to explain why this is the case. Hanna argues that we cannot appeal to a moral asymmetry between doing harm and allowing oneself to have done harm to accommodate Button II, as one might have done to accommodate Poisoner II. This is because both of your options involve allowing oneself to have done harm: if you press Button II, you thereby allow yourself to have killed the five, and if you refrain from pressing Button II, you thereby allowing yourself to have killed the one. So it seems difficult to accommodate Button II regardless of our position on the moral status of allowing oneself to have done harm.

If deontologists are unable to resolve Hanna's Third Challenge, then they must reject the intuitive judgement Agent is morally required to press Button II. Accordingly, DDA will have an implausible implication. In what follows, I will defend a reply to Hanna's Third Challenge.

   B. Reply to Hanna's Third Challenge

Having outlined Hanna's Third Challenge, I will now begin my critique.

We can explain Button II on similar grounds to Poisoner II. For example, refraining from pressing Button II involves saving the five while killing the one by *creating a threat* (i.e., the poisonous fumes). Similarly, refraining from pressing Button II involves killing the one with a means (i.e., the chemical reaction) that *causes* the five to be safe (i.e., by *causing* the life-saving drug to be produced). Yet we have already seen in the previous section that you are prohibited from killing one rather than killing five in this manner (for example, in Two Trolleys).

However, Hanna argues that deontologists cannot appeal to further distinctions beyond DDA to morally distinguish between your two options in Button II. He asks us to suppose that

---

[45] Hanna (2014: 690).

when you initiated the threat to the five last week, *you did so in the course of saving ten others*. This means that if you press Button II the five will be killed in a structurally similar manner to how the one will be killed if you refrain from pressing Button II.[46]

For example, just as refraining from pressing Button II involves saving the five while killing the one by creating a threat, pressing Button II involves saving the ten while killing the five by creating a threat. Similarly, just as refraining from pressing Button II involves killing the one with a means that causes the five to be safe, pressing Button II also involves killing the five with a means that causes the ten to be safe. So it looks like we cannot appeal to such factors to distinguish morally between your two options in Button II.

Hanna's objection assumes that causing harm by, for example, creating a new threat or with a means that causes a greater good, makes imposing this harm *more objectionable* than it would otherwise be. So making it such that you initiated the threat to the five last week in the course of saving ten others is supposed to render killing the five more objectionable than it would otherwise be, thereby counterbalancing the increased objectionableness of killing the one in the course of saving the five. It looks like we should then simply choose the option that minimises the number of killings.

However, we could instead say that securing a good when this involves causing harm by creating a new threat or with a means that causes the good *excludes* this good from our moral consideration.[47] On this view, the fact that saving the ten involves killing the five by creating a threat means that the good of saving the ten is *excluded* as a consideration in favour of killing the five. Similarly, the fact that saving the five involves killing the one by creating a threat means that the good of saving the five is *excluded* as a consideration in favour of killing the one. It is true that if you press Button II you will impermissibly kill the five. But this fact is

---

[46] Hanna (2014: 693-694).

[47] See also Ramakrishnan (2016), Tadros (2018).

itself excluded as a moral consideration in favour of killing the one. Accordingly, you are still prohibited from killing the one in this variant of the case and so are still morally required to press Button II.

Hanna also asks us to suppose that when you initiated the threat to the five last week, you did so in the course of saving *the very same individual* who will now be killed by the poisonous fumes if you do not press Button II. Accordingly, the causal structure that would link the death of the one to the survival of the five mirrors the causal structure that would link the death of the five to the survival of the one.[48]

For example, just as refraining from pressing Button II involves saving the five by creating a threat to the one, pressing Button II involves saving the one by creating a threat to the five. Similarly, just as refraining from pressing Button II involves killing the one with a means that causes the five to be safe, pressing Button II also involves killing the five with a means that causes the one to be safe. So again it looks like we cannot appeal to such factors to distinguish morally between your two options in Button II.

However, we can explain this version of the case on similar grounds as the previous variant. The fact that the one will be killed *by the threat that she was originally under* is excluded as a consideration in favour of killing the five. Similarly, the fact that the five will be killed by the threat you initiated last week is excluded as a consideration in favour of killing the one *with the poisonous fumes*. Since they are *different* threats, the fact that the one will be killed by the *poisonous fumes* is *not* excluded as a consideration against killing him, even though the fact that the one will be killed *by the threat that she was originally under* is excluded as a consideration in favour of killing the five. It is true that if you press Button II you will impermissibly kill the five. But this fact is *itself* excluded as a moral consideration in favour of

---

[48] Hanna (2014: 694).

killing the one with the poisonous fumes. Accordingly, you are still prohibited from killing the one in this version of the case and so are still morally required to press Button II.

Overall, then, Hanna's Third Challenge fails on similar grounds as his Second Challenge. Just as we can distinguish morally between your options in Poisoner II by appealing to factors beyond DDA, we can also distinguish morally between your options in Button II and its variants by appealing to similar factors beyond DDA.

V.   Conclusion

In sum, I have defended DDA against Hanna's three challenges. I began, in Section II, by arguing that his First Challenge fails. This is because we should evaluate Agent's plan to save his friend in Double Rockslide. On this view, although saving Agent's friend involves killing Victim in Rockslide, it involves merely letting her die in Double Rockslide. Thus, DDA prohibits Agent from saving his friend in Rockslide but not in Double Rockslide. I then argued, in Section III, that Hanna's Second Challenge fails. This is because in Poisoner I <poisoning the teapot and not warning the would-be poisoning victim> is itself morally equivalent to killing the single victim. But we can still accommodate Poisoner II by appealing to factors beyond DDA. Finally, in Section IV, I argued that Hanna's Third Challenge fails on similar grounds as his Second Challenge. This is because we can also distinguish morally between your options in Button II and its variants by appealing to factors beyond DDA.

References

Bennett, Jonathan (1995). *The Act Itself* (New York: Oxford University Press)

Bratman, Michael (1987). *Intentions, Plans, and Practical Reason* (Cambridge: Harvard University Press).

Foot, Philippa (1967). The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review* 5:5-15.

Foot, Philippa (1984). Killing and Letting Die, in *Abortion: Moral and Legal Perspectives*, ed. Jay L. Garfield and Patricia Hennessey (Amherst: University of Massachusetts Press): 177-183.

Hanna, Jason (2014). Doing, Allowing and the Moral Relevance of the Past. *Journal of Moral Philosophy* 12 (2014): 677-698

Hanna, Jason (2015). Enabling Harm, Doing Harm, and Undoing One's Own Behavior. *Ethics* 126 (2015): 68-90.

Kamm, Frances (2006). *Intricate Ethics: Rights, Responsibilities and Permissible Harm* (New York: Oxford University Press): 17-21

Kamm, Frances (2015). How Was the Trolley Turned? in Eric Rakowski, ed. *The Trolley Problem Mysteries* (Oxford University Press).

Kamm, Frances (2023). Supererogation and Duty. In David Heyd (ed.), *Handbook of Supererogation*. Springer Nature Singapore. pp. 29-49.

McMahan, Jeff (1993). Killing, Letting Die, and Withdrawing Aid. *Ethics* 103: 250-279.

McMahan, Jeff (2006). Paradoxes of Abortion and Prenatal Injury. *Ethics* 116: 625-645.

Persson, Ingmar (2013). *From Morality to the End of Reason* (Oxford: Oxford University Press)

Quinn, Warren (1989a). Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing. *Philosophical Review*, 98 (3): 287-312

Quinn, Warren (1989b). Actions, Intentions, and Consequences: The Doctrine of Double Effect. *Philosophy and Public Affairs*, 18 (4): 334-351

Quong, Jonathan (2009). Killing in Self-Defense. *Ethics* 119: 507-537

Ramakrishnan, Ketan H. (2016). Treating People as Tools. *Philosophy and Public Affairs* 44 (2):133-165.

Tadros, Victor, 'Dimensions of Intentions: Ways of Killing in War', in Seth Lazar, and Helen Frowe (eds), *The Oxford Handbook of Ethics of War*, Oxford Handbooks.

Tadros. Victor (2015). Wrongful Intentions without Closeness. *Philosophy and Public Affairs*, 43 (1) (2015): 52-74

Thomson, Judith Jarvis (1986). The Trolley Problem. In William Parent (ed.), *Rights, Restitution, and Risk: Essays in Moral Theory*. Cambridge, Mass.: Harvard University Press. Edited by William Parent.

Thomson, Judith Jarvis (2008). Turning the Trolley. *Philosophy and Public Affairs* 36 (4): 359-374.

Unruh, Charlotte Franziska (2021). Letting Climate Change. *Journal of the American Philosophical Quarterly* (2021): 368-386.

Woollard, Fiona (2015). *Doing and Allowing Harm*. Oxford, GB: Oxford University Press.

McMahan, Jeff (2018). Doing Good and Doing the Best, in Paul Woodruff (ed.), *The Ethics of Giving: Philosophers' Perspectives on Philanthropy* (New York).

Pummer, Theron (2016). Whether and Where to Give. *Philosophy and Public Affairs* 44 (1):77-95.