

Why More Than What Happens Matters:
Robust Rights and Harmless Wronging

Joseph Bowen



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the University of St Andrews

February 2020

Candidate's declaration

I, Joseph Bowen, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 79,510 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 was admitted as a research student at the University of St Andrews in September 2016.

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

Date: 11th May 2020

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.

Date: 11th May 2020

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, Joseph Bowen, 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: 11th May 2020

Signature of candidate:

Date: 11th May 2020

Signature of supervisor:

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Joseph Bowen, 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: 11th May 2020

Signature of candidate:

Acknowledgments

First off, I would like to thank all those I have discussed my work with at conferences, between seminars, at the pub, and anywhere else I could check one's intuitions. A thanks to those who have taken the time to give me comments on various papers that have found their way into the thesis in some form: Matt Clark, Luke Davis, Antony Duff, Claire Field, Helen Frowe, Katherine Hawley, Jacob Hienze, Mark LeBar, Sandra Marshall, Bodhi Melnitzer, Daniel Muñoz, Matthew Oliver, Tom Parr, Justin Snedegar, Kevin Sharp, Julian Savulescu, Ravi Thakral, and Lucas Sierra Velez. Thanks also to my annual reviewers Adam Etinson, Kent Hurtig, and Ben Sachs. And a special thanks to Thomas Douglas, who supervised my BPhil thesis where the Safety Condition first saw the light of day, as well as Cécile Fabre, whose question during a supervision first prompted my appealing to what could otherwise be the case. A thanks, also, to Marvin Backes, Karamvir Chadha, James Kirkpatrick, Quentin Parr, and Lewis Ross.

Thanks to my team of proof-readers: Lisa Bastian, Clotilde Torregrossa, Miguel Eglar, Ethan Landes, and Janis Schaab. Thanks to my office mates, both in the Hut and Arché: Colin McLean, Stefano Lo Re, Thomas McKenna, and Saranga Sudarshan (not to mention those close enough that they could hear the noise carrying). And a thanks to the Arché Whisky Club (as well as The Keys) for getting me through my final year.

A huge thankyou to the best supervisory team one could ask for, Rowan Cruft and Theron Pummer. I have felt so supported during the PhD and have never stewed over a question I thought too stupid ask.

Thanks to my other office mates, Mouse and Oscar, for a combination of taking me on walks when I was stuck, interrupting skypes, listening to me practice talks, and leaving muddy footprints on my work. Thanks to my parents for everything; without them, I would not have had the qualities that made writing this thesis over the years actually quite enjoyable.

And finally, the biggest thanks of all to Stephanie for, as well as moving 500 miles north, always listening to me moan about philosophy (which really can't be fun), supporting where I wanted to take things, and lifting me up when I was down—without her, I am very sure I wouldn't have been able to write this.

This work was supported by an AHRC Scottish Graduate School for the Arts and Humanities doctoral training partnership studentship (grant no. 1804818) and a Society for Applied Philosophy Doctoral Scholarship.

Abstract. This thesis examines a range of cases in which it appears one's rights against harm are violated by another's behaviour, even though this behaviour has done one no harm. Call these cases of harmless wrongdoing. These cases raise a serious problem for most theories of rights, though the problem is most pronounced on the Interest Theory of Rights. According to that theory, rights necessarily protect their holder's wellbeing. At first glance, one might think that the person's wellbeing cannot be said to be protected by the right in cases of harmless wrongdoing because they are not harmed in such cases—so, the necessary condition set for the ascription of a right is not satisfied. I offer a novel, welfare-based explanation of why we have rights against harmless wrongs, the Safety Condition. This holds that for someone to hold a right against us that we not perform some action, we look to whether our performing that action could easily leave them sufficiently worse off to place us under a duty. In addition to extensional accuracy, one reason for this focus on modality—on what might have been—is that it removes an objectionable form of luck from rights. And, it matters that rights do not depend on luck in this objectionable way for this requires that we, as duty-bearers, are sensitive to others' wellbeing: that we do not only not harm others, but that we could not easily have harmed them.

Contents

PART I: RIGHTS	5
1. INTRODUCTION	7
1. <i>Why Rights?</i>	7
2. <i>The Problem of Harmless Wrongs</i>	10
3. <i>The Argument</i>	12
4. <i>Rights Theory</i>	15
4.1 Hohfeld	15
4.2 Correlativity	17
4.3 Abstract Versus Specified Rights	20
4.4 All-Things-Considered Versus Pro Tanto View	22
5. <i>Harmless Wrongs</i>	25
2. THE WILL THEORY	27
1. <i>Introduction</i>	27
2. <i>Introducing the Will Theory</i>	28
3. <i>A Terminological Refinement</i>	32
4. <i>Undergenerations of Rights</i>	34
5. <i>Being Revisionary About Rights</i>	36
5.1 Being Revisionary About Right-Holders	37
5.2 Normative Control	40
6. <i>What's in Being a Power-Holder?</i>	43
7. <i>Different Measures of Control</i>	47
8. <i>Conclusion</i>	51
3. THE INTEREST THEORY	53
1. <i>Introduction</i>	53
2. <i>Introducing the Interest Theory</i>	54
3. <i>Unpacking the Interest Theory</i>	56
3.1 Justificatory and Nonjustificatory	57
3.2 Sufficient Weight, Harm, and Wellbeing	59
3.3 Other Considerations	63
3.4 Rights Beyond Claims	65
3.4.1 Liberties	65
3.4.2 Powers and Immunities	67
4. <i>Returning to the Problem of Harmless Wrongs</i>	69
PART II: HARM	73
4. PREEMPTION	75
1. <i>Introduction</i>	75
2. <i>Revising the Necessary Condition</i>	77

2.1 Pro Tanto	77
2.2 Normality.....	82
3. <i>The Temporal Account of Harm</i>	86
3.1 Other Reasons to be Sceptical of the Counterfactual Account.....	87
3.2 Introducing the Temporal Account	88
3.3 Temporal Problems	89
4. <i>The Noncomparative Account</i>	93
4.1 Introducing the Noncomparative Account	93
4.2 Noncomparative Problems	97
5. <i>Back to the Counterfactual Account</i>	99
5. SAFETY	103
1. <i>Introduction</i>	103
2. <i>The Safety Condition and Preemption</i>	104
2.1 Introducing Safety	104
2.2 Closeness.....	107
2.3 Where Safety Comes into the Picture	109
3. <i>In (Partial) Favour of Safety</i>	110
3.1 Why Safety?	111
3.2 Other Cases of Harmless Wronging.....	113
3.2.1 Suboptimal Benefit.....	113
3.2.2 Suboptimal Supererogation	114
4. <i>In Defence of Safety</i>	117
4.1 The Obscure Problem	117
4.2 Very Preempted Harm.....	122
4.3 Safety and Harm	126
5. <i>Conclusion</i>	129
PART III: RISK.....	131
6. RISK AND HARM	133
1. <i>Introduction</i>	133
2. <i>The Desire Solution</i>	135
2.1 Actual and Informed Desires.....	136
2.2 Derivative and Non-Derivative Desires.....	139
3. <i>The Autonomy Solution</i>	142
3.1 Autonomy and Harm	142
3.2 Against the Autonomy Solution	145
3.2.1 Theoretical Commitments	145
3.2.2 Non-Autonomous Individuals.....	147
3.2.3 Roulette and Autonomy.....	148
3.3 Moving Beyond Autonomy	149
4. <i>The Magnitude of Harm Problem</i>	150
5. <i>Conclusion</i>	155

7. FACTS, EVIDENCE, OR BELIEFS?	157
1. <i>Introduction</i>	157
2. <i>The Distinction</i>	159
3. <i>Beliefs</i>	161
4. <i>Evidence</i>	164
5. <i>Facts</i>	171
4.1 Fact-Relative Oughts.....	171
4.2 Jackson Cases.....	174
5. <i>Conclusion</i>	179
8. SAFER STILL	181
1. <i>Introduction</i>	181
2. <i>The Safety Condition and Pure Risk</i>	182
2.1 Making Safety Safer	182
2.2 Preemption, Pure Risk, and Safety.....	185
3. <i>In Favour of Safety</i>	188
3.1 Protection from Harm	189
3.2 Freedom and Autonomy	190
3.3 Sensitivity.....	192
4. <i>In Defence of Safety (Again)</i>	196
4.1 Violations.....	196
4.2 Non-Violations	199
4.3 Overgenerations	201
4.3.1 Why Proliferations and Conflicts Are Not Themselves Worrying.....	202
4.3.2 Ruling Out Worlds.....	204
4.3.3 Other Considerations.....	208
5. <i>Conclusion</i>	210
BIBLIOGRAPHY.....	213

Part I: Rights

1. Introduction

1. Why Rights?

Rights are important. They are rhetorically important. Often, public debate is held in the language of rights. The debate tends not to be whether people *may* use the bathroom of the gender with which they identify, but whether they have the *right* to do so. Rights are also important for non-rhetorical reasons. They are often taken to offer a particularly robust protection against certain forms of conduct. For example, they are taken to place constraints on promoting the good. Plausibly, it is impermissible to kill a healthy patient in order to donate their organs to save five sick patients, even when doing so would maximise the good. A typical explanation for this verdict is that the healthy patient has a right against being killed that cannot be overridden by the mere fact that killing the patient would bring about more good. In this way, rights are often seen as imposing side-constraints on others' behaviour, as 'trumps' over other types of considerations, as providing others with 'exclusionary reasons', and so on.¹

Rights are also non-rhetorically important because of their underlying deontic structure. The focus of this thesis is primarily on the paradigm form of rights, the claim-right. Claim-rights correlate with duties. We get more precise below, but the idea is that I have a claim-right that you not kill me just in case you are under a duty not to kill me.² However, this duty is not just any sort of duty. It is a *directed* duty. Some duties are *undirected*—they are not owed to anyone in particular. For example, some people think that duties of beneficence and duties to care for the environment are undirected.³ Some duties are *directed*—they are *owed* to particular parties. When I promise my friend something, I am under a duty to satisfy the promise. The promissory duty is directed towards, owed to, my friend.

¹ See, respectively, (Nozick 1974; Dworkin 1978; Raz 1986).

² See this chapter, section 4.1.

³ This is not to imply duties to rescue particular people are not directed. For example, intuitively one owes it to Singer's drowning child to save it.

This is the underlying structure of rights that I am claiming is important—rights and their correlative directed duties are relational.⁴

Now, one might act wrongly by failing to satisfy either an undirected or directed duty. When one does so, it is important whether one has infringed an undirected or directed duty for at least the following two connected reasons. First, if I were to infringe a directed duty, I would not merely act wrongly, as when I infringe an undirected duty, but would also wrong the person to whom I owe that duty.⁵ Because of this, there are normative upshots specific to the infringement of directed duties. If a duty is directed, many think that it is demandable on the behalf of the party to whom it is owed and that its violation triggers apology owed to that party. (Correlatively, they tend to think that there is special standing for blame and forgiveness on the part of the party to whom the duty was owed.)⁶ If a duty is undirected, it is not demandable on behalf of a particular party, and its violation does not trigger further duties owed to particular parties. Second, since directed duties correlate with rights, and most think that rights are enforceable (other things being equal), directed duties are enforceable.⁷ They are potentially enforceable before as well as after the right's violation. Undirected duties do not correlate with rights, so are thought not to be enforceable.

So, rights are important because they allegedly offer this particularly robust form of moral protection and because of how they connect us as moral agents and moral patients.

⁴ There is a growing emphasis on rights' relationality: (Cruft 2013a; Gilbert 2018; Cruft 2019; Wallace 2019).

⁵ I am tempted to think that, at least for many directed duties, we are under only directed duties. Particularly, for duties capable of forfeiture or waiver. When I culpably attack someone, I forfeit my rights against being harmed by them. Forfeiture explains why it is not wrong for my victim to defensively harm me. When I consent to someone slapping me, I waive my rights against being slapped by them. Waiver explains why it is not wrong for the person to whom I have consented to slap me. If we were under an undirected duty not to harm in addition to directed duties not to harm, we could not explain why it is not wrong to harm those who have forfeited or waived their rights against harm. For, despite the forfeiture or waiver extinguishing the directed duty, the undirected duty would still exist. This means the duty not to harm must be only a directed duty, and not also an undirected duty. This argument generalises for all duties capable of forfeiture or waiver. If this analysis is correct, this gives us another reason to care about directed duties and their correlative rights—they alone populate a lot of interpersonal morality.

⁶ (Skorupski 2010; Darwall 2013; Cruft 2019).

⁷ (Thomson 1990, 105–22; James 2003; Cruft 2013b; Wenar 2013; Flanigan 2019).

Because of this, a theory of rights (and, correlatively, directed duties) must explain the following two connected features.⁸ First, it must offer an account of what it is to *owe* a duty to another individual. Second, it must offer an account of why, through infringing a directed duty, the duty-bearer does not merely act wrongly but wrongs the person to whom she owes that duty.

The focus of this thesis is with the

Interest Theory of Rights (Canonical). For X to have a right against Y that $Y \Phi$, X 's wellbeing (her interests) must be of sufficient weight to place Y under a duty to Φ .⁹

The Interest Theory has a lot going for it. These reasons are properly developed below, but here is a taster. The Interest Theory offers a plausible account of the *grounds* of rights—rights are grounded in the wellbeing of their holders. Further, it does a nice job explaining why Y owes her duty to X and why Y does not merely act wrongly but wrongs X if she infringes her duty. The duty owes its existence to X . If Y infringes that duty, Y has failed to respond to morally salient duty-grounding features of X .

⁸ I have not included that a theory of rights must explain why rights are thought to offer a particularly robust form of moral protection for, as Feinberg and Cruft note, this role could be achieved with undirected duties only (Feinberg 1980; Cruft 2019). This is because undirected duties require just as much of us as undirected duties. Again, we act wrongly whether we satisfy an undirected or directed duty. Because of this, whether the duty is directed or not makes no difference to what we were required to do—it makes no *extensional* difference to the normative landscape. This is the sense in which Cruft says we could do without rights. This raises the question, if rights make no extensional difference, what difference do they make? The answer, I think, lies in rights' second important role—their relationality. Making the case for this is not the primary focus of this thesis, though does crop up in the final analysis. It is also something I look to make more of moving forward.

⁹ There are lots of clarifications that may be prompted. See chapter 3: especially note 69 and section 3. Throughout, when it comes to necessary conditions, I mean for these claims to be *explanatory* (that is, I mean “only if, and partly because”). Yet I avoid structuring the conditions in this way for stylistic reasons. So, when it comes to Interest Theory (Canonical), we could rewrite the condition as: X has a right against Y that $Y \Phi$, only if, and partly because, X 's wellbeing (her interests) is of sufficient weight to place Y under a duty to Φ .

2. The Problem of Harmless Wronging

But there is a problem with the Interest Theory. It is possible that one may have a right against another's action, and yet not be harmed by that action. This is problematic because, if one is not harmed by the action's occurrence, one's wellbeing is not of sufficient weight to place others under duties.¹⁰ But this means the necessary condition set for the ascription of a right on the Interest Theory has not been satisfied, and so our right-holder will not actually hold a right against that action.

Here are two examples:

Plane Crash (Preempted Harm). Passenger is about to board a plane. Attendant takes a disliking to Passenger, so denies her admittance onto the plane. On departure, the plane crashes and everybody on board dies.

Roulette (Pure Risk Imposition). Target is asleep. Her housemate, Shooter, comes into her room and decides to play Russian roulette with her. Luckily, no bullet is fired. Shooter, content with having had a round of roulette, will never play roulette again.

Intuitively, Passenger and Target have their rights violated. However, Passenger is better off in the world in which her rights are violated than that in which they are not. Target's life is as it would have been had Shooter not made Target the subject of her risky behaviour. Given the standard *Counterfactual Account of Harm*, on which Y harms X iff Y makes X worse off than X would have been had Y not acted as she did, this means that Passenger and Target are not harmed by the violation of their rights. Because of this, it is hard to see how either Passenger's or Target's wellbeing is of sufficient weight to place Attendant and Shooter under their respective duties. But this means the necessary condition set for a right-attribution by the Interest Theory is not satisfied. Call this the Problem of Harmless Wronging for the Interest Theory. The problem is that of accommodating our intuitions that Passenger and Target have their rights violated, given a commitment to the

¹⁰ This claim is defended in chapter 3, section 3.2.

Interest Theory. More generally, the problem is that of accommodating agents having rights against harmless wrongs, given a commitment to the Interest Theory.

This thesis offers a principled solution to the Problem of Harmless Wronging for the Interest Theory by revising the canonical statement of the theory with what I call the Safety Condition. The Safety Condition looks beyond what happens in the actual world to close possible worlds to normatively ensure that right-holders' wellbeing is robustly protected across circumstances that could easily come about. In *Plane Crash*, it could easily have been that the plane Passenger would have caught, were she not to have been denied admittance, would not have crashed.¹¹ Were this to have been the case, Attendant would have harmed Passenger by denying her admittance onto the plane. In *Roulette*, it could easily have been that there was a bullet in the chamber when Shooter pulls the trigger. Were this to have been the case, Shooter would have harmed Target by playing roulette with her. It is in virtue of these nearby possibilities that Passenger and Target hold rights against the respective actions in the actual world.

I argue that the Safety Condition correctly generates rights in our two cases of harmless wrongdoing. Further, it offers a unified solution of *why* people are attributed rights against harm in these different types of cases of harmless wrongdoing: even if one is not made sufficiently worse off as things turn out, one easily could have been. The Safety Condition's *principled* extensional accuracy is the primary virtue of the account that I would like to stress in this thesis.¹² I also argue that we have reason to endorse the Safety Condition in addition to its extensional accuracy. Roughly, this focus on how things could otherwise have been formally requires that duty-bearers are sensitive to others' wellbeing: it ensures not only that we do not harm others, but that we could not easily have harmed others.

¹¹ This is ambiguous between the close possible world in which (a) Passenger is denied admittance onto the plane, the plane takes off, and lands, and (b) Passenger is not denied admittance onto the plane, the plane takes off, and then lands. As we see in chapter 5, it is (a) that is relevant.

¹² We see further in chapter 8, section 2.2, that the two cases of harmless wrongdoing differ from mundane cases of *harmful* wrongdoing in symmetrical ways to each other. This gives us reason to think the Problem of Harmless Wronging is a unified problem that we should want a unified solution to.

3. The Argument

Here is the shape of this thesis. It is made up of three parts. Part I concerns rights. By way of motivating the Interest Theory, chapter 2 examines its main rival, the Will Theory of Rights. The Will Theory says that rights are grounded in normative control. I argue defenders of the Will Theory face a dilemma. They can remain faithful to the Will Theory's account of control as the grounds of rights, in which case the theory counterintuitively undergenerates right-ascriptions. Or, they can revise the Will Theory to increase its extensional accuracy, but at the cost of obscuring the theory's focus on normative control.

Chapter 3 shows how the Interest Theory's focus on wellbeing as the grounds of rights solves the problems facing the Will Theory identified in chapter 2. Though that chapter focuses on the Will Theory, the arguments generalise: to avoid analogous problems to those raised with the Will Theory, we need to recognise wellbeing as the grounds of at least some rights. (And, once we open the door to wellbeing as the grounds of at least some rights, we open the door to the Problem of Harmless Wronging.)¹³ I also argue there are other good reasons to endorse the Interest Theory. The chapter then introduces the Interest Theory in sufficient detail to be the focus for the remainder of the thesis, before returning to the Problem of Harmless Wronging.

Part II concerns the problem posed by cases of preempted harm. Call this the Problem of Preemption. The problem is that of accommodating our intuitions that people have rights against harms that are preempted, as in *Plane Crash*, given a commitment to the Interest Theory. The Problem of Preemption is a more specific version of the Problem of Harmless Wronging. Chapter 4 begins by showing that we cannot weaken the Interest Theory (Canonical) in two seemingly plausible ways to solve the Problem of Preemption: first, by saying that rights are grounded in wellbeing considered pro tanto; second, by saying rights are grounded in our wellbeing under normalcy. Another tempting way to solve the Problem of Preemption is to replace the Counterfactual Account of Harm. Chapter 4 argues against proceeding in this way on the grounds that alternative accounts of harm are more

¹³ It is because of this that the Will Theory is my main focus for a rival to the Interest Theory. See chapter 3, section 4, in which I explicitly argue that the Problem of Harmless Wronging applies to other theories of rights in addition to the Interest Theory.

problematic than the Counterfactual Account even with the noted problem of preempted harm. Chapter 5 then introduces the Safety Condition as a solution to the Problem of Preemption. It offers a taster of why one ought to endorse the Safety Condition in addition to extensional adequacy, and defends the Safety Condition against a few objections.

Part III is concerned with the problem posed by pure risk. Call this the Problem of Pure Risk. The problem is that of accommodating our intuitions that people, sometimes, have rights against being exposed to risk of harm. Now, if risk of harm were *itself* harmful, cases of pure risk imposition would not be cases of harmless wronging. There would be no Problem of Pure Risk. Chapter 6 argues that risk of harm is not itself harmful. For those who do not follow me all the way to thinking that risk is not harmful, I argue, at the least, risk is harmful only in a small range of cases. If risk is harmful only in a subset of the much larger set of cases in which it looks like we have rights against risk of harm, we cannot answer the Problem of Pure Risk by arguing that risk is itself harmful.

Another seemingly plausible way to respond to the Problem of Pure Risk is by arguing that rights depend on duty-bearers' beliefs or the evidence available to them, rather than the facts. For example, in *Roulette* we can suppose Shooter thinks there is a one in six chance that she will kill Target, which the best available evidence also supports. If rights depend on either duty-bearers' beliefs or the evidence available to them, Target's wellbeing is of sufficient weight to place Shooter under a duty not to play roulette with her. Chapter 7 argues that rights depend on the facts. So, the Problem of Pure Risk is not solved in this way. Finally, chapter 8 demonstrates that the Safety Condition solves the Problem of Pure Risk, before explaining why we have reason to endorse the Safety Condition in addition to its extensional accuracy.

Another way to see the shape of the thesis is by focusing on the Problem of Harmless Wronging. Here is the problem:

1. Assume the Interest Theory (Canonical): For X to have a right against \mathcal{Y} that $\mathcal{Y} \Phi$, X 's wellbeing (her interests) must be of sufficient weight to place \mathcal{Y} under a duty to Φ .
2. Whether or not X 's wellbeing is of sufficient weight to place \mathcal{Y} under a duty to Φ depends partly on the extent to which X is harmed by \mathcal{Y} 's not Φ -ing or benefited by \mathcal{Y} 's Φ -ing.
3. Assume the Counterfactual Account of Harm and Benefit: \mathcal{Y} harms X iff \mathcal{Y} makes X worse off than she would have been had \mathcal{Y} not acted as \mathcal{Y} did. \mathcal{Y} benefits X iff \mathcal{Y} makes X better off than she would have been had \mathcal{Y} not acted as \mathcal{Y} did.
4. In *Plane Crash* Attendant does not harm Passenger by denying her admittance onto the plane, nor would she benefit Passenger by allowing her onto the plane (from 3).
5. Passenger's wellbeing is not of sufficient weight to place Attendant under a duty not to deny her admittance onto the plane (from 2 and 4).
6. Passenger has no right against being denied admittance onto the plane (from 1 and 5).
7. Risk of harm is not itself harmful.
8. In *Roulette* Shooter does not harm Target by playing roulette with her, nor would she benefit Target by not playing roulette with her (from 3 and 7).
9. Rights are determined by the facts.
10. Target's wellbeing is not of sufficient weight to place Shooter under a duty not to play roulette with her (from 2 and 8).
11. Target has no right against Shooter playing roulette with her (from 1 and 10).

The Problem of Preemption is the conclusion reached in 6. The Problem of Pure Risk is the conclusion reached in 11. These problematic conclusions could be avoided by denying any of the other premises. Chapters 2 and 3 defend the first premise. Chapter 4 defends

the first and third premise. Chapter 5 then suggests we ought to revise the first premise with the Safety Condition to avoid the problematic conclusion in 6. This is in keeping with the arguments for the first premise from chapters 2 and 3. (In section 2.3 of chapter 5, we also see that the Safety Condition could be seen as revising the second rather than the first premise. Roughly, the idea is that we could read whether or not X 's wellbeing is of sufficient weight to place \mathcal{Y} under a duty to Φ as depending on whether X is harmed by \mathcal{Y} 's not Φ -ing or benefited by \mathcal{Y} 's Φ -ing *across close possible worlds*.) Chapter 6 defends the seventh premise. Chapter 7 defends the ninth premise. Chapter 8 then shows how the revision to the first premise from chapter 5 extends to avoid the problematic conclusion in 11.

4. Rights Theory

Before beginning, a little rights theory is required to form a background for the rest of the thesis. 4.1 introduces the Hohfeldian framework. 4.2 unpacks and then defends Hohfeldian correlativity. 4.3 discusses how we can have rights at different levels of specificity and argues that the Problem of Harmless Wronging cannot be circumvented by focusing on rights at an abstract level. 4.4 discusses how rights might connect to what we are required to do all-things-considered.

4.1 Hohfeld

This thesis operates within Wesley Hohfeld's structural analysis of rights (Hohfeld 1919). Hohfeld suggested that all rights refer to a relation between two parties. Let X refer to the right-holder, \mathcal{Y} to the person against whom she holds that right, and Φ refer to the action X has a right to the performance (or non-performance) of.¹⁴ First,

X has a claim-right that \mathcal{Y} Φ , against \mathcal{Y} , iff \mathcal{Y} is under a duty to Φ , owed to X .

¹⁴ Φ -ing should be read broadly enough to include omissions. For discussion of whether we hold rights to things other than the actions or non-actions of others, see (Cruft 2004, 350).

If you have a claim-right against me that I give you my pen, I am under a duty, owed to you, to give you my pen. I say more about correlativity between claim-rights and directed duties in the following subsection.

Second,

X has a liberty to Φ , against Y , iff X is under no duty not to Φ , owed to Y .¹⁵

If I have a liberty-right against you that I not give you my pen, I am not under a duty, owed to you, to give you my pen. Many think that I can have a liberty-right, against you, that I not give you my pen, but be under a duty to give you my pen nonetheless—it is just that *this* duty cannot be owed to you. For example, I may have promised a third-party that I will give you my pen. Were I not to give you my pen, I would not wrong you, because I did not *owe* my duty to you, but would wrong the third party. Inasmuch, liberties are directed just as claims are directed.¹⁶

Claims and liberties are *first-order* incidents. There are also *second-order* incidents.

X holds a power over Y iff X is able to alter the normative relationship between Y and herself (or between Y and some other party).

From Y 's perspective, she is *liable* to X changing their normative relationship. Suppose that I have the power-right to make promises with respect to my pen. If I promise to give you my pen, I have changed our normative relationship from one in which I held a liberty-right against you not to give you my pen, to one in which you hold a claim-right against me that I give you my pen. Alternatively, I might make it that you have to share my pen with some third-party, thereby altering the normative relationship between you and them.

¹⁵ I have not baked into our definition of liberties, powers, and immunities that they are always liberty-, power-, or immunity-*rights* (see note 36 in chapter 2 and note 68 and section 3.4 in chapter 3). Whereas, I have baked into our definition of claims that all claims are rights (though *cf.* the following subsection); this is because I see the theories of the nature of rights as accounts of *what it is* to owe someone a duty.

¹⁶ E.g., (Van Duffel 2012b, 107). *Cf.* (Wenar 2015)'s definition of liberties that do not account for the directionality of liberties (X has a liberty to ϕ iff X is not under a duty to ϕ).

Second-order incidents also interact with other second-order incidents. I might bestow upon you the power-right to decide what to do with the pen.

Finally,

X holds an *immunity* over *Y* iff *X* is *not liable* to *Y* changing their normative relationship (or between her and some other party).

Before I give you my pen, I might hold an immunity-right that you not divest me of my rights to my pen.

Often, when we hold a “right” against someone, there is actually a combination or cluster of these Hohfeldian incidents at play. For example, my right that you not hit me contains, at the least, a claim that you not hit me, a power to waive my claim that you not hit me, and liberties to exercise or not exercise this power. For the remainder, unqualified references to “rights” refer to claim-rights (though, I do include “claim-” in definitions). This is because claim-rights are going to be our focus.

4.2 Correlativity

Claim-rights are going to be our focus for two reasons. First, they are the only type of right that, when all else is equal, requires the performance (or non-performance) of certain actions. Second, relatedly, they are the only type of right that are violable. And, this thesis is focused on cases in which the violation of rights to the performance (or non-performance) of certain actions does not cause harm—what is more, because our focus is on the harmless violation of claim-rights, they are also of interest because their violation alone triggers the interesting normative upshots introduced above (in section 1).¹⁷

¹⁷ This is not to say that the other incidents are not normatively important in virtue of lacking these upshots. This is because the other incidents are all bound up with claims. For example, since liberties are the absence of directed duties *on one* (the absence of others holding a claim against one), having a liberty to Φ means, other things being equal, one can Φ without triggering any of these normative upshots. See chapter 3, section 3.4.

Since the claim-right is going to be our focus, and there is this correlativity at the heart of claim-rights, let us expand on our definition of a claim from above. Particularly, let us call the definition that appeared above

Strong Correlativity. X has a claim-right that $Y \Phi$, against Y , iff Y is under a duty to Φ , owed to X .

Some deny Strong Correlativity—they think that one can be owed a duty that another Φ , without having a right against them that they Φ . But they do tend to endorse

Weak Correlativity. X has a claim-right that $Y \Phi$, against Y , only if Y is under a duty to Φ , owed to X .

On this view, all claim-rights correlate with directed duties, but not all directed duties correlate with claim-rights. I do not find it immediately obvious that I can be owed things but not have rights to those things.¹⁸ So, I am tempted to accept Strong Correlativity.

Still, perhaps there are counterexamples to Strong Correlativity. Duties of friendship seem to be directed, though do not always seem to correlate with rights (you might owe me a duty to repair our friendship, but do I have a right that you do so?). Similarly, duties of gratitude seem directed, but we might be sceptical of rights to be shown gratitude. Some have argued that these examples do not merely speak in favour of Weak Correlativity, but also that they offer an explanation of when it is that directed duties correlate with claim-rights. This is because these people think that a salient feature of these two examples is that it is inappropriate to enforce the duties. Because of this, they endorse

Enforceability. *Ceteris paribus*, a directed duty correlates with a claim-right iff that duty is appropriately enforceable (Cruft 2013b, 209; Wenar 2013, 214).

I take it “*ceteris paribus*” excludes instances in which rights are overridden, and so enforceability is impermissible. Enforceability will also be impermissible if the only means of

¹⁸ Though, I do not go as far as Rainbolt in thinking that it is ‘extensional[ly] counterintuitive’ to think that one can be owed something, though not have a right to it (Rainbolt 2006, 115).

doing so are disproportionate. For example, it is impermissible for me to enforce my right that you not steal my pen by breaking your leg. But I take it this is the work that “appropriately” is supposed to be doing in the principle.

Happily, we do not need to settle this dispute between Strong and Weak Correlativity here. If we endorse Weak Correlativity, we can endorse the Interest Theory, for example, as an account of directed duties; we can then see the Interest Theory *and* Enforceability as an account of rights. I introduce Strong and Weak Correlativity here without settling the debate to make clear that we are after an account of what it takes to owe a duty to another. For ease of exposition, let us begin by assuming Strong Correlativity.¹⁹

(Notice, we have a distinction between directed and undirected duties, and between Weak and Strong Correlativity. These distinctions cut across each other, so there are several combinations of views here.²⁰ Some people think all directed duties correlate with rights and that there are no undirected duties (Kramer 2000). Some people agree there are no undirected duties but think some directed duties do not correlate with rights (Wallace 2019). Some people agree that some directed duties do not correlate with rights but think there are undirected and directed duties (Cruft 2019). And, finally, some agree there are both directed and undirected duties but think all directed duties correlate with rights (Rainbolt 2006). And, to complete our picture, some people think there are no directed duties—and so, *a fortiori*, no rights (Arneson 2001; Bedi 2009).)

Before moving on, let me say a final thing about correlativity: some people deny its existence altogether. The thinking here tends to be that rights cannot be the logical correlatives to duties since rights are the *grounds* of duties. For example, Joseph Raz writes: ‘A right of one person is not a duty on another. It is the ground of a duty, ground which, if not counteracted by conflicting considerations, justifies holding that other person to have the duty’ (Raz 1986, 171). Similarly, Jeremy Waldron thinks that correlativity must be denied ‘if rights are to be seen as justifications: one statement cannot justify another if the two are

¹⁹ Cf. chapter 2, section 5.1.

²⁰ They also cut across the distinction between perfect and imperfect duties.

equivalent' (Waldron 1988, 84; 1985, 10–11; MacCormick 1982, 161; Dworkin 1978, 171).

Yet, rights and duties can be logical correlatives while one stands in a justificatory relationship to the other. This is because the grounding relationship—the “in virtue of”, the “because of”—is asymmetrical.²¹ Take an example from Matthew Kramer. I could desire a slope in my garden *because of* the fact that I want to enjoy the downwards perspective from the top. Yet, this slope will also provide me with an upwards perspective from the bottom of my garden. Even though either perspective always comes with the other perspective, one can ground the other.²² So, this is not a good reason to doubt correlativity.

4.3 Abstract Versus Specified Rights

In *Plane Crash*, we are focusing on Passenger's right that Attendant not deny her admittance onto the plane and, in *Roulette*, Target's right that Shooter not play roulette with her. One might note that these are fairly specific rights. However, does not Passenger have a right not to be arbitrarily discriminated against and Target a right not to have a gun fired at her? Further, do not Passenger and Target hold these rights against the world (*in rem*)? What is more, though Passenger's and Target's wellbeing might not be of sufficient weight to place their respective parties under duties not to perform the specific actions that we focus on, is not Passenger's wellbeing of sufficient weight to place others under duties not to arbitrarily discriminate against her? Is not Target's wellbeing of sufficient weight to place others under duties not to harm her? Might one object, the Problem of Harmless Wronging for the Interest Theory only gets off the ground because I focus on these fairly specific rights? There is no problem if we focus on Passenger's and Target's rights at a more abstract level.

The answer to this problem comes from looking back to

²¹ See, among others, (Berker 2017).

²² See (Kramer 2000, 39), though he does not relate this to the asymmetrical nature of grounding. Another reason some are sceptical of correlativity is they think the same right can ground many duties: ‘there is no closed list of duties which corresponds to the right [...] A change of circumstances may lead to the creation of new duties based on the old right’ (Raz 1986, 171). The response to this argument can be found in the following subsection.

Strong Correlativity. X has a claim-right that $Y \Phi$, against \mathcal{I} , iff Y is under a duty to Φ , owed to X .

Correlativity needs to be maintained at the same level of specificity (Kramer 2000, 42). When we discuss Passenger's (abstractly stated) right not to be discriminated against, this entails Attendant's duty, owed to Passenger, not to discriminate against her. When we discuss Passenger's (specifically stated) right not to be denied admittance onto the plane, this entails Attendant's duty, owed to Passenger, not to deny Passenger admittance onto the plane.

Speaking about rights and their correlative duties at the abstract level is rhetorically useful. They make nice premises in arguments, for example. However, it also comes with its drawbacks. Principally, it is not obvious what an abstract duty not to discriminate against others entails. To work out exactly what it entails, we need to look to duties at a more specified level. And that is where the problem of Harmless Wronging creeps back in. Passenger may well have an abstract right, against Attendant, not to be discriminated against, but this does not entail that she has a right against Attendant that she not deny her admittance onto the plane (at least, on the canonical formulation of the Interest Theory)—for, her wellbeing is not of sufficient weight to place Attendant under a duty not to deny her admittance onto the plane.

Think about a different right. For example, Ann's right not to be harmed. Does that abstract right mean that Beth is under a duty not to perform some specific action that harms Ann? Well, it depends. Perhaps Ann has waived her right against Beth performing that specific action. Perhaps Ann has forfeited her right against Beth performing that specific action. Perhaps Ann simply never had a right against that specific action, though it is harmful—for example, Ann has no right that Beth not return Ann's love, Ann has no right that Beth give Ann the last lifejacket, and so on.²³ We need, then, to look to the potential specific right and specific duty.

²³ Some people are sceptical of forfeiture and think that when we would usually say someone has “forfeited” their rights, the pro tanto duty not to harm remains (and so does the pro tanto right) but has been overridden (Thomson 1986a, 33–37; Griffin 2008, 65–67); *cf.*, (Thomson 1991). This view seems problematic since it cannot distinguish between victims justifiably defending themselves against unjustified threateners, where

4.4 All-Things-Considered Versus Pro Tanto View

In this final subsection, I speak a little of the relationship between rights and all-things-considered requirements. I rely on this discussion in a few places throughout, but it is principally here to provide a foundation for how rights relate to the rest of the normative domain.

While we are assuming Strong Correlativity, that much does not settle the nature of the correlative duty. Some people endorse the

All-Things-Considered View. If X has a claim-right against Y that $Y \Phi$, Y is required to Φ .²⁴

“Required” should be thought of as an all-things-considered notion.²⁵ There might be good reason to endorse the All-Things-Considered View. As Judith Jarvis Thomson writes, ‘very often when we ought [or are required] to do a thing, we ought [or are required] to do it precisely because someone has a claim[-right] against us that we do it’ (Thomson 1990, 79). If the All-Things-Considered View is not correct, our conception of

unjustified threateners have forfeited their rights against being harmed, and people harming others with a lesser evil justification, where those whom one harms have not forfeited their rights against being harmed. In the case of self-defence, there is no moral remainder—no explanation, apology, or compensation owed. In the case of lesser evil justifications, there is a moral remainder. Appealing to the pro tanto right not to be harmed (and correlative pro tanto duty not to harm) in the case of the lesser evil justification explains this moral remainder. We focus more on the moral remainder in the following subsection.

²⁴ Among others, (Wellman 1995; Shafer-Landau 1995; Oberdiek 2004; Steiner 2013). The All-Things-Considered View is only conditional, rather than biconditional, because of the directionality of claim-rights, the possibility of undirected duties, and the possibility of other factors contributing to all-things-considered requirements. Suppose that Beth does not have a claim-right, against Ann, that Ann mow Beth’s lawn. If the All-Things-Considered View were biconditional, this would mean that Ann is not required to mow Beth’s lawn. But Ann may nonetheless be required to mow Beth’s lawn. For example, Ann may have promised a third party that she will mow Beth’s lawn. Or, for whatever reason, Ann may be under an undirected duty to mow Beth’s lawn. Or, if other factors contribute to the all-things-considered requirement in addition to duties, Ann may be required to mow Beth’s lawn without even being under a duty to do so.

²⁵ The notion that I am calling “requirement” is often called “ought”. However, as Snedegar among others has shown, “ought” is too weak. “Requirement” (or “must”, “have to”, and so on) is the stronger notion we are after. This can be seen from the felicity of “You ought to help that old lady cross the road, but you’re not required do” and the infelicity of “You are required to help that old lady, but it’s not the case that you ought to”. The first of these sentences show that “ought” does not entail “requirement”. The second suggests that “requirement” does entail “ought” (Snedegar 2016, 159–62).

what we are required to do will be more complicated. We must answer the question of what a right's content is, and then 'the separate question of a right's normative implications' (Oberdiek 2008, 128).²⁶

Some people deny the All-Things-Considered View. They endorse the

Pro Tanto View. If X has a claim-right against Y that $Y \Phi$, Y has a pro tanto duty to Φ .²⁷

On this view, rights and their correlative duties are pro tanto considerations: they have 'genuine weight, but nonetheless may be outweighed by other considerations' (Kagan 1989, 17). Since rights do not entail requirements, a defender of the Pro Tanto View must say something about how rights interact with requirements. Following Thomson, let us say, if X has a claim-right against Y that $Y \Phi$, and Y does not Φ , she *infringes* X 's right; if Y does not Φ and she is required to have Φ -ed, she *violates* X 's right to Φ (Thomson 1986a, 40; 1986b, 51–55; 1990, 122).

Why might one prefer the Pro Tanto View? Suppose that Ann owns some land. Her neighbour, Ben, falls ill and needs to get home to take some medicine. Time is tight and the quickest way for Ben to get home is to cut across Ann's land.²⁸ It seems plausible both that:

- (1) Ann has a claim-right against Ben that he stay off her land,

and that

²⁶ Only claim-rights easily translate into requirements. Ann's having a liberty-right to Φ against Beth means neither that Ann is not required to Φ nor that she is required to Φ . Rather, it means only that she is not under a duty, owed to Beth, not to Φ . It is entirely open whether she is required to Φ . We face similar issues when translating rights into talk of reasons. Suppose that X has a (claim-)right against Y that $Y \Phi$. It is plausible that Y thereby has a reason to Φ . However, what if Y has a liberty-right not to Φ against X . This does not mean that Y has reason not to Φ ; she might have plenty of reason to Φ , and no reason not to Φ . For example, suppose Ann has a liberty-right, against Beth, not to mow Beth's lawn. Does this mean Ann has no reason to mow Beth's lawn? No: she may have plenty of reasons.

²⁷ Among others, (Thomson 1990; Feinberg 1980; Kamm 2007; Cruft 2019).

²⁸ The case is taken with slight alternations from (Thomson 1990, 98).

(2) Ben is not required to stay off Ann's land.

Both of these things cannot be true if the All-Things-Considered View is true. The combination of the All-Things-Considered View and (1) implies the negation of (2).

Similarly, suppose we have a standard trolley case in which a trolley is headed towards some number of people, which it will kill if not stopped. Bystander can divert the trolley onto a side-track. Unfortunately, there is a person on the side-track (call this person The One). It seems plausible both that:

(3) The One has a claim-right that Bystander not kill her,

and yet that there is some number of people that makes it permissible for Bystander to divert the trolley. Suppose we have that number of people. This means

(4) Bystander is not required not to turn the trolley (i.e., it is permissible for Bystander to turn the trolley).

The combination of the All-Things-Considered View and (3) implies the negation of (4).

Now, to accommodate cases like this (cases in which it looks like X has a claim-right against Y that $Y \Phi$, and yet it is false that Y is required to Φ), defenders of the All-Things-Considered View suggest that the content of X 's right *specifies* that the right does not obtain in those particular circumstances. *Moral Specificationists* think that X only has a claim-right that Y not *unjustifiably* or *wrongfully* Φ (Oberdiek 2004; 2008). For example, The One has a claim-right that Bystander not unjustifiably turn the trolley onto her. *Factual Specificationists* think that X only has a claim-right that $Y \Phi$, unless: list of exemptive clauses (Shafer-Landau 1995). For example, Ann has a claim-right against Ben that he not cross her land, unless he needs to save his life, and so on.

While defenders of the All-Things-Considered View may be able to explain away this contradiction by endorsing Specificationism, one might wonder how they can deal with what is called the moral remainder. For example, if Ben causes some damage to Ann's land while permissibly cutting across it, some think he is liable to pay compensation to Ann. Similarly, even if Bystander acts permissibly in turning the trolley, it is plausible that

she still *wrongs* The One, owes her an apology, and so on. Yet, if there was no right violated in either case, as Specificationism admits, what might explain these features?

For this reason, let us assume the Pro Tanto View. As it happens, not too much turns on our choice—though I will draw on the moral remainder at times. The reason for this is two-fold. First, the Problem of Harmless Wronging applies to either view. Second, the debate between the All-Things-Considered and Pro Tanto Views is *neutral* on substantive first-order normative questions. A defender of the All-Things-Considered View says that a reason that countervails a putative right directly determines the content of the right (that it does not obtain in those particular circumstances). A defender of the Pro Tanto View says that this reason does not have any bearing on the right's existence but does impact what the putative correlative duty-bearer is required to do.²⁹

5. Harmless Wronging

Before beginning in earnest, a final preliminary is required. This thesis centres around discussion of two examples of harmless wrongdoing: preempted harm and pure risk imposition. One might wonder why we are not discussing more prominent cases of harmless wrongdoing found in the literature. For example:

Promise. Beth promises to meet Ann at the pub at 8. Ann only goes to the pub because Beth has promised to meet her. Ann would prefer, and it would be better for Ann, if Beth were not to show up. Beth does not show up.³⁰

²⁹ See Shafer-Landau: 'the problem of specifying the content of a right is the very same problem as that of [...] knowing when a right is permissibly infringed. In each case, conflicting moral considerations generate a moral conclusion either about what a full-fledged right contains, or about the circumstances under which a full-fledged right can be permissibly infringed' (Shafer-Landau 1995, 215). The views might not even differ on the moral remainder if Specificationists can build it into their view.

³⁰ E.g., (Owens 2012).

Paternalism. Carl has stopped smoking but, in a moment of weakness of will, has bought a packet of cigarettes. It would be better for Carl if Dave were to take the cigarettes to avoid his relapse into nicotine addiction.³¹

Property. Erica has some ugly garden gnomes left in her front garden from when she moved into her house. Fran quite likes the gnomes, so takes them while walking past. Erica does not even notice.³²

Neither Ann, Carl, nor Erica are harmed, all-things-considered, in these cases. Yet, intuitively, they have their rights violated. So, we have three more examples of harmless wrongdoing. If the Interest Theory (Canonical) is correct, we have a problem. We cannot account for why Ann, Carl, and Erica have their rights violated.

The reason we focus on our two cases of harmless wrongdoing, as opposed to these problems raised by harmless promise-breaking, paternalism, and trespass is because cases of preemption and pure risk concern rights against harm that are violated: Attendant's and Shooter's actions wrong Passenger and Target in a very similar way to if they *would* have straightforwardly harmed their victims. It is less obvious that *Promise, Paternalism, and Property* are about harm. However we solve harmless promise-breaking, paternalism, and harmless trespass, we should want to solve the Problem of Preemption and Pure Risk in a way that appeals to harm. The Safety Condition offers such an explanation.

Similarly, suppose we look at all these cases of harmless wrongdoing and conclude the Interest Theory is false. Other things being equal, we should still want harm to explain the grounding of Passenger's and Target's rights even if we do not think all rights are grounded in their holder's wellbeing. Those who deny the Interest Theory should still want a solution to the Problem of Harmless Wronging. Suitably refined, Safety Condition offers such an explanation to those who deny the Interest Theory.³³

³¹ E.g., (Feinberg 1988; Tadros 2016a; May 2017).

³² E.g., (Ripstein 2009; Cruft 2019).

³³ One would just deny the necessary aspect of the definition. We return to why the Problem of Harmless Wronging is problematic to other theories in chapter 3, section 4.

2. The Will Theory

1. Introduction³⁴

Suppose that Ann justifiably owns some property. She plays in its garden with her child. Carl, her neighbour, is under a duty not to drive across her garden. Beth does not have a garden, so plays in the park with her child. Carl is under a duty not to drive across the park. Carl drives across Ann's garden and the park, making both unfit for purpose. In both cases, Carl does not act as his duty dictates. Let us suppose, he harms Ann and Beth. Commonly, it is thought that Carl wrongs Ann but not Beth, and manifests disrespect for Ann but not for Beth. Perhaps this is because Ann holds a right that Carl not drive across her lawn, whereas Beth does not hold a right that Carl not drive across the park. Correlatively, Carl is under a directed duty, owed to Ann, not to drive across her garden, whereas he does not owe Beth a directed duty not to drive across the park.

Recall that I said that a theory of rights (and, correlatively, of directed duties) must explain the following two connected features. First, it must offer an account of what it is to *owe* a duty to another individual. Second, it must offer an account of why, through infringing a directed duty, the duty-bearer does not merely act wrongly, but wrongs the person to whom she owes that duty. We can add a third connected feature: why, through infringing a directed duty, the duty-bearer manifests disrespect towards the person to whom she owes that duty (Cruft 2013b).³⁵

Ann has control over Carl's duty not to drive across her lawn, whereas Beth has no control over his duty not to drive across the park. The Will Theory of Rights says that having control over another agent's duties grounds rights. Perhaps this explains why Carl owes

³⁴ At points in this chapter, I thank two anonymous reviewers from the *Canadian Journal of Philosophy*.

³⁵ The domain in which we find the right and correlative duty affects the nature of the wrong and disrespect. If the right is moral, we have a moral wrong and moral disrespect. If legal, we have a legal wrong and legal disrespect. For the most part, I speak about rights without specifying their nature. When this is salient (e.g., section 4), I distinguish between moral and legal rights. See also Wenar on being "game-wronged" and Cruft on students owing duties to their lecturers (Wenar 2013, 203–4; Cruft 2013b, 207); cf. section 5.1, where I consider whether wronging and directed duties come apart from rights.

his duty *to* Ann. If Carl infringes his duty, he fails to respect that Ann has normative control over him—this may explain why he wrongs and manifests disrespect towards her.

The Will Theory has commonly been objected to on the grounds that it undergenerates right-ascriptions along three fronts. This chapter systematically examines, in a manner not explored before, a range of positions open to a defender of the Will Theory in response to these counterexamples. I argue that none of these initially plausible ways of responding are wholly satisfactory. The argument generalises to form a dilemma for the defender of the Will Theory. Either, one can remain faithful to the Will Theory's account of control as the grounds of rights, in which case the theory undergenerates right-ascriptions. Or, one can revise the Will Theory to increase its extensional accuracy, but at the cost of obscuring the theory's focus on normative control. The dilemma applies for other theories that have the right-holder play some *active* role in right-holding (such as controlling the correlative duty, being able to demand compliance of the correlative duty, and so on); to avoid such problems we need to recognise wellbeing as the grounds of at least some rights.

Section 2 offers an exposition of the Will Theory. Section 3 defends the Will Theory against an important recent objection: that the paradigm Will Theory right-holder does not hold the necessary pair(s) of powers over another's duties. Section 4 considers the standard three dimensions along which the Will Theory undergenerates right-ascriptions. Sections 5 to 7 move beyond much of the literature on the Will Theory, systematically examining how the Will Theory might answer these undergenerations. We want both to be faithful to the Will Theory's account of the grounds of rights while properly accounting for the direction of duties.

2. Introducing the Will Theory

The Will Theory says that rights give agents normative control. As Neil MacCormick puts it, it recognises the right-holder's will as 'preeminent over that of others in relation to a given subject matter and within a given relationship' (MacCormick 1977, 189). Let us offer an initial definition.

Will Theory. For X to have a (claim-)right against Y that $Y \Phi$, X must have the power to alter Y 's duty to Φ .³⁶

Given that the idea behind the Will Theory is that rights endow their holder with normative control, we can see why, for X to have a *right* that $Y \Phi$, she must hold a power to alter Y 's duty to Φ —this grants her *normative control*. Suppose that you are under a duty of non-interference with respect to my pen. If I have the power to waive your duty, I have normative control over my pen: I am permitted to keep hold of it, give it to you, and so on. If it were a third-party who had the power to free you from your duty, on the face of things I would not have normative control over the pen. At any point, the third-party could free you from your duty. It is for this reason that Hillel Steiner sees rights as demarcating ‘spheres of practical choice within which the choices made by designated individuals [...] must not be subjected to interferences’ (Steiner 2000, 238). Inasmuch, on this picture, rights are ‘*normative allocations of freedom*’. In H. L. A. Hart’s terms, ‘[t]he individual who has the right is a small scale sovereign to whom the duty is owed’ (Hart 1982, 183).³⁷

In section 7, we consider a revision to the Will Theory on which it is not necessary that X has the power to *alter* Y 's duty, but that X has some weaker measure of control over Y 's duty. We begin with this stronger formulation since it most clearly endows X with control over Y 's duty.

The Will Theory offers an account of the grounds of rights. Talk of “the grounds of rights” might be ambiguous. First, one might think that what is being identified is the moral base of Y 's correlative duty. That is, the normative control afforded to X by her right explains *why* Y is under a correlative duty. Suppose that Dana has a right, against Eric, that Eric

³⁶ (Hart 1955; 1982; Wellman 1985; Sumner 1987; Steiner 1994; 2000). This definition suffices only for claim-rights for Y only has a correlative duty to Φ when X has a claim to Φ . As suggested by my exposition of Hohfeld in the previous chapter (section 4.1), I think that liberties, powers, and immunities can classify as rights. The Will Theory can be reformulated to accommodate this. See, e.g., (Cruft 2004, 367–68); Steiner makes room for immunities (Steiner 1994, 61). I also omit that Y 's duty to Φ must be owed to X for I see the Will Theory as offering an account of what it is to owe a duty to another.

³⁷ In its strongest formulation, the Will Theory would require that the right-holder has exclusive control over the duty, as it typically the case. If the right-holder does not have exclusive control, this complicates matters. We turn to similar issues in section 7.

not interfere with her movement. On this view, the reason why Eric is under a duty of non-interference is that this allows Dana control of her life. Alternatively, second, the Will Theory may recognise a plurality of moral bases of duties but suggest, when a duty protects X 's will, it becomes *directed* (and, correlatively, confers a right upon X). Suppose that Eric promises Dana that he will water her flowers. Eric's duty might *not* owe its existence to the fact that this serves Dana's autonomy (promise-keeping might be grounded in some other way). Rather, because Dana has control over Eric's duty, *this feature* makes Eric's duty directed.

There are two dialectical reasons often given in support of the Will Theory. First, the Will Theory solves the third-party beneficiary problem that putatively plagues the Interest Theory. The Interest Theory is introduced in more detail in the following chapter but, roughly, on one version of the Interest Theory (the Nonjustificatory version), for X to have a (claim-)right against Y that $Y \Phi$, her interest must be protected by Y 's duty (e.g., Kramer 2000).³⁸ While this is only a necessary condition for right-ascriptions, some people are sceptical that any further condition could explain why third-party beneficiaries are not counterintuitively owed the duties from which they benefit. Suppose Dana promises Erica that she will pay Fran £10. What might explain why Dana does not owe her duty to Fran, given that Fran's interests are served by the duty (Hart 1982; Steiner 2000; Sreenivasan 2005; *cf.*, Kramer 2010)? The Will Theory avoids this conclusion as Fran does *not* have the necessary control over Dana's duty.

Second, the Will Theory can deal with *referred* rights that are not justified by their holder's interest, such as a journalist's right not to disclose her sources. These rights are problematic for the Justificatory version of the Interest Theory, on which, for X to have a (claim-)right against Y that $Y \Phi$, X 's interests must be of sufficient weight to place Y under a duty to Φ (for example, Raz 1986, 166; *cf.* May 2012).³⁹

³⁸ We return to this in chapter 3, section 3.1.

³⁹ This was the version of the Interest Theory introduced in the previous chapter. We return to this choice point between these two versions of the Interest Theory in the following chapter, section 3.1.

A final piece of exposition is required for the Will Theory. Above, I said, let us proceed with the assumption that *X* needs to hold the power to alter *Y*'s duty for this endows her with normative control over her duty. However, there are several ways that *X* might be able to alter *Y*'s duty. These other powers become relevant below (sections 3 and 7).

Hart thought there were three disambiguations of *X*'s power (Hart 1982, 183–84). First, if the duty has not been infringed, *X* might hold the power to waive or enforce *Y*'s duty. Second, if it becomes clear that *Y* will infringe her duty or if *Y* does infringe her duty, *X* might hold the power to waive or enforce proceedings for the remedy of *Y*'s failure to respect her duty (for example, in the legal case, *X* may hold the power to sue for an injunction or compensation). Third, *X* might hold the power to waive or enforce those remedies.

Steiner separates each of Hart's powers into two powers (Steiner 2000, 240). It is useful to do so for reasons that become clear in the following section:

1. to waive compliance with the duty;
2. to leave the duty in existence;⁴⁰
3. to waive proceedings for the enforcement of the duty;⁴¹
4. to demand proceedings for the enforcement of the duty (i.e., sue for an injunction or compensation);
5. to waive enforcement of the duty;
6. to demand enforcement of the duty.

⁴⁰ This is not a Hohfeldian-power for it is not an ability to alter the normative relationship between two parties but the ability to leave things as they are. This will be made clear in the following section. I delay refining it until then for ease of exposition.

⁴¹ Matthew Kramer has suggested in private communication that, to waive proceedings for the enforcement of the duty in advance of a contravention of the duty is to waive the duty. This is incorrect—you may be under a duty not to assault me, and I say, “You are under a duty not to assault me, but I waive my power to enact proceedings for the enforcement of the duty.” My doing this does not waive the duty. You are still under a duty not to assault me.

Let us begin by assuming that, for X to have a right that $\mathcal{Y} \Phi$, all six powers must be held by X (compare section 7). With all of this in place, we can move on to assess the Will Theory.

3. A Terminological Refinement

Matthew Kramer has recently raised a novel objection against the Will Theory. Within any legal system, there are two default rules that might be operating concerning the post-violation stage of enforceability, powers 3-6.⁴²

- (1) A default rule under which the nonexercise of a power-of-enforcement in the aftermath of a violation of a duty will result in the noneffectuation of the duty.
- (2) A default rule under which the nonexercise of a power-of-waiver in the aftermath of a violation of a duty will result in the effectuation of the duty.

(Kramer 2013, 250; Kramer and Steiner 2007, 286–88).

The idea is that, for any duty, there must be some default rule in place determining what will happen if the duty is (likely to be) violated and the power-holder is silent.

Suppose that the second default rule is operative. If \mathcal{Y} is likely to or does violate her duty, X 's nonexercise of powers 3-6 result in the enforcement of that duty. For example, suppose that \mathcal{Y} does violate her duty and X wants to sue for compensation. Given the second default rule, X neither holds nor exercises any power of enforcement because the enforcement of \mathcal{Y} 's duty will proceed automatically—this is not a *power* of enforcement for ‘a putative legal power is no legal power [...] if it does not make any *difference* to anyone’s legal positions’ (Kramer 2013, 252; my emphasis).

The worry for the Will Theory is that, rather than X holding two powers,

5. to waive enforcement of the duty; and,

⁴² *Post*-violation because there is a default rule build into power 2. See note 40.

6. to demand enforcement of the duty,

X actually holds one power, 5. But, according to the standard definition of the Will Theory, right-holders need to hold the *pair(s)* of Hohfeldian powers.⁴³ So, the Will Theory cannot account for *any* rights, given the operativeness of either default rule.

However, even though X does not hold the relevant pair of Hohfeldian powers, the combination of X having the power of waiver and the operativeness of the second default rule means that X has *control* over \mathcal{Y} 's duty. Either, X waives \mathcal{Y} 's duty to pay compensation, or she does not waive \mathcal{Y} 's duty to pay compensation (in which case \mathcal{Y} is under that duty).⁴⁴ Although the Will Theory has traditionally been formulated to require that X hold the pair(s) of Hohfeldian powers over \mathcal{Y} 's duty, there is no reason why we should not refine the Will Theory to reflect that X has the necessary control over \mathcal{Y} 's duty to count as a right-holder in this case:

Will Theory 2. For X to have a (claim-)right against \mathcal{Y} that $\mathcal{Y} \Phi$, either:

- (i) if default rule 1 is operative, X must have the power to enforce \mathcal{Y} 's duty to Φ , or
- (ii) if default rule 2 is operative, X must have the power to waive \mathcal{Y} 's duty to Φ .

Kramer writes, if one were to ‘retreat from the notion that a right consists in three sets of paired powers over a correlative duty, they would be abandoning their rationale for the Will Theory’ (Kramer 2013, 257). This is incorrect. The refinement is motivated by, first, a recognition that some default rule must be operative with reference to the waiver or enforcement of duties and, second, that a right-holder must have the relevant kind of normative control over the correlative duty. Given that X has the power to waive \mathcal{Y} 's duty,

⁴³ This is not a strawman-characterisation of the Will Theory: ‘something is a right if it is [...] a claim [...] to which are attached powers of waiver *and* enforcement’ (emphasis mine) (Steiner 1994, 61).

⁴⁴ Things get more complicated if X is not *uniquely* empowered to waive the enforcement of \mathcal{Y} 's duty. In an appendix on file with author, I take up this question, as well as, first, if X does not have an immunity with regards to exercising her power and, second, if X does not hold the liberty to exercise her power.

she is free to reshape her normative space; given that the duty will be enforced if she does not act, she is free to have her normative space shaped as she would like.

4. Undergenerations of Rights

With Kramer's objection dealt with, let us turn to the standard objections to the Will Theory.

The Will Theory requires that a right-holder has control over the duty that correlates to her right. This means that 'potential rightholders [are] only those beings that have certain capacities: the capacities to exercise powers to alter the duties of others' (Wenar 2005, 239). This precludes agents with undeveloped, compromised, or damaged rational capacities (for example, very young children, the severely mentally disabled, and some of those suffering from Alzheimer's disease).⁴⁵ These agents cannot meaningfully waive or enforce the duties of the individuals against whom they hold their rights. So, they cannot hold powers. Accordingly, they cannot hold rights on the Will Theory (MacCormick 1982, 156–59; Kramer 2000, 69–70). This is an implausible implication of the Will Theory. Call this the Incapacity Undergeneration. It is both an objection to the Will Theory as an account of *moral* and *legal* rights.

The Will Theory also undergenerates right-ascriptions for *inalienable* rights. Perhaps my rights against servitude and against grievous bodily harm are of such a character that I cannot waive them. This means that I do not hold a power of waiver over them. Because the Will Theory requires that we hold powers over any claims in order for them to be rights, inalienable rights cannot be accounted for. MacCormick has nicely captured this problem. In relation to minor interferences, physical sports, or bona-fide surgical operations, *X* can waive *Y*'s duty not to harm or interfere with her. The Will Theory can account for these rights. Yet, in relation to serious harm or operations by unqualified surgeons, *X* cannot waive *Y*'s duty not to act in such ways. Accordingly, the Will Theory cannot account for these more serious rights: 'How odd that, as the [normative]

⁴⁵ It also precludes animals and, perhaps, the dead. Whether these are right-holders is more contentious.

protection is strengthened, the right disappears!’ (MacCormick 1977, 197–98). Call this the Inalienability Undergeneration. It is an objection to the Will Theory as an account of *moral* and *legal* rights.

Finally, the Will Theory undergenerates rights that are not inalienable in general but are not alienable by their holders. The usual example offered is some of our rights pertaining to the criminal law.⁴⁶ Crudely put, criminal law duties are not waivable by their correlative right-holders. My neighbour’s legal duty not to cut off my arm (correlating with my right that she not do so) is not waivable by me. But, at least on positivist views of the law, the state is able to waive this duty. The Will Theory implies that I do not actually hold this right against my neighbour. Call this the Inability Undergeneration.⁴⁷ (Since the state can waive the duty, the Will Theory says the state holds the right. We return to this in section 5.1.)

The Inability Undergeneration contains rights that are alienable by others aside from the putative right-holder, whereas the Inalienability Undergeneration contains rights that are inalienable in general. Drawing the precise boundary between these two classes of undergenerations is going to be hard. It depends on whether we think certain criminal law duties *could* be waived by legal officials, or whether they are inalienable in general. It is also unclear whether the Inability Undergeneration applies to *moral* rights pertaining to the criminal law. Perhaps the moral rights that individuals have pertaining to the criminal law are either inalienable (so captured by the Inalienability Undergeneration) or morally waivable by their holders, though, for practical reasons, not legally waivable. Alternatively, if the legal rights are justified, their holders do not have the moral power to waive the duties correlating to them. Regardless of which way we go on this question, there are some moral rights pertaining to the law that are not inalienable in general, but are not waivable by their holders. For example, in countries with minimum wage laws, employers

⁴⁶ “Some of” since it is likely that the rest will be inalienable in general.

⁴⁷ For those with less positivistic views of the nature of the law, judges have the legal power to incarcerate people for long periods of time (or even execute people), legal powers the individual likely will not have. Thanks to an anonymous reviewer for this.

are under duties to pay their employees certain amounts; these duties are not waivable by potential employees, though is waivable by the state.

Steiner thinks that the Incapacity Undergeneration is a *moral* objection to the Will Theory, but the Inalienability and Incapacity Undergenerations are *conceptual* objections because the Will Theory is ‘logically incapable of explaining how the undisputed existence of some perfectly enforceable duties [criminal law duties] implies rights at all’ (Steiner 2000, 248). It is unclear why this conceptual point is not true of the Incapacity Undergeneration. The objection is not “Wouldn’t it be bad if children didn’t hold rights, how immoral of the Will Theory!” but “We take children to have rights just as we take adults to have rights with respect to the criminal law—the Will Theory logically precludes this.” For this reason, I see no reason to distinguish the character of the Incapacity Undergeneration from the Inalienability and Inability Undergenerations.

Steiner also suggests that the three dimensions along which the Will Theory undergenerates right-ascriptions suffer from ‘severely disabling circularity inasmuch as [they] simply presuppose the truth of the [Interest Theory]’ (Steiner 1994, 66).⁴⁸ However, they do not presuppose the Interest Theory but suggest that the Will Theory is overly restrictive.

With our three classes of undergenerations in place, the following three sections put forward a range of seemingly plausible ways that a defender of the Will Theory can respond, while laying clear the implications of those positions. We see there is a general problem. Either, we can keep the Will Theory faithful to its account of control as the grounds of rights, in which case the theory undergenerates right-ascriptions, or we can make the theory extensionally accurate, but at the cost of obscuring the theory’s focus on control.

5. Being Revisionary About Rights

One might think that we need not refine the Will Theory in reply to our three classes of undergenerations, but simply draw out the Will Theory’s implications more carefully.

⁴⁸ Kramer makes similar remarks in reply to the third-party beneficiary problem (Kramer 2000, 67–68).

Subsection 5.1 examines how defenders of the Will Theory may respond in this way to the Incapacity and Inalienability Undergenerations. Subsection 5.2 examines how they may respond in this way to the Inalienability Undergeneration.

5.1 Being Revisionary About Right-Holders

The Incapacity and Inability Undergenerations occur because someone who is intuitively not the right-holder is in possession of the relevant powers of demand and waiver. Why not think that it is *these* individuals—the holders of these powers—that possess the rights? With reference to the Inability Undergeneration, Steiner writes ‘there’s simply no obstacle to extending the application of the Will Theory to criminal-law duties and holding that, whereas civil law confers Will Theory rights on private citizens, criminal law (now) rests them in state officials’ (Steiner 2000, 250).

It is more tempting to say that it is not the state officials, but society at large who is the right-holder, and that officials act as fiduciaries on society’s behalf. While comparatively more attractive, this view still has flaws. First, the objections I raise below apply equally. Second, the Will Theory cannot, without alteration, make sense of fiduciaries. We consider such an amendment below (Will Theory 3), so I set it aside for now.

There are obstacles to thinking that criminal law rights and the rights of agents without rational capacities are actually held by others. Both as a mark of first-order and theoretical intuitions, it is extremely counterintuitive. The right that correlates with your being under a duty not to break my leg seems to be *my* right, not some state officials’. I said that it is more tempting to think of the revisionary right-holder as being society. Perhaps, if you break my leg, you *wrong* society through breaking its laws in some sense, in addition to, or in place of, acting wrongly in an undirected sense. Nonetheless, you also wrong me in a way that is distinct from my being a member of society. Put differently, I have different grounds for complaint than society.

This is because being revisionary about right-holders in this way is inattentive to the directionality of duties. Take the Incapacity Undergeneration. Children, the mentally disabled, and individuals suffering from Alzheimer’s seem to have rights to the provision of certain goods and against certain interferences. Failure to respect duties pertaining to these entitlements seems to wrong *those* individuals. This feature marks that these duties

are directed to the individuals under consideration—as we saw in the introduction, one is wronged by a duty’s violation only if one is owed that duty; if one is not owed a duty, one is not wronged by its violation.

There are two ways a defender of revisionism about rights might reply. First, they might welcome the implication and argue that *it is* the holders of the relevant legal powers (*it is* the legal officials, the guardians of children) who are wronged when these rights are violated (Hart 1955, 181).⁴⁹ This is an implausible implication of being revisionary about rights. In Michael Thompson’s terms, we think of these individuals as ‘*not just raw materials for wrongdoing, but someone whom someone might ‘wrong’’*’ (Thompson 2004, 352).

Second, we might separate rights from wrongings. We might hold that some feature(s) grounds wronging and say, of that set of wronging, only those instances when an agent has normative control over the individual who wronged them are right-violations.⁵⁰ Above, I distinguished two senses of grounding—this reply is closer to the second sort, on which all manner of things ground duties; it is when an agent has normative control over another’s duty that it becomes directed. This reply goes further, saying that all manner of things ground *directed* duties, infringements of which constitute wronging, and there is a further subset, those directed duties that agents have normative control over, that correlate with rights. Now, we might be sceptical of such a complicated picture as it appears ad hoc. Relatedly, we need to give an account of what makes a directed duty, since the normative control that previously delineated directed from undirected duties now delineates rights from directed duties.

Further, there is a worry that the defender of the Will Theory has changed the subject in dissociating rights from directed duties. This implication can be seen by examining the following dilemma concerning the normative seriousness of right-violations. Suppose that two threateners are about to inflict two equal harms on a child and her guardian. A defender of the view outlined above says that both individuals are wronged by their

⁴⁹ Cf. note 58.

⁵⁰ This is not the way Steiner would go—he characterises rights as ‘essentially about who is owed what by whom’ (Steiner 2007, 459).

respective threateners. However, *only* the guardian is the holder of a right against the threateners' actions. She holds both a right against her threatener's action and a right against her child's threatener. We could say that right violations are normatively significant for the person who holds that right, in addition to that she has been wronged. Yet this implies, counterintuitively, that something more normatively significant happens to the guardian—like her child, she is wronged; unlike her child, she has *two* rights violated, whereas her child has no rights violated. This is odd. Instead, we could say that right violations are *not* normatively significant in addition to wronging. Yet this means *all* of the normative work is being done by directed duties; aside from being counterintuitive, we need not care about the Will Theory of *rights* anymore.

We have been proceeding on the assumption that a person is wronged only if they have a directed duty violated (often, correlating with a claim-right). We might abandon this assumption. This allows us to keep the Will Theory as an account of directed duties and rights. Nicolas Cornell thinks that we should separate rights from wrongs, where rights concern reasons for action *ex ante* and wronging concerns complaint *ex post* (Cornell 2015, 119).⁵¹ Cornell offers several cases in support of this verdict, including a third-party beneficiary case, a case in which harm results from an overheard lie, and a case in which a drunk driver kills someone's child. Cornell believes that the third-party beneficiary, the overhearer, and the child's parents are wronged, though they have no right violated. Because of this, he thinks: 'parties may sometimes be put in a special moral position to complain and seek justification *ex post*, not by conduct over which they could have asserted any rights claim of their own *ex ante*, but rather by conduct that was wrong for other reasons, like violating someone else's rights' (Cornell 2015, 113). He thinks the Will Theory gives a good account of rights, and the Interest Theory gives a good account of wronging.⁵² Regardless of whether Cornell is correct that the third-party beneficiary, overhearer, and parents are wronged, a child or a person without rational capacities seem to

⁵¹ Cornell does not discuss that Y might be under a directed duty to X to Φ , but X not have a claim-right against her that she Φ (notable in its omission, e.g. (Cornell 2015, 111)).

⁵² Cornell takes it that children have rights, so he requires some revision to the Will Theory.

be on the “rights” side of Cornell’s taxonomy. This is because what happens to the child or the person without rational capacities generates reasons for action *ex ante*.

We have considered whether one can respond to the Incapacity and Inability Undergenerations by being revisionary about who the right-holder is, attributing it to whoever is in possession of the relevant powers of demand and waiver. This line of thinking gets wrong who is wronged by a duty’s violation because it gets wrong the directionality of duties. We have also seen that separating rights from directed duties and separating rights & directed duties from wrongs in reply to this problem does not come without its own problems.

For the remainder of this chapter, let us see how we might respond on behalf of the Will Theory without recourse to separating rights from directed duties or wrongs. In Section 6, we see how we might revise the Will Theory to respond to the Incapacity and Inability Undergenerations. Before that, let us see how we might respond to the Inalienability Undergeneration without revising the Will Theory.

5.2 Normative Control

The Inalienability Undergeneration occurs because, intuitively, there are some rights that are not waivable, either by their holder or by anyone else. Above, I gave the example of rights against servitude and against grievous bodily harm. Now, recall from the introduction that I said many think that I can have a liberty, against you, not to give you my pen, but be under a duty to give you my pen nonetheless—it is just that *this* duty cannot be owed to you (for example, it may be owed to a third-party, to whom I have promised my pen). Another example of this phenomena is, even if I free you from your duty not to take my pen, you may still be under a duty not to take my pen, owed to someone else (for example, you may have promised a third-party that you will not take my pen). What is going on here is that I have removed the duty that you owed *to me* not to take my pen without removing *all* duties that you are under not to take my pen.

We might use this machinery to try and explain inalienable rights on the Will Theory.⁵³ Take the example of my legal right, against you, that you not inflict grievous bodily harm on me. Perhaps, were I to attempt to waive my right, I would free you from your duty, *owed to me*, not to inflict grievous bodily harm on me, but you remain under a legal duty not to do so nonetheless (for example, a legal duty owed to the state). Put differently, you are under (at least) two duties not to inflict grievous bodily harm on me. I have the power to waive one of those duties. Since I have the power to waive that duty, I have sufficient control over that specific duty to attribute a right to me on the Will Theory. Correlatively, this means you owe me a duty not to inflict grievous bodily harm on me. An important upshot of this is that, if I waive the duty, you would not wrong me by inflicting grievous bodily harm on me, though you may still act wrongly.⁵⁴

What to make of this way of responding to the Inalienability Undergeneration on behalf of the Will Theory? Let me draw out three implications. First, the type of normative control that grounds rights is somewhat weaker than the picture the Will Theory began with. Recall Hart's insight that, on the Will Theory, '[t]he individual who has the right is a small scale sovereign to whom the duty is owed'. Given what we have said in this subsection, rights are not grounded in the control one has to make others' actions permissible—the type of control that Hart's language evokes.⁵⁵ While this certainly does not count decisively against the Will Theory, it is an interesting implication of this line of thinking.

⁵³ Thanks to an anonymous reviewer for suggesting this way of responding to the Inalienability Undergeneration.

⁵⁴ On the importance of this kind of control, see Owens's discussion of normative interests (Owens 2012). Might we be able to tell a similar story for moral rights? Perhaps I am able to free you from your moral duty, owed to me, not to inflict grievous bodily harm on me, and yet you still be under a moral duty not to inflict such harm on me. For example, perhaps one remains under an undirected duty not to inflict grievous harm on me despite my valid consent. Tadros endorses a view of this sort, though the details are somewhat different (Tadros 2016a, 265–80, esp. 273–74). However, a defender of this view owes us an account of why, on the one hand, the undirected duty not to harm is present in cases of grievous bodily harm, despite one's valid consent waiving the directed duty, while the undirected duty is not present in cases of non-serious harm and, on the other hand, why this does not simply mean the directed duty not to commit grievous harm remains present. I look to develop this in future work.

⁵⁵ Can we appeal to the fact that one has control over whether others wrong one in performing certain actions inasmuch as one has control over whether to leave the duty, owed to one, in existence? As an

Now one might say, *other things being equal*, X , the right-holder, has control over whether T 's not Φ -ing is permissible. This allows defenders of the Will Theory to keep X 's making T 's not Φ -ing permissible central on the account.⁵⁶ Yet, we need to ask what is meant by other things being equal, here. And, what must be being held equal is *all* other duties that T is under—only in holding those other duties absent can we say that X has the power to make T 's not Φ -ing permissible. There is still a departure from Hart's picture of right-holders as small-scale sovereigns.

Second, this way of thinking might still strike one as extensionally inadequate. This way of thinking does allow that one can waive one's rights and others' directed duties not to perform certain actions, though the performance of those actions remain impermissible. Nonetheless, it implies that one can waive *all* of one's rights and correlative directed duties owed to one. But, it seems possible that a legal system or moral theory may imply that inalienability is possible. For example, that one *cannot* waive others' duties not to enslave one, seriously harm one, or even kill one. One could think, even to perform these actions in the knowledge that the right-holder consents would be to undermine the right-holder's moral status—as someone to whom these sorts of things cannot be done—and thereby constitute a wrong *to them*. Of course, this is only a sketch of such a moral view. But it does not seem at all contradictory to have inalienable moral rights. That a legal system may have such implications is even easier to imagine.

Third, and somewhat relatedly, regardless of what one actually thinks of the extensional adequacy of this way of thinking, it implies that it is still conceptually impossible for one to hold a right that one is incapable of waiving. Correlatively, it is conceptually impossible for one to owe a duty to someone else that they are incapable of waiving. And, ruling this out as a matter of conceptual fiat might strike one as a mark against the Will Theory. For example, the *volenti non fit injuria* principle states that, to one who consents, no wrong is

anonymous reviewer pointed out to me, this risks assuming in the analysis one of the features the account is trying to explain.

⁵⁶ Thanks to an anonymous reviewer for this suggestion.

done. If the Will Theory is correct, there is no argument concerning the volenti principle—it simply falls out of the nature of rights. But this seems like an open question.

We have considered how problematic the Inalienability Undergeneration is without recourse to revising the Will Theory. It does not come without its problems. Below, we turn to another way that a Will Theorist might respond to the Inalienability Undergeneration by revising the Will Theory (section 7).

6. What's in Being a Power-Holder?

We have seen that it will not do to be revisionary about rights. Perhaps:

Will Theory 3. For X to have a (claim-)right against \mathcal{Y} that $\mathcal{Y} \Phi$, either X or some other agent, \mathcal{Z} , must have the power to alter \mathcal{Y} 's duty to Φ on X 's behalf.

Will Theory 3 responds to the Incapacity and Inability Undergeneration. If there are inalienable rights, they cannot be rights on Will Theory 3.⁵⁷

One might wonder how principled of a refinement Will Theory 3 is. Hart came to endorse (something like) Will Theory 3.⁵⁸ By way of supporting the refinement, he suggests,

since (a) what such representatives can and cannot do by way of exercise of such power is determined by what those whom they represent could have done if *sui juris* [of age] and (b) when the latter become *sui juris* they can exercise these powers without any transfer or fresh assignment; the powers are regarded as belonging throughout to them and not to their representatives, though they are only exercisable by the latter during the period of disability. (Hart 1982, 184, fn. 86)

⁵⁷ Herein lies the difference between Will Theory 3 and Sreenivasan's Hybrid Theory—the Hybrid Theory allows that no-one hold the power over \mathcal{Y} 's correlative duty, provided this is in the right-holder's interest (Sreenivasan 2005; 2010).

⁵⁸ (Hart 1982, 184, fn. 86); *cf.* Hart's earlier view, note 49.

Hart's two points hold only for the Incapacity Undergeneration. So, one will either need to tell a different story to motivate Will Theory 3 with regards to the Inability Undergeneration, or else find some other way of responding to that class of undergenerations. Further, it is not obvious that Hart's first point, (a), holds for all rights. There are moral and legal powers that one can exercise when of age, though a fiduciary cannot exercise on one's behalf. For example, when of age, one can consent to sexual intercourse with others. But fiduciaries cannot exercise powers of sexual consent. So, it is not true in general that 'what such representatives can and cannot do by way of exercise of such power is determined by what those whom they represent could have done if *sui juris* [of age]'.

Hart's second point, (b), holds only for children who go on to develop sufficiently sophisticated rational capacities. Perhaps an analogous story could be told for those who lose their rational capacities, though the transfer of legal powers does not proceed by default as Hart emphasises with children. Hart's second point does not hold for those who will never develop sufficiently sophisticated rational capacities, such as those with severe mental disabilities. So, it is unclear how principled of an account Hart has actually offered. This points to a more general problem with Will Theory 3 that is best got at through considering two further problems with this refinement.

First, Will Theory 3 weakens the grounding relation supporting the Will Theory.⁵⁹ The idea behind the Will Theory is that rights protect agents' will. However, in what way does a fiduciary acting on the claim-holders' behalf manifest the protection of the claimholder's will? Perhaps, if we consider the rights of children, their fiduciaries are acting in ways that the child would will, were she able. Yet, I think we are conflating what we take to be in a child's will with what is in their interests in general. In this way, I am unclear of whether I can understand the child's will in the same way as I understand an autonomous agent's will—the autonomous agent's will is morally salient because it allows her to shape her own life; this is not true of the child. This line of argument gets even more obscure

⁵⁹ Put differently, despite Hart's attempts to the contrary, Will Theory 3 is ad hoc.

when considering the rights of those with severe mental disabilities and agents who have permanently lost their rational capacities.⁶⁰

Relatedly, second, we might ask how Will Theory 3 interacts with the directionality of wrongdoing associated with violating rights. On the Will Theory, if \mathcal{Y} fails to respect X 's right, she wrongs X *because* she has failed to respect X 's will. The protection of X 's will grounds her right—it explains why her right has normative importance. Once we weaken the Will Theory so that the power might be held by another agent, \mathcal{Z} , why does \mathcal{Y} 's failing to respect X 's right not wrong \mathcal{Z} instead of X ? When \mathcal{Y} fails to respect X 's claim that \mathcal{Z} holds a power over, \mathcal{Y} fails to respect \mathcal{Z} 's will; on the Will Theory, failing to respect someone's will was supposed to account for wrongdoing.

This leads us to the more general problem with Will Theory 3—we need an explanation of how the fiduciary relationship links up with the Will Theory's focus on normative control, as well as a way to determine when someone is holding a power as a fiduciary, rather than holding a power as a right-holder. Moving forward, those wanting to defend Will Theory 3 need to offer us such an account.

One way of cashing out this relationship between the right-holder and others holding powers on their behalf that is faithful to the Will Theory's focus on normative control as the grounds of rights is:

Will Theory 4. For X to have a (claim-)right against \mathcal{Y} that $\mathcal{Y} \Phi$, either (i) X must have the power to alter \mathcal{Y} 's duty to Φ or (ii) that duty must be aimed at satisfying the minimum conditions necessary for the exercise of normative agency for X .

Will Theory 4 appears to deal with most of the undergenerations. First, perhaps children are afforded rights *so that* they can become normative agents; second, perhaps inalienable

⁶⁰ Perhaps there are even some people who do not counterfactually have consciousness because their disability is essential to their identity.

rights are aimed at protecting normative agency; third, perhaps the rights we have pertaining to the criminal law protect us in our capacity as normative agents.

There are at least three reasons to be sceptical of Will Theory 4. First, Will Theory 4 will not generate rights for agents without autonomy in prospect, so still undergenerates right-ascriptions for those with permanently compromised or deteriorating rational capacities. While Will Theory 4 lessens the scope of the Incapacity Undergeneration, it is still deeply problematic.

Second, it is misplaced to say that, for example, children are afforded rights *only* insofar as those rights are aimed at satisfying the minimum conditions necessary for normative agency. (Note that I am not denying that these sorts of considerations ground *some* of the rights afforded to children.) Suppose that Dana is under a duty not to break Eric's child's arm. It is misplaced to say that Eric's child has this right only insofar as, were Dana to break her arm, this would inhibit her ability to arrive at the minimum conditions necessary for normative agency. (If Eric's child is young enough, will it inhibit her arriving at normative agency whatsoever?) The same seems true, to a great extent, of inalienable and criminal law rights.

Third, and this is somewhat of a dialectical point, this account now seems structurally identical to the Interest Theory. Without having offered the Interest Theory, I cannot fully develop this criticism. But, Will Theory 4 says that children are afforded rights so that they can become beings who can lead autonomous lives. Because it is not the *exercise* of autonomy that grounds rights (*vis-à-vis* clause (ii)) but autonomy in prospect, proponents of Will Theory 4 require that rights are aimed at something valuable (autonomy). This looks structurally identical with the Interest Theory, though a version of the Interest Theory on which autonomy-interests *exclusively* ground rights.

That Will Theory 4's clause (ii) is structurally identical to the Interest Theory (with interests restricted to autonomy) does not mean it is false, though it does have dialectical implications on those arguments that can be used against the Interest Theory. And, in any case, the two previous objections (undergenerations pertaining to those without autonomy in prospect and the strange grounding of lots of rights) heavily speak against the view.

7. Different Measures of Control

Here are two further ways we might revise the Will Theory.

Will Theory 5. For X to have a (claim-)right against Y that $Y \Phi$, X must have *some* power over Y 's duty to Φ .⁶¹

Will Theory 6. For X to have a (claim-)right against Y that $Y \Phi$, either (i) X must have the power to alter Y 's duty to Φ or (ii) be specially placed to demand the enforcement of the right.⁶²

There are some problems with Will Theory 5 and 6. First, Will Theory 5 cannot answer all of the Inalienability and Inability Undergenerations since those who are intuitively the right-holders will not *always* hold any of the six powers identified above. (While the person who is the object of a criminal law duty is well placed to see to it that the duty is enforced, she is not necessarily uniquely placed.) Will Theory 6 is in a better position on this score since one can argue that the right-holder *is* specially placed to demand the enforcement of the duty, even if powers of waiver are held by other parties or if the duty is inalienable.

Second, both Will Theory 5 and 6 cannot answer the Incapacity Undergeneration. This is because young children, for example, cannot exercise any of the six powers offered above, nor are they specially placed to demand the enforcement of their rights. So, it is not clear how far Will Theory 5 and 6 get us.

Third, Will Theory 5 and 6 seem at odds with the intuitive stringency of rights on the Will Theory. This is easiest to see on Will Theory 6. In discussion of the Inalienability Undergeneration, I noted MacCormick's observation that agents hold the power to waive others' duties not to inflict minor but not major suffering upon them. This means that any agent, X , satisfies conditions (i) and (ii) with respect to minor suffering and only (ii) with respect to major suffering. X 's right against major suffering is more stringent than her right

⁶¹ This view is relied upon in Steiner, especially his discussion of plea bargaining (Steiner 1994, 70; 2000, 251).

⁶² (Feinberg 1980; Skorupski 2010, 307–13; Darwall 2013).

against minor suffering. Rights are grounded in the normative control they endow their holder with. It is not clear how these two features link up on Will Theory 6. We can tell a similar story with Will Theory 5. Suppose that X has only the power to waive enforcement of the duty (powers 5 and 6), but not the power to waive the duty. If X 's having normative control over Y 's correlative duty grounds X 's right, how can the right that manifests less of this be more important?⁶³ This seems difficult to explain.

Fourth, there might be problems with Will Theory 5 and 6 depending upon how the normative control of the duty is distributed. Let us begin by seeing how this follows on Will Theory 5. Suppose that X holds only the powers to waive or demand enforcement of Y 's duty (powers 5 and 6), but Z holds the power to waive Y 's duty. The Will Theory says that rights are grounded in the normative control they endow their holder with. Yet, while X has *some* control over Y 's duty—she is in control of what happens as concerns enforcement if Y violates her duty—she only has this control conditional on Z 's not waiving Y 's duty. And, one might be sceptical of whether this level of control is sufficient for X to have a right over Y that $Y \not\Phi$. Even if one thinks this level of control is sufficient, it is important to appreciate the somewhat weakened notion of normative control that grounds rights given the move to Will Theory 5.

Alternatively, perhaps it is Z that has control over the duty, and so Z that holds the right. First, this means Will Theory 5 will not help with the Inability Undergeneration (since it is those who are not, intuitively, the right-holder that have the power to waive the duty). Second, it is not even obvious it gets us the result that Z holds the right. Suppose that Z leaves the duty in existence for she wants Y to Φ . Y violates her duty. Z has no control over what happens now— X may simply waive Y 's duty to pay compensation. Again, one might be sceptical that this level of control is sufficient to endow X with a right. (Steiner thinks the right shifts from Z to X , though does not comment on its intuitive strangeness (Steiner 2000, 247).)

Things are a little different for Will Theory 6. But, the general point holds when X is only specially placed to demand enforcement of the right, but someone else has the power to

⁶³ Sreenivasan makes a similar point (Sreenivasan 2005, 261).

waive the duty correlative to the right. For example, suppose X wants to demand enforcement of the right, yet Z waives the duty. More generally, we might be sceptical that X has sufficient normative control over the correlative duty in virtue of being specially placed to demand enforcement of the right if someone else has the power to waive the duty in the first place. At the least, again we should note the marked shift in the type of control over others' duties that grounds rights on the Will Theory.

At this stage, one might reply that all this fourth objection shows is that we need to be more careful in formulating the right in question. For example, when X has the power to enforce \mathcal{V} 's duty, and Z has the power to waive \mathcal{V} 's duty, why not say: X has a (claim-)right that $\mathcal{V} \Phi$ conditional on Z not waiving \mathcal{V} 's duty.⁶⁴

The first thing to note is this does not help us with the first three objections raised against Will Theory 5, and the second and third objections, which hold against Will Theory 6. And, in any case, this proposal is not going to come without its issues. This is because the conditional nature of these rights might strike one as extensionally odd. Let us take two concrete examples.

Suppose we want to use Will Theory 5 to deal with rights correlating with criminal law duties. Sometimes, victims of the violation of criminal law duties have some measure of control over the enforcement of the duty after its violation through judges taking into account their wishes in sentencing. In virtue of this control afforded to victims, let us say they are endowed a legal right correlative with the criminal law duty. Now, this does not mean they have a right against the harm. Rather, they have a *conditional* right against the harm, conditional on the duty's not being waived (and on their wishes being taken into account in sentencing—for if their wishes are not taken into account, they have no control over the duty). And, one might find it implausible that the right is conditional in this way.

Let us take another example. In section 4, I said that, in countries with minimum wage laws, employees have the right to a minimum wage. This is problematic for the Will Theory since the state has the power to waive employers' duty to pay a minimum wage. On

⁶⁴ Thanks to two anonymous reviewers for this suggestion.

Will Theory 5 and 6, employees have the right to a minimum wage if they hold, for example, the power to demand enforcement of the duty. The fourth objection I have raised to Will Theory 5 and 6 is that they only have the control endowed by this power if the state does not waive employers' duty to pay a minimum wage. And, this does not endow employees with the type of normative control necessary for rights on the Will Theory. The conditional strategy of replying to this objection would say that the right employees have is the right to a minimum wage conditional on the state not waiving the duty—this links the control employees do have with the content of their right. But, like with rights corresponding to criminal law duties, the conditional nature of this right might strike one as extensionally odd—I am inclined to think employees have the right to a minimum wage. The right may well be more secure if the state did not have the power to waive the employers' duty to pay a minimum wage, but this does not affect the content of the right.

Whatever one thinks of the extensional plausibility of these conditional rights, there is a more general problem that all of this discussion highlights: the more moves defenders of the Will Theory make in replying to our three classes of undergeneration (for example, by moving to either Will Theory 5 or 6, by drawing out that putative right-holders do have control conditional on others not waiving the correlative duties, and so on), the weaker the normative control the right-holder has over the duty. But, the insight of the Will Theory was supposed to be that rights endow us with normative control. Recall Steiner's remark that rights, on the Will Theory, demarcate 'spheres of practical choice within which the choices made by designated individuals [...] must not be subjected to interferences' (Steiner 2000, 238). One might be sceptical of how much of this insight of the Will Theory remains true of either Will Theory 5 and 6. Put differently, how connected to the idea of normative control grounding rights are Will Theory 5 and 6?

It is difficult, then, to link the idea of normative control with there being a distribution of the powers one holds over another's duty. It is even more difficult to see why, the more stringent a right appears, the less normative control one has over it. This is exemplified

with *inalienable* rights—we are only (elusively) specially placed to demand the enforcement of these right.⁶⁵

8. Conclusion

In section 4, I introduced the Incapacity, Inalienability, and Inability Undergenerations. Sections 5-7 put forward a range of seemingly plausible ways that a defender of the Will Theory may respond, while laying clear the implications of those positions. At the end of section 4, I suggested that there was going to be a general problem with these responses—either, we can keep the Will Theory faithful to its account of normative control as the grounds of rights, in which case the theory undergenerates right-ascriptions; or, we can make the theory extensionally accurate, but at the cost of obscuring the theory’s focus on control.⁶⁶

We have seen that, (5.1) if one wants to be revisionary given the undergenerations that plague the Will Theory, one faces trouble in attempting to account for the connection between rights, directed duties, and wrongings. That is, the counterintuitiveness of the undergenerations cannot be explained away. And, we have seen that (5.2) emphasising that one may be able to waive some of others’ duties not to Φ , without waiving all of others’ duties not to Φ , still leads to undergenerations of rights.

In section 6, we found that, on the one hand, while the idea of a fiduciary acting on a right-holder’s behalf may get a better extension than the standard formulation of the Will Theory, it obscures the connection between rights and wronging. On the other hand, we saw that though it may be faithful to the Will Theory’s account of the grounds of rights

⁶⁵ Rowan Cruft argues, further, that Will Theory 6 will be, in many circumstances, circular, for whether or not X is *specially placed* to demand the enforcement of another’s duties seems to depend upon whether we think the duty is owed to them in the first place (Cruft 2019, 32–34).

⁶⁶ A further point to note is that none of the revisions across sections 5 to 7 solved all three undergenerations; each would have to be taken with other revisions, thereby intensifying the weakened focus on normative control as the grounds of rights.

to see some rights as being grounded in securing the minimum conditions necessary for autonomy, this account is both stretched and itself extensionally inadequate.

Finally, in section 7, we saw that the Will Theory has trouble if revised to allow right-holders to have lessened powers of control over others' duties. While this revision may achieve a better extension than the standard formulation of the Will Theory, again it is difficult to link these revisions up with the Will Theory's focus on normative control as the grounds of rights.

Moving forward, defenders of the Will Theory need either to take a stand on this dilemma or else find some other way of dissolving the dilemma not discussed in this chapter. I am sceptical this dilemma can be dissolved. The central insight of the Will Theory is that rights endow their holders with control—but, some rights just are not about control. So, if rights are not necessarily grounded in control, what are they grounded in?

3. The Interest Theory

1. Introduction

Think of some of one's most important rights: one's rights against being killed, against being tortured, against being raped. Why might one think those rights are so especially important compared with one's other rights? One thought is that it would be awful for one were those things to happen. According to the Interest Theory of Rights, it is precisely *because* these things would be so awful for one that one has rights against the performance of those actions by others. Correlatively, that is why others owe one duties not to perform those actions: *because* they would be so awful for one.

This way of introducing the Interest Theory brings to light a problem with the Will Theory, a problem that was lurking in the background of the previous chapter. The Will Theory says rights are grounded in the control with which they endow their holders. However, this fails to capture why (at least some) rights are important. Some rights are important *because* of the harm that would befall one were they to be violated. This point is well put when Cécile Fabre writes that 'being tortured without one's consent is bad because of the immense physical suffering one incurs, and/or because of the resulting [...] impairment of one's [autonomy]' (Fabre 2000, 15).⁶⁷ The Will Theory cannot recognise this first feature—that badness of certain states of affairs—as playing any grounding role in rights, for rights are, on that theory, exclusively grounded in normative control. That is not a satisfactory account of the grounds of rights.

In this chapter, we see that the Interest Theory solves the problems raised against the Will Theory in the previous chapter (section 2). I also argue we have other reasons to endorse the Interest Theory. The Interest Theory is then introduced in sufficient detail to be our focus for the remainder of this thesis (section 3). Finally, we return to the problem for the Interest Theory outlined in chapter 1, the Problem of Harmless Wronging (section 4). I

⁶⁷ Similarly, Cruft, Liao, and Renzo ask: 'The fact that torture undermines one's agency by undermining one's capacity to decide how to act and to stick to the decision is certainly an important factor in justifying the existence of a human right not to be tortured but is it the only factor?' (Cruft, Liao, and Renzo 2015, 11). See also (Liao 2010; Tasioulas 2010, 663–66).

also suggest the problem is troublesome for other theories in addition to the Interest Theory.

2. Introducing the Interest Theory

Let us offer an initial formulation of the

Interest Theory (Canonical). For X to have a right against Y that $Y \Phi$, X 's well-being (her interests) must be of sufficient weight to place Y under a duty to Φ .⁶⁸

The Interest Theory offers a plausible account of the *grounds* of rights—rights are grounded in the wellbeing of their holders.⁶⁹ Further, the Interest Theory neatly explains the distinctive structural features of rights and their correlative directed duties. First, it gives a good account of why Y owes her duty to X —the duty owes its existence to features about X . Second, and relatedly, it offers a plausible account of why, if Y infringes her duty, she does not merely act wrongly but *wrongs* X — Y has failed to respond to morally salient duty-grounding features about X .

I unpack the Interest Theory in the following section. In the remainder of this section, let us see why else one might want to endorse the Interest Theory (in addition to the Interest

⁶⁸ As we see directly below in the text, the Interest Theory offers an account of *what it is* to owe a duty to another. Accordingly, I omit that Y 's duty must be owed to X as is usually contained in these definitions. A few more clarifications. First, this version of the Interest Theory more closely resembles that of (Raz 1986, 166) rather than that of (Kramer 2017) (see section 3.1). Second, the definition suffices only for claim-rights for Y is under a correlative duty to Φ only if X has a *claim-right* that she Φ (*cf.* section 3.4). Third, one might wonder how the Interest Theory can be called a *theory* given that it is merely a necessary but sufficient condition; this is addressed below (see section 3.3).

⁶⁹ This is ambiguous between saying that wellbeing makes the right exist and that wellbeing makes the duty correlate with a right. On the Interest Theory as I have defined it, grounds should be disambiguated in the first, stronger sense. This is because we are assuming a Justificatory version of the Interest Theory on which the grounds of the duty (the justification of the duty) determine its direction (see section 3.1). But, were we to go for a Nonjustificatory version, on which the direction of the duty is not tied to its justification, grounds could be disambiguated in the second, weaker sense. See the similar discussion in section 2 of chapter 2 on the ambiguity of grounds.

Theory's plausible grounding of rights and its nice explanation of some of the structural features of rights).

The Interest Theory can recognise rights that the Will Theory does not. The Will Theory says right-holders must have some measure of control over the duties that correlate with their rights. This precludes those with undeveloped, compromised, or damaged rational capacities (for example, very young children, the severely mentally disabled, and those suffering from the later stages of Alzheimer's disease). This was the Incapacity Undergeneration. The Interest Theory can recognise the rights of these people. All it requires is that they have wellbeing sufficiently important to place others under duties.

Importantly, the Interest Theory does not simply gain extensional accuracy where the Will Theory struggles; it also explains why these people are potential right-holders. It is the facts about their wellbeing that explains why others are under duties not to treat them in certain ways.

Since the Will Theory requires that right holders have some measure of control over the duties correlative with their rights, the Will Theory rules out inalienable rights. It also rules out holding rights where control over the correlative duties is held by someone else. These were the Inalienability and Inability Undergenerations respectively. The Interest Theory gets these cases correct. In either case, it does not matter that the right-holder does not have the power to control the duties correlative with those rights. All that matters is whether the holders' wellbeing is the grounds of those duties. What was so puzzling about the Will Theory's failure to recognise inalienable rights was that the best candidates for inalienable rights are very stringent. For example, the right not to be killed, rights against servitude, and so on. The Interest Theory makes sense of this—these are incredibly weighty rights because there is so much at stake for the holder's wellbeing. Below, we turn to whether the Interest Theory can *explain* why we do not hold powers of control over some of our rights (section 3.4.2).

This gives us another reason to endorse the Interest Theory: it offers a plausible account of the *stringency* of rights that is in keeping with the grounds of rights. (Differently, it gives

a plausible account of the severity of the wrong of a right's violation.)⁷⁰ My right that you not kill me is much more stringent than my right that you not hit me, and the wrong of killing me much graver than the wrong of hitting me, *because* being killed would be a lot worse for me than being hit.⁷¹

3. Unpacking the Interest Theory

Above, we canvassed some reasons to endorse the Interest Theory: it offers a plausible account of the grounds of rights; building on this, it offers a plausible account of the stringency of rights. It nicely explains some structural features of rights and directed duties. And, it deals well with the three classes of undergenerations that the Will Theory struggles with (as well as explaining why the Incapacity Undergeneration is not a problem for the Interest Theory).

⁷⁰ Some people have suggested to me that wrongs are all of the same weight (these people have exclusively been card-carrying Kantians). I find this incredibly implausible. Intuitively, the wrong one does to me when one breaks my leg is graver than the wrong one does to me when they break my finger. This is easily explained by saying that the respective right violated is more stringent in the leg breaking than the finger breaking.

⁷¹ This may be in conflict with what McMahan calls the Equal Wrongness Thesis: 'the wrongness of killing should reflect our commitment to the fundamental equality of persons [and] not vary with such factors as the degree of harm caused to the victim' (McMahan 2002, 235). Space does not permit me a full analysis of this (for other discussion, see (Hanser 2012)). But let me say a few things. Some of this conflict is just the Problem of Preemption: when death is not particularly bad for one, this may be because the badness is preempted by some other bad; to anticipate discussion of what is to come, whether this preempted harm will affect what rights we hold given the Safety Condition will depend on how easily it could have been that the harm was not preempted.

Even setting that aside, an analogous thesis, that the wrongness of inflicting some harm should not vary with such factors as the degree of harm, is implausible as suggested in the text: my right that you not kill me is much more stringent than my right that you not hit me. (The same holds comparing my right that you not kill me with Bloggs' right that you not hit them.) Further, the Equal Wrongness Thesis is, to my eye, intuitively implausible in intrapersonal comparisons: other things being equal, the wrongness of killing me as a child is graver than killing me as I approach the end of my life. And, even when it comes to interpersonal cases, I am tempted by the thought that, other things being equal, the wrong of killing a child *is* graver than the wrong of killing someone at the end of their life, and that it is harder to justify killing the child than the person at the end of their life. Finally, even if one is not tempted by this thought, this might be because the comparative utilities are irrelevant, in a way analogous to discussion of nonconsequentialist aggregation (Scanlon 1998, 229–41; Kamm 2007, 48–77).

In this section, I unpack the Interest Theory. My aim here is to provide a solid foundation for the Interest Theory. Because of this, at places we go into more detail than is strictly necessary for the arguments raised in the remainder of the thesis.

3.1 *Justificatory and Nonjustificatory*

This subsection compares two versions of the Interest Theory. The Problem of Harmless Wronging applies to either version, so this subsection is not strictly necessary for the argument of the thesis. The Interest Theory (Canonical) links the grounds of the correlative duty with the person to whom the duty is owed. However, there is another way defenders of the Interest Theory can go. Compare:

Interest Theory (Justificatory). For X to have a right against Y that $Y \Phi$, X 's wellbeing (her interests) must be of sufficient weight to place Y under a duty to Φ .

Interest Theory (Nonjustificatory). For X to have a right against Y that $Y \Phi$, X 's wellbeing (her interests) must be protected by Y 's duty to Φ .⁷²

Whereas the Justificatory version links the grounds of the duty to its direction (its justification, so to speak), the Nonjustificatory version does not go this way. On the Nonjustificatory version, we need not care what grounded the duty. We need care only about whether the duty *protects* the right-holder's wellbeing. In this way, we can see that the Justificatory version is in some sense more *demanding* than the Nonjustificatory version. Often, one's wellbeing will be served by some duty without being the grounds of the duty.

⁷² I have omitted "Canonical" to make clear that the choice-point is between Justificatory and Nonjustificatory versions of the Interest Theory. The Interest Theory's most prominent defenders are Raz and Kramer. Raz endorses a Justificatory version of the Interest Theory: "'X has a right" if and only if X can have rights, and, other things being equal, an aspect of X's well-being (his interests) is a sufficient reason for holding some other person(s) to be under a duty' (Raz 1986, 166). Kramer endorses a Nonjustificatory version of the Interest Theory: 'Necessary though insufficient for the holding of a legal right by X is that the duty correlative to the right, when actual, normatively protects some aspect of X 's situation that on balance is typically beneficial for a being like X (namely, a human individual or a collectivity or a non-human animal)' (Kramer 2017, 49). Kramer's view has subtly changed over the years, but that does not concern us here (Kramer 2000; 2001; 2010; 2013); for discussion of the changes, see (Bowen, n.d.).

There are at least two reasons to endorse the Justificatory version. First, the Nonjustificatory version struggles with so-called third-party beneficiaries, briefly introduced in the previous chapter (Hart 1982; Steiner 2000; Sreenivasan 2005).⁷³ Again, suppose that Dana promises Erica that she will pay Fran £10. Intuitively, Dana owes Erica a duty to pay Fran £10 and does not owe Fran a duty to pay her £10. While the definition of the Nonjustificatory version of the Interest Theory is only a *necessary* condition, what additional necessary conditions will explain why Dana does not owe her duty to Fran (and, correlatively, will explain why Fran does not hold a right against Dana to the promise) given that Fran's interests *are* served by the duty?

Now, intuitions go both ways on this example.⁷⁴ Even so, it is very hard to accept that Dana owes her duty to others who might benefit from the duty. For example, what of Fran's daughter, for whom Fran was going to buy a present with the £10?⁷⁵ Since Dana's promissory duty is not grounded in Fran's or her daughter's wellbeing, the Justificatory version easily explains why Dana's duty is not owed to them.

The most promising reply to the problem of third-party beneficiaries is that Fran's daughter does not *necessarily* benefit from the duty's being fulfilled (Kramer 2000, 79–84; 2010, 36–39). For example, Fran may decide not to buy her daughter a present. Whereas, so the thought goes, Fran's wellbeing is necessarily protected by the promise's fulfilment; so, the duty is owed only to those whose wellbeing is necessarily protected by the duty. Now, there may be some bugs with going this way. For example, we can likely bake into the case that the third party will necessarily be harmed when the right-holder is harmed.⁷⁶ But perhaps these bugs can be worked out. In any case, this appeal to necessity at least appears ad hoc. What principled reason do we have to move from, on the one hand, rights being grounded in their holder's wellbeing being protected by the correlative duty to, on

⁷³ See chapter 2, section 2.

⁷⁴ Compare (Kramer 2017, 74–76) with (Sreenivasan 2005; Wenar 2013, 213).

⁷⁵ Some think so-called harmless promise breaking is wrong because it erodes the practice of promise keeping, a practice that is beneficial to everyone (Rawls 1999, 301–7). If this is correct, we all benefit from a promise being upheld since we all benefit from the practice of promise keeping. This means promissory duties will be owed to everyone on the Nonjustificatory version, an unwelcome implication.

⁷⁶ For example, see the previous note.

the other hand, rights being grounded in their holder's wellbeing being *necessarily* protected by the correlative duty?

This is related to our second reason to endorse the Justificatory version. The Justificatory version's account of the grounds of rights is more plausible than the Nonjustificatory version's account of the grounds of rights. On the Justificatory version, one holds a right only if, and (partly) *because*, one's wellbeing is of sufficient weight to place the person against whom one holds the right under a duty. That is why the duty-bearer owes her duty to the right-holder—the duty exists because of them. On the Nonjustificatory version, one holds a right only if, and (partly) *because*, one's wellbeing is necessarily protected by the duty. But why would one's wellbeing being necessarily protected by some duty make it that the duty is owed to that individual?⁷⁷

Now, the Problem of Harmless Wronging applies to both versions of the Interest Theory. Since I think the Justificatory version's account of the grounds of rights is more plausible than the Nonjustificatory version's account, I will here adopt the Justificatory version.

3.2 *Sufficient Weight, Harm, and Wellbeing*

What does it mean, then, for someone's wellbeing to be of sufficient weight to place someone else under a duty? On the face of it, things would be easier if we had gone for the Nonjustificatory version of the Interest Theory: we have an intuitive idea of what it is for someone's wellbeing to be *protected* by a duty—we look to whether they would be harmed or benefited were the duty not to be respected. But, discovering whether someone's

⁷⁷ This is not to say that there is no reason to prefer the Nonjustificatory version. Consider again an example of what we can call a *referred* right, introduced in section 2 of chapter 2: a journalist's right not to disclose her sources (Raz 1986, 178–80; 1994). The journalist's right does not seem to be grounded in her own interests in not disclosing her sources, but in the public's interest in having a free press. But, on the Justificatory version of the Interest Theory, rights must be grounded in their holder's interests (in the wellbeing at stake for the holder). So, it looks like that version of the Interest Theory cannot explain how the journalist holds a right not to disclose her sources. However, since presumably the journalist's interests are protected by others' duties not to force her to disclose her sources (it makes carrying out her job easier for one thing), she can hold a right on the less demanding Nonjustificatory version. For potential replies on behalf of the Justificatory version, see (Raz 1994; May 2012). Addressing this problem for the Justificatory version would take us too far afield. But, for what it is worth, I am not even sure the costs of biting the bullet on referred rights is too great, especially when it comes to referred *moral* rights. Perhaps the duty is owed to the public at large.

wellbeing is sufficiently weighty to place others under duties is quite similar to discovering whether someone's wellbeing is protected by another's duty. It depends, in the first place, on the extent to which the potential right-holder would be harmed or benefited were the potential duty-bearer not to act as the would-be duty dictates.⁷⁸ For example, Fabre says: 'the degree to which an interest of X's is important enough to warrant the imposition of a duty on Y is a function of the degree to which X would be harmed if Y desisted from acting as required by the duty' (Fabre 2008, 227).⁷⁹

This means we need an account of harm and benefit. Let us begin by assuming the dominant

Counterfactual Account of Harm and Benefit. *Y* harms *X* iff (and because) *Y* makes *X* worse off than *X* would have been had *Y* not acted as she did. *Y* benefits *X* iff (and because) *Y* makes *X* better off than *X* would have been had *Y* not acted as she did.

The Counterfactual Account gets lots of cases right. For example, you harm me by cutting off my leg because you make me worse off than I would have been had you not cut my leg off. You benefit me by curing me of a cold because you make me better off than I would have been had you not cured me of the cold. What is more, the Counterfactual Account has great *explanatory* power: it is very plausible that you harm me by cutting my leg off *because* you make me worse off than I would have been had you not cut my leg off. The Counterfactual Account has this significant explanatory power because it tracks the difference harm makes to the harmed person—inasmuch, it is a *difference-making* account

⁷⁸ There is a distinction between benefiting from a duty's existence and benefiting from the performance of the action required by that duty. For example, suppose that Threatener will punch Victim regardless of whether she is under a duty not to. In this case, Victim would not be benefited by the duty's existence, since its existence makes no difference to whether Victim is punched either way; but, Victim would be benefited by the performance of the action required by the duty (i.e., not being punched). Cruft thinks the Interest Theory is formulated in the first way, and that this is problematic for the Interest Theory on these grounds (Cruft 2010). But we should formulate the Interest Theory in the second way, so there is no problem. This comes up again in chapter 8, section 4.2.

⁷⁹ See also (Scanlon 1984, 146; Raz 1986, 165–92; Cruft 2010, 441–45; Tasioulas 2015, 50–56).

(Bradley 2009, 50–51). We see some competing accounts of harm and benefit in the following chapter (sections 3-4).

It is worth clarifying two features of the Counterfactual Account of Harm. First, it is restricted to one person harming another, since we are mostly concerned with one person harming another person. More generally, we can say, an event, e , harms someone, X , iff X is worse off than she would have been had e not occurred. On this more general understanding, a tree falling from the wind can harm X just as much as a person's cutting down a tree can harm X . We can also read X broadly enough to include anything we think capable of being harmed. For example, Mouse (my dog) can be harmed by an event if that event leaves her worse off than she would have been had the event not occurred.

Second, the Counterfactual Account it is morally neutral. This is a very plausible feature of an account of harm and benefit but strikes some people as odd.⁸⁰ Suppose we were to go for a morally loaded Counterfactual Account: Y harms X iff Y *unjustifiably* makes X worse off than X would have been had Y not acted as she did. Now suppose Threatener comes at Victim with a knife, culpably looking to kill them. The only way Victim can defend herself is by shooting Threatener in the leg. Victim acts justifiably in defending herself. This implies that, on the moralised Counterfactual Account, Victim does not harm Threatener. But that is very implausible. She harms Threatener. It is merely that the harm is justified.

So, we discover whether one's wellbeing is sufficiently weighty to place others under duties by looking, in the first place, to how one would have fared were the duty not to be respected. Another thing we look to is the cost to the potential duty-bearer of acting as the would-be duty dictates. This plausibly differentiates rescues that one has rights to the performance of and those that one does not have rights to the performance of. For example, if the only way that you can save a drowning child is by subjecting yourself to some sufficiently large harm, intuitively, the child has no right to be rescued. But, if the costs are

⁸⁰ In conversation, Antony Duff objected on these grounds. Similarly, Feinberg thinks Y harms X only if Y violates X 's rights (Feinberg 1984, 36); in addition to the reasons raised in the text to avoid moralised accounts, Feinberg's view would leave us in an awful circle since we are grounding rights in harm. For more on why not to go moralised, see (Tadros 2016a, 181–82).

trivial, intuitively, the child does have a right to be rescued—it is not only wrong for you not to save the child, but you would wrong the child.⁸¹

There might be other things we look to in determining whether potential right-holders' wellbeing is sufficiently weighty to place others under duties. Some of these things I consider in the following subsection as separate considerations on rights; for example, whether the harm is intended or merely foreseen, whether it is a doing or an allowing, and so on. As I say below, we could build these other considerations into the theory here as partly constitutive of what it is for one's wellbeing to be sufficiently weighty to place others under duties.

We are building up to an account of what determines whether one's wellbeing is sufficiently weighty to place others under duties. A complete analysis would take a stand on what wellbeing consists of. We can distinguish three types of theory of wellbeing (Parfit 1984, 493–502). On *Hedonistic Theories*, wellbeing consists only of pleasure. On *Desire Theories*, wellbeing consists only of the satisfaction of one's desires. On *Objective List Theories*, wellbeing consists only of the advancement of certain aspects of our lives, regardless of our attitudes toward those aspects. Finally, on *Hybrid Theories*, wellbeing consists of some combination of these accounts (it could be disjunctive, for example, “pleasure or desire satisfaction”, or conjunctive, for example, “desire satisfaction one takes pleasure in”).⁸² We begin by being agnostic on this.

⁸¹ This discussion is in keeping with the discussion of liberty-rights on the Interest Theory below, section 3.4. For classic discussion of cost sensitive duties to rescue, see (Kagan 1989; Scheffler 1994). For explicit discussion of cost sensitive *directed* duties to rescue, correlative to rights, see (Quong 2015; Frowe 2019; n.d.)

⁸² Even within each type of theory, there is debate as to the details. For example, Hedonists tend to disagree as to what pleasure amounts to. *Internalists* about pleasure say that there is some quality common to all pleasurable experiences (Crisp 2006), whereas *Externalists* say there is some external feature common to all pleasurable experiences, for example, that they are desired (Sidgwick 1981, 127; Sumner 1996, 87–91; Parfit 1984, 493). Hedonists also disagree about whether all pleasures are commensurable (Mill 1969; Crisp 2006). Desire Theorists tend to disagree over which desires count for one's wellbeing. We turn to two choice-points in chapter 6, section 2. Objective List Theorists disagree about what items get to go on the list, as well as how the items on the list are linked. *Monists* think that one sort of thing unifies the list; for example, *Perfectionists* think that what is on the list is determined by the kind of thing that we are (Hurka 1993), whereas *Pluralists* think there is no such story to be told (Rice 2013). And Hybrid Theorists also disagree about which theories to combine, whether to make them conjunctive or disjunctive, and so on (Kagan 2009). The keen reader might also notice that I have only defined what is good for one, with no mention of what is bad for

3.3 Other Considerations

The Interest Theory (Canonical) is a necessary but insufficient condition on rights. In this subsection, I list some other potential considerations that, together, are jointly sufficient for rights on the Interest Theory. Some of these conditions are explicit necessary but insufficient conditions. Some modify the other necessary but insufficient conditions. I also explain how we might see these considerations as themselves partly constitutive of whether others' wellbeing is sufficiently weighty to place one under duties.

First, we have not said anything of the relation between the potential duty-bearers not acting as the duty dictates and the harm (or failure to benefit) to the potential right-holder. There are a few ways Interest Theorist could go. They could say that the potential duty-bearer needs to *cause* the harm to the right-holder were she not to act as the potential duty dictates. Suppose Carla will hit Ann if, and only if, Ben Φ s (for example, Carla will hit Ann unless Ben tells Ann he will never see her again). Since Ben does not *cause* the harm to Ann were he not to Φ (were he not to tell Ann he will never see her again), the causal reading of the relationship between the duty and harm (or failure to benefit) is not satisfied. So, Ann would not have a right against Ben that he not Φ . Alternatively, the Interest Theorist could say that the potential duty-bearer would need to “bring about” or “make it the case” that the right-holder was harmed were she not to act as the potential duty dictates.⁸³ Though Ben may not cause the harm to Ann, he may well bring it about or make it the case. And there may be other ways still that the Interest Theorist could go.

There are other things related to the relationship between the harm and the action required by the duty that might be relevant to the duty. Consider the following putative asymmetries:

one. This is because some theorists think wellbeing (what is *good for* one) is different from illbeing (what is *bad for* one) (Kagan 2014).

⁸³ One place to look for locutions like this would be the literature on the Doctrine of Doing/Allowing. See the note below.

Asymmetry Between Harm and Failure to Benefit. Other things being equal, it is harder to justify Y 's harming X than it is to justify Y 's failing to benefit X .⁸⁴

Doctrine of Foreseeability. Other things being equal, it is harder to justify Y 's (reasonably) foreseeing that she will harm X than it is to justify Y 's not (reasonably) foreseeing that she will harm X .⁸⁵

Doctrine of Double Effect (DDE). Other things being equal, it is harder to justify Y intending to harm X than it is to justify Y merely foreseeing that her action will harm X .⁸⁶

Doctrine of Doing/Allowing (DDA). Other things being equal, it is harder to justify Y doing harm to X than it is to justify Y allowing harm to X .⁸⁷

Interest Theorists may make any combination of these four conditions partly constitutive of rights. For example, in making DDE internal to rights we would say that, other things being equal, if Y intends the harm that would befall X were Y to violate the potential duty, it is more likely to warrant a right against that conduct than if that harm were merely foreseen. This means that a harm to X caused by Y might be sufficient to warrant a right if that harm is intended by Y , although it might be insufficient if it was an unintended but foreseen consequence of her action. The same story could be told for our other three putative asymmetries.

There may be other conditions relevant for right-ascriptions. For example, we might look to the distribution of harm among relevant groups of people and draw upon *Prioritarian* or *Egalitarian* considerations. We might look to the nature of the aspect of wellbeing at stake and draw *Sufficientarian* considerations (Fabre 2012, 17–26). We might not be

⁸⁴ (Shiffrin 1999; 2012; Gardner 2017).

⁸⁵ See (Thomson 1990, 229–34) for discussion of a similar principle.

⁸⁶ (Foot 1967; Quinn 1989a; Bennett 1995; McMahan 1994; Kamm 2007; Nelkin and Rickless 2014; Tadros 2015).

⁸⁷ (Foot 2002; Quinn 1989b; Bennett 1995; Woollard 2015).

Sufficientarians, but nonetheless think only some aspects of one’s wellbeing are deserving of right-ascriptions (Tasioulas 2015, 56–63).

Now, we could see all these considerations as themselves partly constitutive of whether others’ wellbeing is sufficiently weighty to place one under duties. If we went this way, we could read Interest Theory (Canonical) as a necessary and sufficient condition. This would be a good thing. Yet, it would leave “sufficient weight” doing a lot of work in the background. Since our focus is on harmless wronging—cases in which people are not harmed by their right’s violation, so it is the canonical necessary condition of the Interest Theory that is not satisfied—I think it is clearer to keep these conditions separate.⁸⁸

3.4 *Rights Beyond Claims*

To round off our unpacking of the Interest Theory, let us return to the other Hohfeldian incidents. So far, we have spoken only of claim-rights. This is because our definitions have referred to the correlative party’s duty to Φ , and there is only such a duty when the right-holder holds a *claim*-right that $\mathcal{Y} \Phi$. However, as my exposition of Hohfeld suggested, intuitively other Hohfeldian incidents are appropriately labeled “rights”. In the following two subsections, I suggest we can extend the Interest Theory to cover liberties, powers, and immunities. As with subsection 3.1, the following discussion does not really affect the arguments to be extended in the remainder of the thesis. However, it does demonstrate the resources available to a defender of the Interest Theory.

3.4.1 *Liberties*

First, consider liberties. Recall,

X has a claim that $\mathcal{Y} \Phi$, against \mathcal{Y} , iff \mathcal{Y} is under a duty to Φ , owed to X ,

and,

⁸⁸ This relates to whether we see the Safety Condition as a revision to the canonical statement of the Interest Theory or as a revision to how we determine whether one’s wellbeing is sufficiently weighty to place others under duties. See chapter 5, section 2.3.

Y has a liberty not to Φ , against X , iff Y is not under a duty to Φ , owed to X .⁸⁹

X 's claim against Y that $Y \Phi$ just is the absence of Y 's liberty not to Φ against X . The same is true in the opposite direction. Because of this, we might endorse the

Interest Theory (Liberty). For Y to have a liberty-right against X not to Φ , Y 's wellbeing (her interest) must be of sufficient weight not to place her under a duty to Φ .

That condition is in keeping with the Interest Theory's account of the grounds of rights—Ann has a liberty-right not to have sex with people at her discretion *because* her being under those duties not at her discretion would be awful for Ann.

Notice, whereas all claims are claim-*rights* (notwithstanding our discussion of Weak Correlativity in chapter 1, section 4.2), not all liberties are liberty-*rights*. Take Ann's absence of a duty to assault Ben. Since Ann does not owe Ben a duty to assault him, she has a liberty against him not to assault him. Yet, it might be a stretch to say she has a *right* against Ben not to assault him. The explanation of this is that Ann's wellbeing is not doing any work in explaining why she is not under a duty to assault Ben; it is Ben's wellbeing that is doing that work. So, Interest Theory (Liberty) is not satisfied. This is different from Ann's liberty not to have sex with Ben, for example. There, her wellbeing is grounding the absence of her duty.

The question arises, however: when is someone's wellbeing the grounds of the absence of a duty? I leave this question hanging for two reasons. First, it is not strictly relevant to the argument of this thesis and this chapter is already getting long. Second, suppose there was no good way to make this distinction, and so we were forced to conclude that all liberties were liberty-*rights*. I am in two minds as to how worrying this would be. While it does seem odd to say that Ann has a liberty-*right* not to assault Ben, I wonder if the strangeness comes from the rare need we have to refer to this right. Focus on Ann's *legal* liberty not to

⁸⁹ Chapter 1, section 4.1. Here, I have switched " X " to " Y " and " Φ " to "not Φ " to help emphasise how claims and liberties are deontic opposites.

assault Ben. Suppose the police began hounding Ann, trying to make her assault Ben. Perhaps she *would* say, “Look, leave me alone—I have the legal right not to assault Ben!” So, even if we cannot offer a comprehensive account of when someone’s wellbeing is and is not the grounds of the absence of a duty, we can still extend the Interest Theory to cover liberty-rights.

3.4.2 Powers and Immunities

Recall, the Interest Theory can accommodate inalienable rights. Suppose the right to life is inalienable (unless one has sufficiently good reason to end one’s life). The Interest Theory can say that others are under duties not to take one’s life because one’s life is sufficiently important for one. Yet, rather than merely accommodating inalienable rights, the Interest Theory may offer a ready *explanation* of why such a right is inalienable: ‘the disability is to [the right-holder’s] own advantage in preventing him ever from bartering away his freedom, whatever the temptation’ (MacCormick 1977, 205). This suggests we should endorse the

Interest Theory (Power). For X to have a power-right against Y , X ’s wellbeing must be of sufficient weight to make Y liable to X altering the normative relationship between Y and herself (or between Y and some other party).

An example might help. Compare the world in which Ann is able to make promises to the world in which she is unable to make promises. In the world in which Ann is unable to make promises, she cannot benefit from the practice of promise keeping. Further, she misses out on lots of deep normative phenomena. Often, making promises strengthens friendships. Ann has a lot of wellbeing at stake in being able to change the normative landscape between her and others through promising. These aspects of her wellbeing *ground* those normative powers.⁹⁰

⁹⁰ For further discussion, see (Owens 2012).

Not only does the Interest Theory (Power) explain why we hold inalienable rights, it nicely explains why lots of our rights are often alienable. To see this, consider the following objection from N. E. Simmonds. Simmonds complains that the Interest Theory is

compatible with a state of affairs where all powers of enforcement and waiver are monopolized by the state [...] There is surely a good deal of force in the Will Theory's claim that, in such circumstances, citizens would have no rights at all, regardless of how effectively their interests might be catered for in the state's policies and enactments. (Simmonds 2000, 225)

Simmonds is talking about legal rights here, but the example is nonetheless illustrative. He is correct about a few things. First, the citizens of this state would hold legal claim-rights given the Interest Theory. Second, this regime is problematic: citizens cannot shape the normative landscape in ways that would allow them to lead autonomous lives. However, if the Interest Theory (Powers) is correct, we can *explain* what is problematic about this regime: citizens' wellbeing is sufficiently weighty to warrant their being empowered to shape the normative landscape in ways that would allow them to lead autonomous lives. This importance is not being reflected in the legal powers they hold.

Just as claims are the flipside of liberties, powers are the flip sides of immunities:

X holds a *power* over Y iff Y is liable to X altering the normative relationship between Y and herself,

and

Y holds an *immunity* over X iff Y is *not* liable to X changing their normative relationship.⁹¹

Ann's not having the power to impose duties on Ben to worship in certain ways just is Ben's immunity, against Ann, that she not impose duties on him to worship in certain ways. This suggests we should endorse the

⁹¹ Again, there is some slight tinkering with these definitions from chapter 1, section 4.1, to draw out that powers and immunities are deontic opposites.

Interest Theory (Immunity). For X to have an immunity over Y , X 's wellbeing must be sufficiently important not to make her liable to Y 's changing their normative relationship.

Now, some people may find the Interest Theory (Power) and (Immunity) implausible.⁹² I say nothing else to defend it here, and the focus of this thesis is going to continue to be only on claim-rights. This subsection has widened the scope of the theory just to show it has the resources to do so.

4. Returning to the Problem of Harmless Wronging

I have argued that the Interest Theory has a lot going for it. However, there is a problem for the Interest Theory—the Problem of Harmless Wronging. I introduced the problem in chapter 1. However, we are now in a better position to fully appreciate the problem. In subsection 3.2 of this chapter we saw that whether one's wellbeing is of sufficient weight to place others under duties depends, in the first place, on the extent to which one would be harmed or fail to be benefited by the potential duty-bearer not acting as the would-be duty dictates. But, in cases of harmless wrongdoing, one is not harmed by the would-be duty's violation.

Take our cases of harmless wrongdoing, *Plane Crash (Preemption)* and *Roulette (Pure Risk)*. Since neither Passenger nor Target is harmed by the violation of the would-be duties, neither party's wellbeing is of sufficient weight to place their respective would-be duty-bearers under their duties. This means that the necessary condition set for right holding on the Interest Theory (Canonical) is not satisfied, so they hold no right.

I said that the Problem of Preemption and the Problem of Pure Risk are more specific versions of the Problem of Harmless Wronging for the Interest Theory. We take up the Problem of Preemption in the following chapter. However, before beginning I want to

⁹² Examples of harmless promise-rights are very easy to come up with (though *cf.* the introduction of the Pro Tanto and Normality Thesis in the following chapter, section 2).

note how harmless wrongdoing is problematic for other theories of rights in addition to the Interest Theory.⁹³

Leif Wenar’s Kind-Desire Theory says that ‘claim-rights correspond to [...] duties that the members of the relevant kind want to be fulfilled’ in their capacity as members of that kind (Wenar 2013, 218–19).⁹⁴ Wenar offers the following example of a kind-based desire:

When traveling through King’s Cross station in London, you may hear a recorded announcement saying that members of the station’s staff will be happy to help passengers. Even if every individual who works at King’s Cross is entirely misanthropic, we still say that all staff members, qua staff members, are happy to help passengers. (Wenar 2013, 215)

By grounding rights in kind-based desires, which are independent of right-holders’ well-being, Wenar hopes to avoid problems caused by cases of harmless wrongdoing.⁹⁵ Yet, while “what is in one’s kind-based desire in one’s capacity as a member of that kind” and “what is good for one” may come apart when thinking about roles one finds oneself in, it is difficult to see how one’s rights against *harm* will not end up being cashed out in terms of what is good for one—one’s kind-based desires will just be what is good for one.⁹⁶ Because of this, the Kind-Desire Theory will struggle in a similar way to the Interest Theory when it comes to the cases of harmless wrongdoing I discuss.⁹⁷

⁹³ This is in addition to my suggestion in chapter 1 (section 5) that cases of preempted harm and pure risk concern rights against harm; so, we should want a harm-based explanation to be available. Even if one denies the Interest Theory, one should want harm to play some explanation in why we have rights against *these* harmless actions.

⁹⁴ I have avoided Wenar’s formalisation to avoid unnecessary complications.

⁹⁵ Among others, cases of harmless promise-breaking, paternalism, and harmless trespass (chapter 1, section 5). In note 4, I said speaking of the grounds of rights is ambiguous: we could be after the thing that makes the right exist or the thing that makes it a right. The Kind-Desire Theory disambiguates grounds in the second way (along with the Nonjustificatory version of the Interest Theory and Sreenivasan’s Hybrid Theory introduced directly below). Cruft’s view (introduced below) should be disambiguated in the first, stronger way since it is based on the Justificatory version of the Interest Theory.

⁹⁶ Cruft makes a similar observation (Cruft 2017). Cruft now worries the Kind-Desire Theory fails as a sufficient condition (Cruft 2019, 20–31).

⁹⁷ Another popular reductive analysis of rights is Sreenivasan’s *Hybrid Theory*. On that theory, ‘Y has a claim-right against X that X ϕ just in case: Y’s measure (and, if Y has a surrogate Z, Z’s measure) of control over a duty of X’s to ϕ matches (by design) the measure of control that advances Y’s interests on balance’ (Sreenivasan 2005; 2010). Take two problematic examples from our discussion of the Will Theory. Control over a

Rowan Cruft has recently offered a non-reductive account of directed duties and their correlative rights that we may think fares better with cases of harmless wrongdoing. However, Cruft does think the Interest Theory (Canonical) offers a sufficient condition for natural directed duties in the sense of ‘duties owed to someone whether or not their existence or direction is recognized, and independently of their creation in law or convention’ (Cruft 2019, 89). If we add to this sufficient condition that the duty is ‘everyone’s business’, in the sense of ‘rights that anyone anywhere can, *ceteris paribus*, permissibly demand on behalf of the right-holder’, then we have two necessary and jointly sufficient conditions on *human* rights (Cruft 2019, 17). Cases of harmless wrongdoing pose problems for Cruft’s sufficient condition, then, and his account of human rights.⁹⁸

Finally, even those who offer seemingly fully non-reductive analyses of rights struggle with cases of harmless wrongdoing. For example, many think, other things being equal, the stringency of a right corresponds to the harm that would befall its holder were that right to be violated (Thomson 1990, 149–75; Kamm 2007, 249–75).⁹⁹ Cases of harmless wrongdoing wreak havoc with the intuitively plausible stringencies of rights. For example, what is the stringency of Passenger’s and Target’s rights against Attendant’s and Shooter’s behaviour?

Though none of these theories are committed to the necessity claim of the canonical statement of the Interest Theory, there is a reason they all face problems when it comes to

young child’s right is vested in a fiduciary because that advances the child’s interests. Similarly, ordinarily one’s lack of control over duties correlative to one’s inalienable rights (if one has them) advances one’s interests. Yet, the lack of control one has over duties that one is a third-party beneficiary to the performance of does not best advance one’s interests. However, since Srinivasan’s theory is ultimately grounded in well-being, again it has trouble with cases of harmless wrongdoing. For example, since Shooter does not harm Target in *Roulette*, what level of control over the duty is in her interests on balance?

⁹⁸ Elsewhere, I argue Cruft ought to see the Interest Theory (Canonical) as *grounding* his non-reductive analysis of rights (Bowen Forthcoming).

⁹⁹ Their views are hard to pin down. Here goes. Thomson thinks *X*’s claim against *T* is equivalent to *T*’s behaviour being constrained in a certain way, principle of which is that the constraint is (moral) remainder inducing when permissibly not acted in accordance with (Thomson 1990, 200–202). Kamm thinks rights reflect our inviolability as moral persons: ‘Inviolability is a reflection of the worth of the person. On this account, it is impermissible for me to harm the person in order to save [the many], because doing so is inconsistent with his having this status’ (Kamm 2007, 449).

harmless wronging.¹⁰⁰ This is the plausible thought that we began with: that part of the grounds of our rights against certain forms of harm is that it would be awful for one were those things to happen to one. Even though the focus of this thesis is offering a solution to the Problem of Harmless Wronging *for the Interest Theory*, moving to a focus on modality when it comes to rights against harm is of use for others in addition to defenders of the Interest Theory.

¹⁰⁰ Well, Cruft's account of human rights accepts the necessity claim.

Part II: Harm

4. Preemption

1. Introduction

In this part of the thesis we are focusing on preempted harm. Here is our example of harm that is preempted:

Plane Crash. Passenger is about to board a plane. Attendant takes a disliking to Passenger, so denies her admittance onto the plane. On departure, the plane crashes and everybody on board dies.

Passenger is better off in the world in which her putative rights are violated than in the world in which they are respected. Whether one's wellbeing is of sufficient weight to place others under a duty depends, in the first place, on how one would fare were the would-be duty not to be respected. Given this, Passenger's wellbeing is not of sufficient weight to place Attendant under a duty not to deny her admittance onto the plane. And because of this, on the canonical statement of the Interest Theory, Passenger does not have a right against Attendant that Attendant not deny her admittance onto the plane. This is an example of the Problem of Preemption.

Now, there are at least two general ways that we might solve the Problem of Preemption. First, we might say that people are clearly harmed when harm is preempted. So, any view of harm that says otherwise (for example, the Counterfactual Account) is mistaken. All we need to do is replace our account of harm. Once we find an account that arrives at the putatively correct verdict that agents are harmed when harm is preempted, the canonical statement of the Interest Theory can arrive at the correct verdict that agents have rights against preempted harm.¹⁰¹ (In effect, this strategy suggests that *Plane Crash*, and examples of preempted harm more generally, are not cases of *harmless* wrongdoing.)

Second, we might think that we need not even get involved with whether agents are or are not harmed by preempted harm. This is because we might think that the way I have set up the Interest Theory is *too* demanding in requiring that agent's wellbeing need *actually*

¹⁰¹ This strategy only offers a solution to the Problem of Preemption.

ground the potential duty. Rather, we might revise the Interest Theory's necessary condition in some way.¹⁰² This strategy has the advantage over the first strategy of not having to get involved in controversial debates over the metaphysics of harm.

In this chapter, I argue that the first strategy and two seemingly plausible ways of pursuing the second strategy do not work. First, in section 2, we consider two seemingly plausible ways we may revise the Interest Theory (Canonical). First, that only one's wellbeing considered *pro tanto* needs to ground duties, and not one's wellbeing all-things-considered. Second, that one's wellbeing needs to ground duties only under normal circumstances and not necessarily as things turn out. Neither of these accounts offer a full solution to the Problem of Preemption. The rest of the chapter then turns to the first strategy—replacing the Counterfactual Account. In sections 3 and 4, I present and object to Temporal and Non-Comparative Accounts of Harm and Benefit respectively. Finally, in section 5, I bolster my support for the Counterfactual Account by showing that other potential problems with that account are not too worrying. This sets the stage for the following chapter in which I introduce the Safety Condition as a solution to the Problem of Harmless Wronging.

Before we begin, a preliminary. *Plane Crash* is an example of harm that is preempted. Cases in which harm is preempted are similar to cases in which harm is overdetermined. Suppose I shoot Victim. Had I not shot Victim, Threatener would have. There, the harm I cause Victim preempts the harm Threatener would have caused were I not to have shot. Now suppose that Threatener and I shoot Victim simultaneously, where either shot alone would have been sufficient to kill Victim. There, the harm that each of us causes is overdetermined. Since the Counterfactual Account says *Y* harms *X* iff *Y* makes *X* worse off than *X* would have been had *Y* not acted as she did, it cannot accommodate either preempted or overdetermined harm. Victim would have been equally worse off in either case had I not shot.

¹⁰² It is an open question whether this is a full solution to the Problem of Harmless Wronging. It is open because it depends on whether the weakened necessary condition can handle our cases of pure risk.

Our focus is going to be on preemption. When harm is overdetermined, each person lays an equal claim to having caused the harm; when harm is preempted, intuitively, only one person lays claim to having caused the harm—in our case, me. This is important. In either case Victim is worse off than she would have been had Threatener *and* I not shot. So, it is open to a defender of the Counterfactual Account to say, though neither of us individually harm Victim, we together harm Victim. Now, while it might be somewhat plausible to say we together harm Victim when we simultaneously shoot, it is counterintuitive when harm is preempted. So, Threatener can say: “Wha’d’ya mean, *we* harmed Victim—I didn’t shoot her, you did. Yes, I *would’ve* harmed her had you not harmed her, but you did harm her so I didn’t!” This is even stranger in cases in which the thing that causes the harm is not an agent: suppose a tree falls on Victim and, had the tree not fallen on Victim, Threatener would have shot. It is odd to think that the tree and Threatener together harm Victim. Because of this, it is best to see how we can deal with the Problem of Preemption without saying *we* caused the harm.¹⁰³

2. Revising the Necessary Condition

2.1 *Pro Tanto*

A tempting thought to begin with is that, in *Plane Crash*, although Passenger is not harmed all-things-considered by being denied admittance onto the plane, she is made worse off in a regard (she is harmed *pro tanto*). For example, she suffers inconveniences that she would not have suffered were she not to have been arbitrarily denied admittance onto the plane. And, though the extent to which she is made worse off is insufficient to make her worse off all-things-considered, there is nonetheless *this regard* in which she is worse off through being denied admittance onto the plane. Perhaps she ought to be afforded a right protecting those aspects of her wellbeing.

¹⁰³ For defence of these plural harm views, see: (Parfit 1984, 82–83; Feit 2015). For a similar objection to mine, see (Hanser 2008, 436). For independent problems, see (Norcross 2005).

If we want to say this, we would need to endorse the following revision to the Interest Theory (Canonical):

Pro Tanto Thesis. For X to have a right against \mathcal{Y} that $\mathcal{Y} \Phi$, some aspect of X 's wellbeing must be of sufficient weight to place \mathcal{Y} under a duty to Φ .

Because Passenger is made worse off in a regard, she can be afforded a right protecting that aspect of her wellbeing.¹⁰⁴

Many see the Pro Tanto Thesis as plausible. For example, Raz says 'what one can have a right to may be in one's interest to have in some respects but not in others' (Raz 1994, 45).¹⁰⁵ Simon Căbulea May goes as far as to say that: 'No plausible version of the interest theory should assert that the ascription of rights turns on a comprehensive assessment of each person's wellbeing' (May 2017, 90).

Now, a task for the defender of the Pro Tanto Thesis is to determine how we parse one's wellbeing all-things-considered from one's wellbeing pro tanto. Since the Pro Tanto Thesis is ultimately not going to solve the Problem of Preemption, let us assume that task can be fulfilled.

Why might one endorse the Pro Tanto Thesis? One reason might be that such an endorsement solves a few other problems for the Interest Theory. Recall our other cases of harmless wronging (chapter 1, section 5):

Promise. Beth promises to meet Ann at the pub at 12. Ann only goes to the pub because Beth has promised to meet her. Ann would prefer, and it would be better for Ann, if Beth were not to show up. Beth does not show up.

¹⁰⁴ That is, assuming the other necessary and sufficient conditions on right ascriptions are satisfied. Since we are assuming Passenger would have a right against being denied admittance onto the plane were it not to have crashed, we can assume these other conditions are satisfied—it is Interest Theory (Canonical) that causes us trouble with the Problem of Harmless Wronging. Because of this, if our revision to Interest Theory (Canonical) is satisfied, I assume hereafter these other conditions will be satisfied.

¹⁰⁵ See, also, (Cruft 2004, 372; Tasioulas 2015, 48–49).

Paternalism. Carl has stopped smoking but, in a moment of weakness of will, has bought a packet of cigarettes. It would be better for Carl if Dave were to take the cigarettes to avoid his relapse into nicotine addiction.

May says of *Paternalism*, while it might be better if Dave were to take the packet of cigarettes, '[Carl] holds a claim-right against him that he not steal the cigarettes insofar as [Carl] has an interest in keeping [his] possessions' (May 2017, 90). We could tell a similar story of *Promises*.¹⁰⁶

As it turns out, we need not take a stand on whether to endorse the Pro Tanto Thesis. While the Pro Tanto Thesis might solve *Plane Crash*, in that it accommodates the intuition that Passenger has her rights violated, it will not solve all cases of preempted harm—it will not provide a full solution to the Problem of Preemption. This is because it provides a solution to *Plane Crash* only because Passenger is made worse off along some specific dimension—a specific dimension that she would not have been made worse off in were Attendant to have allowed her onto the plane (while the inconveniences of not catching a plane are greatly outweighed by death, these two sorts of harm are different). However, this feature does not hold in all cases of preempted harm. Consider:

Hitmen. Suppose that we have two hitmen. Hitman₂ admires Hitman₁. Hitman₂ secretly follows Hitman₁ on every job she has in the hope that, one day, Hitman₁ will fail to complete a hit and she will be able to do so instead, thereby impressing Hitman₁.

For any victim that Hitman₁ is contracted to kill (call her Victim), what aspect of her wellbeing is setback by Hitman₁ that would not be setback by Hitman₂? I am sceptical there is any such aspect, assuming Hitman₂ would complete the hit the instant that

¹⁰⁶ We need to be careful in specifying these pro tanto aspects of one's wellbeing in a non-circular way. We cannot appeal to an aspect of one's wellbeing *created* by the duty and correlative right (e.g., that it makes Ann's life go better for her when Beth keeps her promises because this means Beth does not disrespect Ann by not keeping her promises made to Ann). Yet, Ann may have an interest in having promises to her kept as it allows her to rely on future promises, forge normatively significant bonds, and so on. For more on this, see (Cruft 2017; 2019, 26–30).

Hitman₁ fails to.¹⁰⁷ Accordingly, the Pro Tanto Thesis cannot account for why Victim has a right against Hitman₁ killing her. We ought to look beyond the Pro Tanto Thesis, then.

The way I am thinking about pro tanto harm is that we see if there is a regard that Victim is worse off in that she would not otherwise have been worse off in. And, there is no aspect of Victim's wellbeing that is diminished in the actual world compared to the closest counterfactual world in which Hitman₁ does not shoot. But one might find it implausible that my simple counterfactual account of pro tanto harm cannot recognise even pro tanto harm in *Hitmen*. So, perhaps we should think of pro tanto harm differently. In section 4, we look at a drastically different account of pro tanto harm. In the meantime, let us consider a less drastic revision.

Ben Bradley endorses the Counterfactual Account when it comes to all-things-considered harm. But he thinks:

Something is a [pro tanto] harm for a person if and only if either (i) it is intrinsically bad for that person, or (ii) it brings about something intrinsically bad for that person, or (iii) it prevents something intrinsically good for that person. (Bradley 2009, 66)¹⁰⁸

Bradley's way of thinking about pro tanto harm differs from mine. In *Hitmen (Nonlethal Variant)*, suppose Hitman₁ *nonfatally* shoots Victim and, had they not shot, Hitman₂ would have nonfatally shot Victim in the same way. In this case, my view does not recognise any pro tanto harm. There is no regard that Victim is worse off in that she would otherwise not have been worse off in had Hitman₁ not shot. Why? Because Hitman₂ would have shot her in the same way. However, Bradley's view *can* recognise a pro tanto harm: being shot is intrinsically bad for Victim, so is a pro tanto harm for Victim; since Hitman₁ caused that pro tanto harm, Hitman₁ pro tanto harms Victim. As Bradley's account can recognise

¹⁰⁷ Perhaps one could postulate an interest in not being killed *by* Hitman₁, an interest that Hitman₂ does not setback. It is difficult to translate this "interest" into an aspect of Victim's wellbeing. And, when all else is equal, whether one is actually killed by Hitman₁ or Hitman₂ does not seem to matter (hold equal that Victim has no personal relationship with either party, and so on).

¹⁰⁸ Bradley actually speaks about prima facie harm, but he means pro tanto: 'harms that are bad for a person in one way, but might also be good for that person in another more important way' (Bradley 2009, 66). The objections to the Noncomparative Account below (section 4) give us reason to doubt whether (i) and (ii) are sufficient conditions on pro tanto harm.

pro tanto harm, we can accommodate Victim having a right against being shot in *Hitmen* (*Nonlethal Variant*). Problem of Preemption thus solved.

Despite Bradley's view being able to deal with the Nonlethal Variant of *Hitmen*, it cannot deal with our original lethal variant. Let us proceed on the plausible assumption that when one dies, one ceases to exist. Given this, death cannot be bad because of its intrinsic features—'For nonexistence has no intrinsic properties, positive or negative' (McMahan 2002, 98). This means that the badness of death must be *comparative*: bad in virtue of what good it excludes.¹⁰⁹ Finally, let us specify that Victim's death in *Hitmen* is painless.

Since *Hitmen* involves painless death, only condition (iii) is relevant for our purposes (it cannot be intrinsically bad for that person since it, death, is nothing). However, Hitman₁'s shooting Victim does not *prevent* anything intrinsically good for Victim. Why not? Because Hitman₂ would have shot Victim anyway. So, condition (iii) is not satisfied. Given Bradley's definition of pro tanto harm, Victim is not harmed pro tanto by Hitman₁ shooting her. Accordingly, the Pro Tanto Thesis cannot account for why Victim has a right against Hitman₁ killing her even if we think of pro tanto harm differently from how I have been thinking about it. This, then, gives us reason to move beyond the Pro Tanto Thesis.

(Offhand, it is not only cases involving death that Bradley's account cannot deal with. Rather, it is cases that (a) do not involve intrinsic bads and (b) in which the prevention of intrinsic goods is preempted. Suppose a secret admirer has slipped some tickets for the basketball into Justin's mailbox. Ben comes across the tickets and takes them without Justin ever knowing they were there. Had Ben not taken the tickets, Adam would have taken them. Ben taking the tickets (i) is not intrinsically bad for Justin, (ii) does not bring about something intrinsically bad, nor (iii) does it prevent something intrinsically good. So, it is not pro tanto harmful for Justin.)

¹⁰⁹ This is very much the standard view on the badness of death (McMahan 2002; Bradley 2009; Timmerman 2016).

2.2 Normality

To return to *Plane Crash*, another thought that we might have is that planes do not normally crash. Perhaps we ought to revise the Interest Theory along this dimension as a way to respond to the Problem of Preemption.

Normality Thesis. For X to have a right against \mathcal{I} that $\mathcal{I} \Phi$, X 's wellbeing (her interests) must, *under normal circumstances*, be of sufficient weight to place \mathcal{I} under a duty to Φ .

Passenger is worse off than she would have been under normal circumstances—for, under normal circumstances, had she not been denied admittance onto the plane, she would have landed safely at her destination. The extent to which she is worse off is sufficient to place Attendant under a duty not to deny her admittance onto the plane. So, given the Normality Thesis, Passenger can be attributed a right against being denied admittance onto the plane.¹¹⁰

Appealing to something like the Normality Thesis is prominent in the literature. For example, Neil MacCormick's version of the Interest Theory says,

to ascribe to all members of a class C a right to treatment T is to presuppose that T is, *in all normal circumstances*, a good for every member of C , and that T is a good of such importance that it would be wrong to deny it to or withhold it from any member of C (MacCormick 1982, 160, emphasis added).

Similarly, Kramer's formation of the Interest Theory appeals to what is *typically* beneficial for a being like X (Kramer 2017, 49, emphasis added). And, when considering cases of preempted harm, Joel Feinberg suggests that we ought to amend the Counterfactual Account of Harm with an appeal to normality: ' A harms B only if his wrongful act leaves B worse off than [*sic*] he would otherwise be *in the normal course of events insofar as they were reasonably foreseeable in the circumstances*' (Feinberg 1986, 153). Were we to endorse both Feinberg's revised Counterfactual Account and the Interest Theory (Canonical), Passenger would have a right against being denied admittance onto the plane. Ignoring Feinberg's

¹¹⁰ Subject to the other necessary conditions on rights being satisfied.

“insofar as they were reasonably foreseeable in the circumstances”, this view and the MacCormick/Kramer view are extensionally equivalent. They differ on intension: the MacCormick/Kramer view would say agents are not harmed by, but have rights against, preempted harm; the Feinberg view says we are harmed by preempted harm and therefore have rights against it.¹¹¹

(Let us pause and note that MacCormick and Kramer go further than the Normality Thesis in normalising the agent as well as the circumstances. We ought not go this far. If we normalise the agent, why would someone violating my rights wrong *me*, rather than wrong everyone. After all, I only have that right because the normal or generic person’s wellbeing would be of sufficient weight to place the duty-bearer (on grounds of parity, also considered under normality or generically) under a duty. Or, since MacCormick and Kramer endorse Nonjustificatory versions of the Interest Theory, I have that right only because the normal or generic person’s wellbeing would be served by the duty. On this picture, rights lose what Cruft has called *individualistic justification*; but individualistic justification nicely accounts for why the duty is owed to the individual, why its violation wrongs them, and why it manifests disrespect to them (Cruft only makes this final point) (Cruft 2013b, 205–6).)

Why might one appeal to the Normality Thesis? Think back to our other cases of harmless wrongdoing. Plausibly, under normal circumstances, one is harmed by promise-breaking, interferences with autonomy, and trespass.¹¹² Given the Normality Thesis, we can account for why we have rights that the promises made to us are kept, as well as rights that

¹¹¹ Building normality into harm is odd, though. Suppose Passenger complains that she was made worse off all-things-considered by being denied admittance onto the plane (what can I say; she’s a philosopher). Attendant retorts that the plane has actually crashed, and so they benefited her as things turned out. Passenger replies, “Yes, but normally planes do not crash. Since you made me worse off than I would’ve been under normal circumstances had you not denied me admittance onto the plane, you harmed me.” Attendant: “How have I harmed you along those lines? The plane crashed. Had you been aboard, you’d be dead now. All that your appeal to normality shows is that, *normally*, I would’ve harmed you by denying you admittance onto the plane. Luckily for me, things didn’t go as usual this time, so I didn’t harm you.” Attendant is a bit insensitive, but has a point.

¹¹² It is notable that Kramer, while appealing to normality, goes for the all-things-considered reading of the Interest theory, whereas those who endorse the Pro Tanto Thesis make no appeal to normality.

our autonomy and property is not interfered with even if, as things turn out, violation of those rights is harmless.

However, this does not so much give us what we might call a deep reason to endorse the Normality Thesis. Instead, it gives us a dialectical reason to endorse the Normality Thesis *rather than* Interest Theory (Canonical). So one might say: “Great, you can solve those problem cases with the Normality Thesis. But that doesn’t tell me anything about *why* rights respond to normality.” Further, we have already seen that there are other ways that one might attempt to answer these problems.

A deeper reason for why we might appeal to normality is the following: if rights are determined by whether *X*’s interests are *actually* of sufficient weight to place *Y* under a duty, this opens up an epistemological gap between, on the one hand, what rights *X* holds and correlative duties *Y* is under and, on the other hand, *Y*’s knowledge of whether she is under those duties. This problem might be compounded depending on how we think of duties. For example, following Thomson, we might think that *Y*’s being under a duty amounts to her behaviour ‘being constrained in a certain way’ (Thomson 1990, 200). Perhaps it is implausible that one’s behaviour can be (normatively) constrained while one is unaware of whether they are under that duty. Appealing to normality helps bridge this gap—plausibly, we are always in a position to know what happens under normality. (In fact, I postulate this is the reason Feinberg includes “*insofar as they were reasonably foreseeable in the circumstances*” in his appeal to normality.)

However, there might be other ways that we can bridge this epistemological gap between what duties people are under and their knowledge of being under those duties.¹¹³ We might not even be worried about this epistemological gap, thinking it is running together questions of whether we are under duties with whether we are blameworthy for infringing those duties.

Alas, we have reached a dead end again. While the Normality Thesis does accommodate our intuition that Passenger has a right against Attendant, it will not solve all cases of

¹¹³ For example, we could move to an Evidence-Relative version of the Interest Theory (chapter 7, section 4). The Safety Condition may also help dissolve this gap.

preempted harm. For, harm may be preempted in the normal world and yet we still want to say that people should be attributed rights against those preempted harms. Consider *Hitmen* again. Because Hitman₂ *always* follows Hitman₁ on every job she has, any victim that Hitman₁ is contracted to kill will not have a right against Hitman₁'s action. This is because Victim would be no worse off under normal circumstances were Hitman₁ not to shoot since Hitman₂ would make Victim worse off to an equal extent. The harm, then, is preempted under normal circumstances.

One might reply that it is not obvious that the circumstances described in this case are normal of hitmen in general, but only of Hitman₁. Perhaps we ought to restrict the reference class to *normal* hitmen. (We are moving into the territory of MacCormick's and Kramer's appeal to normality; with that comes the worry I raised above.) But this will not do. Suppose that, so eager to make sure that their hits are completed, assassination agencies begin to *always* send their hitmen out in pairs. This means that for any victim it is normal that, were the first hitman to fail, another hitman would always be there to kill that victim.¹¹⁴

A less fanciful example than *Hitmen* is to imagine a polluting factory. Suppose that, because of certain factors, these kinds of polluting factories always pop up in close proximity to each other. For example, the factories always require a river to dump waste in, cheap land, close proximity to (cheap) work forces, and so on. On the supposition that one polluting factory is sufficient to harm on its own, any harm caused by a particular factory would under normal circumstances be preempted by another factory. This is because there would always be another factory nearby, that is itself polluting, that would have caused an equal harm.¹¹⁵

¹¹⁴ I think this holds on most accounts of normality, including non-statistical ones. For example, some people think, the less cause for explanation an event's occurrence would prompt, the more normal that event's occurrence is (Smith 2016, 38–45). But, if hitmen were always sent out in pairs, and people knew this, a second hitman's presence would not call for explanation. So, it would be normal. We return to this in section 4.3.2 of chapter 8.

¹¹⁵ "Cancer Alley" comes to mind, an 85 mile stretch of the Mississippi that produces one quarter of America's petrochemicals, each year dumping more than one billion pounds of toxic chemicals into the river (Shrader-Frechette 2002, 8–9). See, also, (Kagan 2011).

Another sort of normalcy we might appeal to is what we can call *deontic* normalcy. On this view, we look only to counterfactual worlds in which people are acting as they are required when considering how things would have fared had the potential duty-bearer not performed the action required by the would-be duty. This would mean the closest admissible counterfactual world in which Hitman₁ does not shoot Victim is not the possible world in which Hitman₂ shoots Victim—in that world, Hitman₂ is not acting as she is required. Rather, it is one in which no one shoots Victim. Since Victim is sufficiently worse off in the world in which Hitman₁ shoots her than the world in which neither hitman shoots, Hitman₁ is under a duty not to shoot Victim correlating with Victim having a right against Hitman₁ that she not shoot her.

However, appealing to deontic normalcy will not help with *Plane Crash*. Everyone is acting as they are required to act in the closest counterfactual world in which Attendant does not deny Passenger admittance onto the plane and the plane crashes. We would need to look to (1) the normal world in which (2) everyone is acting as they are required. Then, we need not worry about the closest counterfactual world in which the plane crashes. This combination of these two senses of normalcy appears ad hoc. There is also a danger of circularity. When looking to those worlds in which everyone is acting as they are *required*, directed duties cannot be contributing to this requirement; we are trying to determine what directed duties people are under. We also cannot allow these undirected duties to be grounded in harm-based considerations; we are trying to circumvent that harm may be preempted in the normal world. This means only agent-centred, non-harm-based, undirected duties can be contributing to requirement. We also need to move to a Nonjustificatory version of the Interest Theory: obviously enough, the grounds of the duty will not be the wellbeing of the right-holder, so the Justificatory version of the Interest Theory is not satisfied. I would prefer not to have to do any of these things.

3. The Temporal Account of Harm

In the previous section, I argued that it will not do to answer the Problem of Preemption by revising the canonical statement of the Interest Theory with either an appeal to pro tanto aspects of one's wellbeing or to normality. Let us leave this general strategy of

responding to the Problem of Preemption by revising the Interest Theory (Canonical) until the Safety Condition is introduced in the following chapter. Instead, let us take up the other strategy of revising the Counterfactual Account of Harm. Recall that strategy said, intuitively, agents are harmed even if the harm in question is preempted. Any view of harm that says otherwise, such as the Counterfactual Account, is intuitively mistaken. All we need to do to solve the Problem of Preemption is reject that account of harm and find a new one. This and the following section take up two alternative accounts of harm. Since this strategy of rejecting the Counterfactual Account objects that the Counterfactual Account is implausible as an account of harm and benefit, in the following two sections we assess competing views *as* accounts of harm and benefit, and leave aside discussion of rights.

3.1 Other Reasons to be Sceptical of the Counterfactual Account

Cases of preempted harm suggest, against the Counterfactual Account, that intuitively it is not necessary for a person to be harmed that they be made worse off than they otherwise would have been. However, it is not obvious that it is *sufficient* for a person to be harmed that they be made worse off than they otherwise would have been. This is because often it seems we do not harm people when we fail to benefit them. But, failing to benefit people leaves them worse off than they would have been had we benefited them. So, this will imply failure to benefit is a harm on the Counterfactual Account. Consider:

Golf Clubs. Suppose Batman purchases a set of golf clubs with the intention of giving them to Robin, which would make Robin happy. The Joker says to Batman, “Why not keep them for yourself?” Batman is persuaded. He keeps the golf clubs. (Bradley 2012, 397)

Batman makes Robin worse off than he would have been had he given Robin the golf clubs. Given the Counterfactual Account, on first pass this implies that Batman harms Robin. Yet, many are sceptical of this: ‘Merely failing to benefit someone does not constitute harming that person’ (Bradley 2012, 397). Let us call this the Problem of Omission.

A closely related problem with the Counterfactual Account, one that we might think parasitic on the Problem of Omission, is that ‘we often consider failing to be benefited as morally and significantly less serious than both being harmed *and* not being safe from

harm' (Shiffrin 1999, 121; 2012, 372). Yet, if we cannot distinguish harm from failure to benefit (given the Problem of Omission), how can we account for the more serious nature of harm rather than failure to benefit? Let us call this the Problem of the Harm/Failure to Benefit Asymmetry.¹¹⁶ These two problems are separate, though the latter parasitic on the former, for one may think there is no asymmetry concerning the moral significance of harm and failure to benefit, though nonetheless think that there is an metaphysical distinction between harm and failure to benefit.

So far, my intention has been only to motivate rejecting the Counterfactual Account. I delay saying anything in response to these problems on behalf of the Counterfactual Account until section 5.

3.2 *Introducing the Temporal Account*

Cases of preempted harm suggest that it is not necessary for a person to be harmed that they be made worse off than they otherwise would have been. Instead, perhaps it is sufficient that they are made worse off than they *were* prior to being harmed. For example, in *Plane Crash*, while Passenger is not worse off than she would have been had she not been denied admittance onto the plane, she is worse off than she was prior to being denied admittance on the plane. Perhaps it is because of this that she is harmed.¹¹⁷ Similarly, in *Hitmen*, perhaps Victim is harmed because she is made worse off through being killed than

¹¹⁶ Shiffrin actually thinks there are two asymmetries that the Counterfactual Account blurs: first, harm/failure to benefit, and second, avoidance of harm/bestowal of benefit. *Golf Clubs* is an example of the first asymmetry. Shiffrin suggest that an example of the second asymmetry is, other things being equal, I may pro tanto harm you to preempt a greater pro tanto harm, but not pro tanto harm you to bring about a larger pro tanto benefit (Shiffrin 2012, 363). See, also, (Gardner 2017).

¹¹⁷ Is Passenger harmed all-things-considered or merely pro tanto? As we see, the Temporal Account thinks it is sufficient for one to be harmed that they be made worse off than they were prior to being harmed. Since she is not made better off than she was prior to being denied admittance onto the plane, and only made worse off, it is difficult to see how a defender of the Temporal Account could plausibly argue that Passenger is harmed pro tanto but benefited all-things-considered. Yet, it is plausible that Passenger is harmed pro tanto but benefited all-things-considered by being denied admittance onto the plane. Straightaway we run into a problem with Temporal Account, similar to that discussed below with *Lessening Decline*. That said, the Temporal Account does handle *Hitmen* very well, where there is no inclination to see any all-things-considered or pro tanto benefit. So, in spite of the worry raised in this note, the view deserves entertaining.

she was prior to being shot, even if she is no worse off than she would have been had Hitman₁ not shot her.

If we want to say this, we should endorse the

Comparative Temporal Account of Harm and Benefit. Y harms X iff (and because) Y makes X worse off than she was prior to Y 's acting. Y benefits X iff (and because) Y makes X better off than she was prior to Y 's acting.¹¹⁸

The Temporal Account also seems not to fall foul of the Problem of Omission: Batman does not make Robin worse off than Robin was prior to Batman not giving him the golf clubs. Inasmuch, Batman does not harm Robin. Since the Temporal Account seems to distinguish between harm and failure to benefit, we might also think that it does not fall foul of the Problem of the Harm/Failure to Benefit Asymmetry.

Returning to the Problem of Preemption, since Attendant and Hitmen₁ harm Passenger and Victim in spite of their not leaving them worse off than they otherwise would have been, they can be attributed a right on the canonical statement of the Interest Theory.

3.3 Temporal Problems

Alas. There are problems with the Temporal Account. I think these problems are more fundamental than those problems with the Counterfactual Account. It is neither necessary nor sufficient for a someone to be harmed that they be made worse off than they were prior to being harmed. Let us begin with the claim that it is not sufficient for an someone to be harmed that they be left worse off than they were prior to being harmed.

¹¹⁸ The Temporal Account is endorsed in some form by (Thomson 1990, 261–63; Perry 2003; Velleman 2008, 242–44; Rabenberg 2014). (Thomson 2011) now rejects the Temporal Account. In chapter 3, section 3.2, I offered a more basic, event-based definition of the Counterfactual Account; we could offer an analogous more basic definition of the Temporal Account.

Lessening Decline. Badly-Off is suffering from a painful disease that is getting worse by the day. Nice gives Badly-Off some medicine that makes Badly-Off feel worse, but which slows the decline of her disease.¹¹⁹

By giving Badly-Off the medicine, Nice makes Badly-Off worse off than she was prior to Nice acting. According to the Temporal Account, she harms Badly-Off. Yet, this is implausible. Intuitively, Nice does not harm but actually benefits Badly-Off. This shows that it is not sufficient for someone to be harmed that they be left worse off than they were prior to being harmed—one does not harm someone else when lessening their decline, even if one makes them worse off than they were prior to being harmed. The explanation of *why* Nice benefits Badly-Off lies in, as the Counterfactual Account suggests, that she leaves Badly-Off better off than Badly-Off would have been had she not given the medicine.¹²⁰ She makes a difference, in a positive way, to how Badly-Off fares.

Let us turn to why it is not necessary for someone to be harmed that they be left worse off than they were prior to being harmed.

Placebo. Badly-Off is suffering from a painful disease. Naughty steals the medicine that would have cured Badly-Off and replaces it with sugar pills. The sugar pills make Badly-Off slightly better, but worse than she would have been had she received the medicine.¹²¹

By swapping Badly-Off's medicine, Naughty makes Badly-Off better off than she was prior to Naughty acting. According to the Temporal Account, she does not harm Badly-Off but benefits Badly-Off. Yet, this is deeply implausible. Intuitively, Naughty harms, and does not benefit, Badly-Off. This shows that it is not necessary to be harmed that one be left worse off than they were prior to being harmed. In a similar way to the preceding problem, the explanation of why she harms Badly-Off lies in that, as the Counterfactual

¹¹⁹ See, also, (Norcross 2005, 149–50; Tadros 2016a, 192).

¹²⁰ As Norcross puts it: 'We compare levels of welfare, not across time, but across worlds' (Norcross 2005, 150).

¹²¹ Hanser calls this a preventative harm since Naughty harms Badly-Off by preventing Badly-Off from receiving a benefit (Hanser 2008, 429; Boonin 2014, 60; Tadros 2016a, 192).

Account suggests, Naughty leaves Badly-Off worse off than Badly-Off would have been had she not swapped the medicine for the sugar pills. This time she makes a difference, in a negative way, to how Badly-Off fares.

I have suggested that it is not necessary for a person to be harmed that they be worse off than they were prior to being harmed. One might reply that, in *Placebo*, Naughty denies Badly-Off the *opportunity* of getting better. And so, perhaps Badly-Off is worse off than she was prior to Naughty acting, in virtue of not having the opportunity of getting better. For example, Thomson says of a similar case, ‘he [Naughty] caused his victim’s chances of getting a benefit to shrink, thereby causing [her] victim to be worse off in a way than he was’ (Thomson 2011, 445).¹²² Let us call this the Opportunity Reply.

Before assessing the Opportunity Reply, a clarification. Intuitively, we want to say that Badly-Off is worse off all-things-considered than she was prior to Naughty acting, and not that she is worse off only pro tanto. This means that the harmfulness of being denied the opportunity of getting better needs to be greater than the benefit she receives from the placebo-effect. To see this, suppose that, at t_1 (prior to Naughty acting), Badly-Off is at wellbeing level x . And suppose the pro tanto benefit she receives from the placebo effect is y and the pro tanto harmfulness of being denied the opportunity of recovering is z . This means that at t_2 (after taking the placebo) Badly-Off is at wellbeing level $x+(y-z)$. From this we can see that z needs to be greater than y , otherwise Naughty only harms Badly-Off pro tanto, though benefits Badly-Off all-things-considered.

First, it is unclear whether the Opportunity Reply extends to show that it is not sufficient for someone to be harmed that they be left worse off than they were prior to being harmed. In *Lessening Decline*, Nice does not furnish Badly-Off with the opportunity of

¹²² It is surprising that Thomson says this given what she says in earlier work. She says ‘if it would be bad for X to get a thing Z, and if Y makes it probable that X will get Z, then Y causes X to be at a *disadvantage*’ (Thomson 1990, 244). Thomson then considers whether Y *harms* X merely by subjecting her to risk of harm—by putting her at a disadvantage. She says: ‘this is a bad argument, for we cannot really say that causing a person to be at a disadvantage is itself causing the person a harm’ (Thomson 1990, 244). But if causing someone to be at a disadvantage is not in itself harmful, why would having opportunities (being at an advantage) be beneficial? (Somewhat ironically, in 1990, Thomson endorsed the Temporal Account but suggests this argument, needed to supplement the Temporal Account, is bad; in 2011, Thomson abandons the Temporal Account, though mounts this ‘bad argument’ on its behalf.)

getting better—she just lessens Badly-Off’s decline. So, it is unclear how the Temporal Account could avoid the counterintuitive result that Nice harms Badly-Off in *Lessening Decline*. This means, even if the Opportunity Reply works to account for why Badly-Off is harmed in *Placebo* (which we see below that it does not!), we still have reason to be sceptical of the Temporal Account because it implies Badly-Off is harmed and not benefited in *Lessening Decline*.

Second, the Opportunity Reply misidentifies the harm in *Placebo*. Straightforwardly, Badly-Off is harmed in virtue of not getting better and not in virtue of being denied the opportunity of getting better. As Victor Tadros suggests, ‘it is the recovery that [Badly-Off] cares about, not the option of recovering’ (Tadros 2016a, 193).

Third, in chapter 6, we see that merely being subjected to risk of harm is not itself harmful. Since benefit is the other side of the coin from harm, this shows that having the opportunity of a benefit is not itself beneficial. So, removing the opportunity of a benefit does not itself make one worse off.

Even if one is not convinced by the previous objections (the third of which the reader is having to take on good faith at this stage), we can mount a fourth, dialectical reply. The Opportunity Reply suggests that Naughty makes Badly-Off worse off than she was prior to Naughty acting in virtue of removing the opportunity of getting better. However, to determine whether the opportunity of getting better is actually beneficial to Badly-Off, we need to look to the counterfactual world in which she receives the medicine and see how she fares. Seana Shiffrin suggests that, because of this, the Temporal Account collapses into the Counterfactual Account (Shiffrin 2012, 368, fn. 25).¹²³ This is not strictly correct: the Temporal Account still says that one is harmed *if and only if*, and *because*, one is made worse off than they were; it is just that one can be made worse off than they were by being denied the opportunity of a benefit, were the value of that opportunity is determined by counterfactual-analysis. However, what this does mean is that we cannot evade the

¹²³ Tadros objects that we need to appeal to counterfactuals to determine the magnitude of harm on the Temporal Account (Tadros 2016a, 192). Perry says we need to appeal to counterfactuals to determine the damages owed as a matter of compensation (Perry 2003, 1296). Neither notice that this means the Temporal Account inherits the Counterfactual Account’s problems with preemption as suggested below.

problem caused by preemption since it may be that, in the closest world in which Naughty does not swap the medicine, the benefit is preempted by some harm. This means that the opportunity Badly-Off is denied will not be valuable. So, the removal of this worthless opportunity does not make her worse off than she was. For example, consider

Placebo (Preempted Variant). Badly-Off is suffering from a painful disease. Naughty steals the medicine that would have cured Badly-Off and replaces it with sugar pills. The sugar pills make Badly-Off slightly better. Were Naughty not to have swapped the pills, Also-Naughty would have swapped the actual medicine for sugar pills.

Since opportunities are valued in counterfactual terms, the problem caused by preemption re-emerges. Naughty does not deny Badly-Off the opportunity of getting better in *Placebo (Preempted Variant)* since Also-Naughty would have swapped the medicine for a placebo anyway.

4. The Noncomparative Account

I have just argued that, against the Temporal Account, it is neither necessary nor sufficient for someone to be harmed that they be made worse off than they were prior to being harmed. Before that, we saw that we have reason to think, against the Counterfactual Account, it is neither necessary nor sufficient for someone to be harmed that they be made worse off than they otherwise would have been. Perhaps the lesson we ought to draw from this is that harm is not *comparative*.

4.1 *Introducing the Noncomparative Account*

This thought—that harm is not comparative—has some prominence in the literature. For example, Shiffrin says:

Accounts that identify harms with certain absolute, noncomparative conditions (e.g., a list of evils like broken limbs, disabilities, episodes of pain, significant losses, death) and benefits with an independently identified set of goods (e.g., material enhancement, sensual pleasure, goal-fulfilment, nonessential knowledge,

competitive advantages) would not generate these puzzles [associated with comparative accounts]. (Shiffrin 1999, 123)

Elizabeth Harman says ‘an action harms someone if it causes the person to be in a bad state. Bad states are understood as states that are in themselves bad, not bad because they are worse than the state the person would otherwise have been in’ (Harman 2009, 139). Let us formulate this view.

Noncomparative Account of Harm and Benefit. Y (pro tanto) harms X iff (and because) Y makes it the case that X is in a noncomparatively bad state. Y (pro tanto) benefits X iff (and because) Y makes it the case that X is in a noncomparatively good state.¹²⁴

The Noncomparative Account is a view of pro tanto and not all-things-considered harm.¹²⁵ This is because, on the all-things-considered reading of the view, one would harm someone only if one made it the case that they were in a noncomparatively bad state all-things-considered. But this is implausible. For example, luckily, to make my life noncomparatively bad all-things-considered, one would have to *significantly* lessen my well-being. But surely I can be harmed all-things-considered by events that fall much shorter than such drastic changes to my life.

Given that the Noncomparative Account is an account of pro tanto harm, ‘it is compatible with the comparative account since they are about different things’ (Bradley 2012, 399). But this disunified account of pro tanto and all-things-considered harm is odd. For example, think again of *Hitmen (Nonlethal Variant)*, in which Hitman₁ non-fatally shoots Victim and, had she not shot, Hitman₂ would have nonfatally shot Victim in the same way.¹²⁶

¹²⁴ (Shiffrin 1999; 2012; Harman 2004; 2009). Woollard thinks there is a noncomparative and a comparative sense of harming, and that, ‘[o]ther things being equal, our reasons against harming are much stronger if we would harm a person in the overall-comparison sense’ (Woollard 2012, 686). Again, we could offer a more basic definition of the Noncomparative Account.

¹²⁵ Harman says that she is not denying that we sometimes use harm in an all-things-considered way, indicating that she sees her condition as merely a sufficient condition on pro tanto harm (Harman 2004, 109; Thomson 2011, 439).

¹²⁶ I use a non-lethal variant of *Hitmen* to avoid discussion of whether the Noncomparative Account can recognise death as harmful (Bradley 2012, 400–401). This depends on what makes a state noncomparatively bad.

Here, given the Noncomparative Account there is (i) pro tanto harm and (ii) no pro tanto benefit, but given the Counterfactual Account there is (iii) no all-things-considered harm and (iv) no all-things-considered benefit. This is very odd. Underlying the oddness is something like the following plausible thought: ‘[all-things-considered] harm and *pro tanto* harm are closely related. In fact, they seem interdefinable. Whether something is all-things-considered harmful is a function of the ways in which it is *pro tanto* harmful or beneficial’ (Bradley 2012, 393). This picture runs afoul of this very plausible thought. So, I am not sure where this leaves the defender of the Noncomparative Account when it comes to all-things-considered harm. (Perhaps they will deny all-things-considered harm.) This does mean one will have to endorse the Pro Tanto Thesis from above.¹²⁷

So, what reason do we have to endorse the Noncomparative Account? It can recognise preempted harm as, at the least, pro tanto harm. In *Hitmen (Nonlethal Variant)*, Hitman₁ makes it that Victim is in a noncomparatively bad state; given this, she (pro tanto) harms Victim. The Noncomparative Account seems also to get the cases right when it comes to the Problem of Omission. In *Golf Clubs*, Batman does not make it that Robin is in a noncomparatively bad state, so Batman does not harm him. Since receiving gifts seems like the sort of thing that would not be out of place on Shiffrin’s list of benefits, Batman fails to benefit Robin. Further, since the defender of the Noncomparative Account can distinguish between harm and benefit, they are in a good place to answer the Problem of the Harm/Failure to Benefit Asymmetry.¹²⁸

¹²⁷ Defenders of the Noncomparative Account will have to make similar meta-normative claims when wanting the Noncomparative Account to do whatever work it is introduced to do. For example, were one to introduce the Noncomparative Account as a solution to the Non-Identity Problem (see the following note), one would also have to claim that we have reason not to harm pro tanto where our harming pro tanto cannot be justified by bestowing pro tanto benefits.

¹²⁸ The Noncomparative Account was primarily offered as a solution to the Non-Identity Problem. Suppose Mother is considering conceiving but is suffering from condition *x*. She is warned by her doctor that if she has a child in her current state, her baby will be born with a serious disability; if Mother waits two months before conceiving, her child will be born without the disability. Mother decides, on a whim, to have the child anyway, whom she calls Child. Child is subsequently born with a serious disability, but one that does not make her life not worth living (i.e., Child would not prefer nonexistence to existence with such a handicap). Some people think Mother harms Child by failing to wait two months before conceiving. However, were Mother to have waited two months to conceive, a different egg would have been fertilised with a different sperm and so a different person would have been born. Accordingly, the Counterfactual Account implies Child is not harmed by Mother’s action since she is not worse off than she would have been had

As Shiffrin appreciates, ‘[i]f one could, in addition, provide an account that explained what unified such items, why they were together classified as harm, one might make further inroads to support [...] the asymmetries’ (Shiffrin 1999, 123; 2012, 376). Her own account is that ‘harm involves a distinctive sort of frustration or impediment of the will of the ability to exert and effect one’s will’ (Shiffrin 2012, 383; 1999, 123). However, this account is not very convincing. This is because Shiffrin is confusing two different questions (Tadros 2016a, 179). First, we can ask whether the sort of thing that is affected when we are harmed is wellbeing, or whether it is those things one has a stake in, or perhaps the relationship between one’s will and the real world, and so on.¹²⁹ Second, we can ask, spotting an answer to this first question, how ought we determine whether one has been harmed. Ought we compare how their [wellbeing/things they have a stake in/relationship between world and will] fares with how it would have fared otherwise (Counterfactual Account), with how it was prior to acting (Temporal Account), or do we not compare it whatsoever (Noncomparative Account)? Shiffrin’s speculation that ‘harm involves conditions that generate a significant chasm or conflict between one’s will and one’s experience’ concerns the first question, and so is not very helpful when trying to defend the Noncomparative Account, which is an answer to the second question. For example, we could think one is harmed when the chasm between one’s will and one’s experience is made greater than it otherwise would have been, thereby endorsing a version of the Counterfactual Account. Perhaps Shiffrin’s thought is, “Look, these states are noncomparatively bad *because* they create a chasm between one’s will and one’s experiences.” But it is not at all clear why that is true only of noncomparatively bad states.¹³⁰

Mother waited two months to conceive (Parfit 1984, 351–80). However, Mother harms Child according to the Noncomparative Account as she makes it that Child is in a non-comparatively bad state. I am inclined to think Mother does not harm Child, nor do I think this would explain why her action is wrong (Boonin 2014). For further discussion, see (Bradley 2012, 398; Tadros 2016a, 189–90).

¹²⁹ The first of these is the view we have been assuming, the second is Feinberg’s view of interests, the setting back of which constitute harms (Feinberg 1984, 33–34), and the third is Shiffrin’s.

¹³⁰ For explicit objections to this focus on autonomy as determining what sort of thing is affected when we are harmed, see (Bradley 2012, 400; Tadros 2016a, 179–80). Given my discussion of autonomy elsewhere in this thesis, one should be able to guess what I would say about such an account (e.g., chapter 6, section 3).

Harman says that her ‘list of harms is unified by comparison with a healthy state (though I haven’t claimed that all harms meet this condition)’ (Harman 2004, 111). I cannot say I like the look of this picture either. But I am not going to take that up. Rather, in the following subsection, we see we have good reason to be sceptical of the Noncomparative Account independently from not having a robust picture of what states are harmful and which are beneficial.

4.2 Noncomparative Problems

We spent quite a while introducing the Noncomparative Account. The reasons to reject the view can be stated relatively quickly. The principal problem with the Noncomparative Account is that it is neither necessary nor sufficient for someone to be (pro tanto) harmed that they be left in a noncomparatively bad state.

First, it is not necessary for one to be harmed that they be left in a noncomparatively bad state, for one can be harmed through being moved from a noncomparatively good state to a worse, though nonetheless, noncomparatively good state. Here is an example:

Well Off. Well-Off has a very high IQ. Naughty poisons Well-Off, reducing her IQ by a few points, though leaving her IQ well above average.¹³¹

Intuitively, Naughty harms Well-Off. However, since Naughty does not put Well-Off into a noncomparatively bad state, the Noncomparative Account cannot arrive at this verdict. The explanation of why Naughty harms Well-Off lies in that, as the Counterfactual Account suggests, Naughty leaves Well-Off worse off than she would have been had Naughty not poisoned her, even though she leaves her well off in noncomparative terms. She makes a difference, in a negative way, to how Well-Off fares.

Often, people with sympathies to the Noncomparative Account respond by attempting to locate a noncomparatively bad state that Well-Off is in after being poisoned (Thomson 2011, 440). For example, might one say that Naughty dominates, controls, or manifests

¹³¹ See: (Hanser 2008, 432; Tadros 2016a, 188).

disrespect towards Well-Off—is not being the subject of these things a noncomparatively bad state to be in?

These sorts of explanations seem misguided to me—Naughty harms Well-Off, in the most straightforward sense, because she diminishes her IQ. Any putative noncomparatively bad state one postulates is beside the point. If one is unconvinced, here are three further reasons to be sceptical. First, the line of argument is appealing to what we might call *respect-based* considerations. Yet, suppose that it is not Naughty that poisons Well-Off, but Well-Off who accidentally poisons herself by drinking from a dirty stream. Well-Off is still harmed by having her IQ diminished. But how so on the Noncomparative Account? The water does not dominate, control, or manifest disrespect towards Well-Off. Second, it is unclear that harm can play much of a normative role on this account; this is because we are already presupposing an account of what is owed to whom, the disrespecting of which puts people into these noncomparatively bad states. And third, in any case, it is unclear that being the subject of domination, control, or disrespect puts one in a noncomparatively bad state in terms of one’s wellbeing independently from the effects this has on one.

So, it is not necessary to be harmed that one be in a noncomparatively bad state. Yet, it is also not sufficient for someone to be harmed that they be put into a noncomparatively bad state. This is because one can be benefited, and not harmed, when being moved from one noncomparatively bad state to a comparatively better, though still noncomparatively bad, state. Consider:

Badly Off. Badly-Off is in a severely, noncomparatively bad state, such as pain so bad it is immobilising. Nice gives Badly-Off some medicine that lessens Badly-Off’s pain to be merely nauseating.

Nice benefits and does not harm Badly-Off. Yet, Nice puts Badly-Off into a noncomparatively bad state (nauseating pain is noncomparatively bad). Given the Noncomparative Account, Nice harms Badly-Off. This plainly will not do.¹³² The explanation for why Nice

¹³² This objection is also raised by (Hanser 2009, 188; Thomson 2011, 441; Tadros 2016a). Shiffrin says if someone’s interests are advanced, though they’re still left badly off overall, ‘it seems strained to say that she has been benefited’ (Shiffrin 1999, 122–23). I do not see how this seems strained in *any* way.

benefits Badly-Off lies in that, as the Counterfactual Account suggests, Nice leaves Badly-Off better off than she would otherwise have been, even though she leaves her badly off in noncomparative terms. She makes a difference, in a positive way, to how Badly-Off fares.

5. Back to the Counterfactual Account

Here is where we are. I began section 3 by saying that we have reason to think the Counterfactual Account fails because it appears to be neither necessary nor sufficient for one to be harmed that one be worse off than they otherwise would have been. This was the problem posed by preempted harm and the Problem of Omission respectively. We then examined whether there are any accounts that can take the place of the Counterfactual Account. We saw that of the two accounts considered, neither comes without its own troubles. I pressed at each stage that both accounts are limited *because* they fail to recognise that one can be harmed or benefited by another's action in virtue of the difference the event makes to how one would otherwise have fared. Because of this, even if there was nothing else that could be said on behalf of the Counterfactual Account in reply to the problem posed by preemption and the Problem of Omission, I think we would be best placed to stick with the Counterfactual Account.¹³³

In the following chapter, I argue that the Problem of Preemption can be solved for the Interest Theory. This means even if individuals are not harmed when what otherwise would be harmful is preempted by another would-be harmful event, there is no problem in our having rights against these actions. (I also argue this is a more plausible way to go than tinkering further with the Counterfactual Account.) That still leaves us with the Problem of Omission and the Problem of the Harm/Failure to Benefit Asymmetry. Though I said we are best placed to stick with the Counterfactual Account despite these problems,

¹³³ For example, *Placebo* showed that the Temporal Account cannot recognise *preventative* harms: Naughty harms Badly-Off through preventing Badly-Off from receiving a benefit (Hanser 2008, 429). Boonin suggests we ought to be more worried if our account of harm cannot recognise preventative harms as harms (as the Temporal Account cannot) than if our account cannot recognise preempted harms as harms (as the Counterfactual Account cannot) (Boonin 2014, 60).

we could bolster this claim if there was something else we could say in reply to the Problem of Omission and the Problem of the Harm/Failure to Benefit Asymmetry.

Space does not permit me to say too much. I focus on the Problem of Omission. If we find some differences between what are, intuitively, harms and what are, intuitively, benefits, we could use this to try to explain the putative asymmetry between harm and failure to benefit.

First, recall our more basic definition of the Counterfactual Account, on which, an event, e , harms someone, X , iff X is worse off than she would have been had e not occurred.¹³⁴ For the vast majority of failures to benefit, it is not the case that there is some event that, had that event not occurred, X would have been better off in the closest counterfactual world. For example, while writing this thesis, I could have benefited my friend Sean in Canada. But it is not the case that, had most of the events that have occurred over that time not occurred, then Sean would have been better off in the closest counterfactual world.¹³⁵

Now, this does not help us with *Golf Clubs*.¹³⁶ In that case there are events that, had those events not occurred, Robin would have been better off in the closest counterfactual world. For example, had Batman not put the clubs back into the Batmobile, we can suppose Robin would have been better off as Batman would have gone on to give him the clubs. Yet, the preceding point does lessen the Problem of Omission when it comes to other cases.

Next, though Batman may harm Robin by failing to give him the golf clubs, he does not *do* harm to Robin in the sense that is familiar from discussion of the doctrine of doing and allowing harm. Rather, he *allows* harm. For example, on Warren Quinn's account, \mathcal{I} does harm to X iff \mathcal{I} 's most direct contribution to the harm to X is an action (Quinn 1989b). But it is unlikely our theory of action will have it that Batman's most direct contribution

¹³⁴ Chapter 3, section 3.2.

¹³⁵ For relevant discussion, see (Hanser 2008, 427; Klockslem 2012, 294; Hanna 2016, 252–53; Feit 2017, 3–5).

¹³⁶ Hanna and Feit note this (see previous note).

to Robin's not getting the golf clubs will be an action. Similarly, Jonathan Bennett's account says \mathcal{T} does harm to X iff most of the ways \mathcal{T} could have moved her body would have led to the harm (Bennett 1995). But most of the ways Batman could have moved his body when failing to give the golf clubs to Robin would not have led to the harm. What this means is that, even if a defender of the Counterfactual Account is committed to saying some of \mathcal{T} 's failures to benefit X are harmful for X , they are not committed to saying that \mathcal{T} does harm to X in this case. (We could even say, \mathcal{T} harms X only if she *does* harm to X ; so, Batman does not harm Robin.)

Finally, even if there are events that, had those events not occurred, Robin would have been better off, there are only these events because of other things that Batman has done; for example, his buying the golf clubs with the intention of giving them to Robin. And it is not obvious that had *those events* not occurred, there would have been some other event which, had *that* event not occurred, then Robin would have been better off in the closest counterfactual world. For example, let us take events associated with Batman's deciding to head into the golf shop where he formed the intention to gift Robin the clubs. True, *if* Batman goes into the shop and forms the intention to buy Robin the clubs, Robin might be worse off if Batman decides not to give him the clubs. But had Batman not gone into the shop, it is not the case that there is some event the non-occurrence of which leaves Robin worse off than he would otherwise have been in the closest counterfactual world. We can learn two things from this. First, roughly, had Batman not bought the golf clubs with the intention of giving them to Robin, then he would not have harmed Robin by not giving him the clubs (Klocksiesm 2012, 294). Second, the harm to Robin that occurs when Batman does not give him the clubs results from the removal of a benefit that Batman himself is responsible for. But we have independent reason from the growing literature on the removal of benefits to think this already confuses things (McMahan 1993; Hanser 1999).

The remarks in this final section have by no means meant to be conclusive. However, I hope to have shown there is more to be said on behalf of the Counterfactual Account to resist the Problem of Omission.

5. Safety

1. Introduction

Let us take stock. At the end of chapter 3, I reintroduced the Problem of Harmless Wronging for the Interest Theory. The problem is that of accommodating our intuitions that people can have rights against the performance of actions that do not harm them given a commitment to the Interest Theory. According to that theory, necessary for X to have a right against Y is that X 's wellbeing is of sufficient weight to place Y under a duty. And, whether X 's wellbeing is of sufficient weight to place Y under a duty depends, at the least, on the extent to which X would be harmed or benefited were Y not to act as the would-be duty dictates. But if X would not be harmed by the non-performance of the action required by the duty, the Interest Theory's necessary condition is not satisfied, and X has no right.

The Problem of Preemption is one specific form of the Problem of Harmless Wronging. Recall the two examples of preempted harm that we focused on in the previous chapter:

Plane Crash. Passenger is about to board a plane. Attendant takes a disliking to Passenger, so denies her admittance onto the plane. On departure, the plane crashes and everybody on board dies.

Hitmen. Suppose that we have two hitmen. Hitman₂ admires Hitman₁. Hitman₂ secretly follows Hitman₁ on every job she has in the hope that, one day, Hitman₁ will fail to complete a hit and she will be able to do so instead, thereby impressing Hitman₁.

In *Plane Crash*, Passenger is no worse off through being denied admittance onto the plane than she would have been had Attendant not denied her admittance onto the plane. This is because the would-be harmfulness of being denied admittance onto the plane is preempted by the greater would-be harmfulness of not being denied admittance. And in *Hitmen*, Victim is no worse off through Hitman₁'s shooting her than she would have been had Hitman₁ not shot. This is because the would-be harmfulness of being shot by Hitman₁ is preempted by the would-be harmfulness of being shot by Hitman₂.

In the previous chapter, we looked at two ways of revising the canonical statement of the Interest Theory to respond to the Problem of Preemption: first, by appealing to pro tanto aspects of one's wellbeing; second, by appealing to normality. I argued that neither refinement offers a full solution to the problem. We then looked at whether to revise the Counterfactual Account of Harm. Yet, I argued that competing accounts come with more worrying problems than those raised by the Counterfactual Account.

In this chapter, I introduce my solution to the Problem of Harmless Wronging: the Safety Condition. Roughly, it says for someone to hold a right against one that one not perform some action, we look to whether performing that action could easily leave that person sufficiently worse off to place one under a duty. Section 2 introduces the Safety Condition. Section 3 begins to build an account of why we should endorse the Safety Condition in addition to its extensional accuracy; the full defence comes in chapter 8 once we have seen how the Safety Condition deals with the Problem of Pure Risk. Section 4 defends the Safety Condition against three objections.

2. The Safety Condition and Preemption

2.1 Introducing Safety

Retracing some ground from the previous chapter leads us naturally to the Safety Condition. In *Plane Crash*, Passenger's wellbeing is not of sufficient weight to place Attendant under a duty not to deny her admittance onto the plane. This is because Passenger is no worse off through being denied admittance onto the plane than she otherwise would have been. We can see this by comparing the following two worlds.

World 1. Attendant denies Passenger admittance onto the plane. The plane takes off and crashes killing everybody on board.

World 2. Attendant does not deny Passenger admittance onto the plane. The plane takes off and crashes killing everybody on board.

Passenger is no worse off in world 1 through Attendant's denying her admittance onto the plane than she would have been in world 2. However, world 2 is not the only possible world available for comparison with world 1. We might look to

World 3. Attendant does not deny Passenger admittance onto the plane.

The plane takes off and lands safely at its destination.

Passenger is worse off in world 1 than she is in world 3. The extent to which she is worse off is sufficient to place Attendant under a duty not to deny her admittance onto the plane. The problem is that when considering what would have happened to Passenger had Attendant not denied her admittance onto the plane, world 3 is not the closest counterfactual world to world 1; that is world 2.

World 3 is what would happen under normal circumstances were Attendant not to have denied Passenger admittance onto the plane. This was the thought behind the Normality Thesis—by looking to what happens in the *normal* world in which Attendant does not deny Passenger admittance onto the plane, we are directed to world 3 and not to world 2. As we have seen, the problem with the Normality Thesis is that it is possible for harm to be preempted in the normal world. In *Hitmen*, because Hitman₂ always follows Hitman₁ on every job she has, Victim will be no worse off than she would have been under normal circumstances had Hitman₁ not shot. There, Hitman₂ would have made Victim worse off to an equal extent.

We are faced with a comparison between the following two worlds:

HM World 1. Hitman₁ shoots Victim. (Hitman₂ was waiting in the wings.)

HM World 2. Hitman₁ does not shoot Victim. Hitman₂ shoots Victim.

Victim is no worse off in HM world 1 than she is in HM world 2. HM world 2 is what would happen under normal circumstances were Hitman₁ not to shoot. However, as with *Plane Crash*, there is another close possible world that we can appeal to:

HM World 3. Hitman₁ does not shoot Victim. Hitman₂ does not shoot Victim.

Victim is worse off in HM world 1 than she is in HM world 3. The question is how to establish the salience of *these* close worlds (world 3 and HM world 3) in order to accommodate the intuition that Passenger and Victim have rights against the relevant conduct.

I suggest we appeal to *modal safety*. The idea behind safety is nicely explained by Timothy Williamson:

Imagine a ball at the bottom of a hole, and another balanced on the tip of a cone. Both are in equilibrium, but the equilibrium is stable in the former case, unstable in the latter. A slight breath of wind would blow the second ball off; the first ball is harder to shift. The second ball is in danger of falling; the first ball is safe. Although neither ball did in fact fall, the second could easily have fallen; the first could not. The stable equilibrium is [safe]; the unstable equilibrium, [unsafe]. (Williamson 2000, 123)

There is a danger an event will occur if that event does occur in some sufficiently similar case. Much like as the ball is not safely balanced on the top of the cone, Passenger's and Victim's wellbeing is not safely protected—though they are not actually made worse off by the violation of their rights, there is a danger they could have been. And, it is plausible that rights ought to safely protect people's wellbeing.

We can make appeal to safety by revising the Interest Theory (Canonical) in the following way:

Interest Theory (Safety). For X to have a right against Y that Y Φ , Y 's not Φ -ing must cause X to be worse off than she would have been in at least one close world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .¹³⁷

¹³⁷ The Safety Condition resembles the safety condition on knowledge. Roughly, for an agent, X , to know some contingent proposition, p , X believes p only when p is true in all nearby worlds (Pritchard 2005, 71). More naturally: 'If one knows, one could not easily have been wrong in a similar case' (Williamson 2000, 147; Sosa 1999). My Safety Condition is not structurally identical to the safety condition on knowledge. Stated generally, a safety condition states that X satisfies some condition, C , just in case there is *no* nearby world in which some state of affairs, a , holds. My Safety Condition is of the form, X satisfies some condition, C , only if there *is* a nearby world in which some state of affairs, a , holds. (X satisfies the necessary condition on holding a right only if there is a nearby world in which she is worse off).

For brevity, I refer to this as the Safety Condition. It works by comparing how X fares when Y does not act as the would-be duty dictates with how X would have fared in close worlds in which Y acts as the would-be duty dictates. Call the world in which Y does not act as the would-be duty dictates our *world of evaluation*. Call the world in which Y does act as the duty dictates our *world of comparison*. In *Plane Crash* and *Hitmen*, the world of evaluation is world 1, the actual world. We evaluate how our potential right-holder fares in this world through comparison with close worlds in which the potential duty-bearer acts as the would-be duty dictates (hence the names: world of evaluation and world of comparison). Since Passenger and Victim are worse off in world 1 than they are in world 3, and since the extent to which they are worse off is of sufficient weight to place Attendant and Hitman₁ under their respective duties, Passenger and Victim hold rights. They hold these rights in the actual world.

Strictly speaking, the Safety Condition is only a necessary condition like the Interest Theory (Canonical). So it does not follow that Passenger and Victim hold rights just because the Safety Condition is satisfied. But we are assuming that the other conditions necessary and jointly sufficient for right-holding are satisfied.¹³⁸ So, for ease of exposition, in our cases of harmless wrongdoing, I assume our individual holds a right if the Safety Condition is satisfied.

2.2 Closeness

For the Safety Condition to be satisfied, X must be worse off as a result of Y 's not acting as the duty dictates than she otherwise would have been in at least one close world. It would be useful to say a little about the closeness of worlds. The view of the closeness of worlds one endorses affects the extension of the Safety Condition (those cases in which it is satisfied). In this subsection, I introduce David Lewis's view of closeness. This should

¹³⁸ I explained why this is in footnote in 104 in chapter 4. Like the Interest Theory (Canonical), we could allow "sufficiently weighty" to do a lot of heavy lifting, thereby making the Safety Condition necessary and sufficient (see this chapter, section 2.3).

give us a good idea of how the Safety Condition works without too much technical detail. I do not take a stand on exactly which view of closeness to endorse.¹³⁹

Lewis has us suppose that the worlds under consideration have deterministic laws: these worlds start and play out, following the deterministic laws of nature. We discover the closeness of two worlds at some time, t , by reference to the size of the violations of those laws that would be required at (or close to) t to render those worlds the same after t . The size of miracles is determined by the following four ordered rules:

1. [First,] avoid big, widespread, diverse violations of law.
2. [Second,] maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
3. [Third,] avoid even small, localized, simple violations of law.
4. [Fourth,] it is of little or no importance to secure approximate similarity of particular fact even in matters that concern us greatly.

(Lewis 1986b, 47–48).

Two worlds are close to the extent that there is a small miracle at t in one world, convergence between those worlds before t , and no further miracles after t . In this way, the following counterfactual is true: “If Richard Nixon would have pushed the button [referring to “the button” to launch nuclear missiles], there would have been a nuclear war”. Though worlds in which there is nuclear war are very different from worlds in which there is not, the worlds are identical up until the point at which a miracle occurs to make Nixon push the button at t , and there are no further violations of the laws of nature after t . It would require more miracles to have made Nixon push the button and war not result because, for example, the wiring was shoddy.¹⁴⁰

We can see why world 3 in both *Plane Crash* and *Hitmen* are very close to the actual world. In *Plane Crash*, we would need two miracles to take us from world 3 to world 1: first, a

¹³⁹ The following two paragraphs are very similar to two paragraphs from my BPhil thesis introducing Lewis’s view of closeness.

¹⁴⁰ The Nixon example comes from (Fine 1975, 452). See Lewis’s discussion of his roulette case also (Lewis 1973, 76).

miracle to make Attendant deny Passenger admittance onto the plane; second, a miracle to make the plane crash. Similarly, in *Hitmen*, we would need: first, a miracle to make Hitman₁ not shoot; second, a miracle to make Hitman₂ not shoot. While in both cases, world 2 is closer than world 3 to world 1 (in both cases we need only the first miracle to take us from world 2 to world 1), world 3 is still extremely close to world 1.

It might help to consider an example in which all the other conditions on right-holding are satisfied, but the Safety Condition is not satisfied because the relevant world of comparison is not close enough to the actual world to matter. Suppose that you bump into me on the street. Were you to have taken more care, you would not have bumped into me. And, it would not take too much effort not to bump into me. There is a possible world in which your bumping into me causes me to bleed out (because, let us suppose, our skins are much thinner). Were it to be the case that everyone was susceptible to bleeding out were they to be bumped, I think we would be under duties to take a lot more care in ensuring that we did not bump up against people when walking on the street: it is a little more costly for us to ensure we do not bump up against others, but it would be very costly if this would lead to us bleeding out. (Think of the care we are required to take when driving.) The reason we do not have rights against others that they not bump us, despite our wellbeing being sufficiently weighty to place others under duties were this to be the case, is because the possible world in which bumping into others causes such problems is not sufficiently close to the actual world to matter.

If we look back to the relevant part of the Safety Condition (\mathcal{Y} 's not Φ -ing must cause X to be worse off than she would have been in at least one close world), we see that we need not only need a handle on closeness but also on a world being close *enough*. We turn to this in section 4.1.

2.3 *Where Safety Comes into the Picture*

Recall, from the breakdown of the argument offered in chapter 1, section 3:

1. Assume the Interest Theory (Canonical): For X to have a right against \mathcal{Y} that \mathcal{Y} Φ , X 's wellbeing (her interests) must be of sufficient weight to place \mathcal{Y} under a duty to Φ .
2. Whether or not X 's wellbeing is of sufficient weight to place \mathcal{Y} under a duty to Φ depends partly on the extent to which X is harmed by \mathcal{Y} 's not Φ -ing or benefited by \mathcal{Y} 's Φ -ing.

We could see the Safety Condition as replacing the second premise. Read in this way, it would tell us what it is for one's wellbeing to be of sufficient weight to place others under duties:

- 2*. Whether or not X 's wellbeing is of sufficient weight to place \mathcal{Y} under a duty to Φ depends, at the least, on how X fares across close possible worlds as a result of \mathcal{Y} 's Φ -ing.

However, as I have formulated it, I see the Safety Condition as a revision to the Interest Theory (Canonical). These two ways of reading the Safety Condition are extensionally equivalent; and, the difference in their intension is minimal. But let us go my way and read the Safety Condition as revising Interest Theory (Canonical): like the Interest Theory (Canonical), the Safety Condition itself appeals to how X fares being of sufficient weight to place \mathcal{Y} under a duty; and, our second premise tells us how we work out whether someone's wellbeing is of sufficient weight to place others under duties.

3. In (Partial) Favour of Safety

Above, I introduced the Safety Condition. Whereas the canonical statement of the Interest Theory and the refinements considered in chapter 4 fall foul of the Problem of Harmless Wronging, the Safety Condition correctly generates rights in the cases of preempted harm considered so far. In chapter 8, we see that the Safety Condition also correctly generates rights in cases of pure risk imposition, where the canonical statement of the Interest Theory flounders. What is more, the Safety Condition offers us a unified account of *why* people hold rights across these two different types of harmless wrongdoing: for someone to

hold a right against us that we not perform some action, we look to whether our performing that action could easily leave them sufficiently worse off to place us under a duty.¹⁴¹ The Safety Condition's *principled* extensional accuracy is the primary virtue of the account that I would like to stress in this thesis.

Once we have seen how the Safety Condition solves the Problem of Pure Risk in chapter 8, I offer a full account of why rights respond to modality in the way the Safety Condition prescribes. In the meantime, in subsection 3.1, I offer a sketch of these reasons. In subsection 3.2, I offer two other examples of harmless wrongdoing that the Safety Condition solves; this further extensional accuracy gives us more reason to endorse the Safety Condition.

3.1 *Why Safety?*

One important feature of the Safety Condition is that it removes an objectionable form of luck from right-ascriptions (and, derivatively, from right-violations). If we assume the Interest Theory (Canonical) then, through sheer luck, Attendant does not violate Passenger's right not to be denied admittance onto the plane. In this case, it could easily have been that Attendant *did* harm Passenger and so would have violated her rights. By focusing on more than what happens in the actual world, the Safety Condition removes this objectionable form of luck from right-ascriptions (and right-violations). This gives us good reason to endorse the modal character of the Safety Condition.

Another reason it matters that rights depend on modality in the way the Safety Condition prescribes is that this formally requires that we, as directed duty-bearers, are sensitive to others' wellbeing: that we do not merely not harm others, but that we could not easily have harmed them.¹⁴²

¹⁴¹ In chapter 8, section 2.2, I explain that preemption and pure risk involve symmetrically different types of harmless wrongdoing.

¹⁴² In chapter 8, section 3.3, I examine the relationship between these two reasons (anti-luck and sensitivity to others' wellbeing), as well as expanding on them. In saying that duty-bearers are sensitive to others' wellbeing because of safety, one might think I am confusing my modal notions. In epistemology, the truth of one's belief that p is safe just in case there is no close world in which p is false; whereas, one's belief is sensitive just in case, in the closest world in which p is false, one would not believe p (Sosa 1999). Given this, we can say: T 's Φ -ing *safely* does not harm X just in case there is no close world in which T 's Φ -ing harms X ; and T is *sensitive* to her Φ -ing not harming X just in case, were Φ -ing harmful, T would not Φ . That T safely

Recall, the Interest Theory starts with the idea that others' wellbeing is sufficiently important to place others under duties. You do not need to be an Interest Theorist to believe that. The Interest Theorist is distinctive because they then say, a nice explanation for when and why you owe someone a duty is that their wellbeing is the grounds of the duty. Commonly, it is taken that duties are just a special type of reason: they have something like exclusionary weight, they leave a moral remainder when not acted on, they are demandable, and so on.¹⁴³ The Safety Condition's focus on what could happen or what could otherwise easily have happened makes it that duty-bearers' reasons of this form are more sensitive to others' wellbeing than is the case on the modally undemanding, canonical statement of the Interest Theory. Since the Interest Theory began with the idea that others' wellbeing is very important (it both places us under duties and exclusively makes those duties owed to others), it is plausible to explain modality's importance in this way.

An example may help. Think back to Williamson's example of the ball on the tip of a cone and the ball in the hole. Suppose you have asked two people to put the ball somewhere and keep it still. When all else is equal, the person that puts the ball at the bottom of the hole has taken more care to ensure that the ball is still than the person who has balanced the ball on the top of the cone. Similarly, the person who turns out not to harm others, but could easily have, such as Attendant, has taken less care not to harm others than those who robustly do not harm others. This offers an attractive picture of what we *owe* to others—that we take care not to leave them worse off through our actions than otherwise could have been the case.

Or, at least, so I argue at greater length in chapter 8.

does not harm X does not mean that she is sensitive to her Φ -ing not harming X : we can suppose Y would still Φ even if Φ -ing is harmful, so her not harming X is not necessarily sensitive, but the world in which her Φ -ing is harmful is not sufficiently close, so her not harming X is safe. To be more precise, then, I should say, the Safety Condition makes Y sensitive to that her Φ -ing could not be harmful in sufficiently close worlds. Since I am already departing from epistemology's notion of safety (see note 137), I stick with just saying the Safety Condition makes duty-bearers sensitive to others' wellbeing.

¹⁴³ See chapter 8, note 228.

3.2 Other Cases of Harmless Wronging

In the introduction to this section, I said the Safety Condition's principled extensional accuracy in reply to cases of preempted harm and pure risk is the primary virtue of the Safety Condition that I want to stress in this thesis. Though preemption and pure risk are the focus of this thesis (for reasons explained in chapter 1, section 5, and chapter 8, section 2.2), the Safety Condition helps explain other cases of harmless wrongdoing. This gives us more reason to endorse the Safety Condition.

3.2.1 Suboptimal Benefit

Consider the following case of what we can call suboptimal obligatory benefit.

Surgeon. Surgeon is operating upon Patient. Patient will die without the surgery. Because Surgeon has taken a disliking to Patient, she does not want to help Patient survive. But she knows that her superiors will be suspicious if Patient were to die and so performs the operation, though not to the best of her ability. The operation is successful, though not as successful as it would have been had Surgeon performed the operation to the best of her ability.

Intuitively, Patient has a right against Surgeon that she performs the operation to the best of her ability. Yet, because Patient is no worse off than she would have been had Surgeon not acted as she did, Patient is not harmed by Surgeon. Because of this, Patient's wellbeing is not of sufficient weight to place Surgeon under a duty to perform the operation to the best of her ability.

One might amend the Counterfactual Account of Harm to deal with *Surgeon*. For example, Feinberg says that *A* harms *B* if '*B*'s personal interest [must be] in a worse state than it would be had *A* acted as he should have instead of as he did' (Feinberg 1986, 150). Patient is worse off than she would have been had Surgeon performed the operation as she should have (that is, properly). Given Feinberg's amended condition, Patient is harmed by Surgeon. On the canonical statement of the Interest Theory, Patient can be attributed a right.

Feinberg's condition will not do—it is viciously circular for our purposes. If we do not know whether Patient has a right that Surgeon perform the operation to the best of her ability, we do not know how Surgeon *should* perform the operation. But, if we do not know how Surgeon should perform the operation, we cannot say that Patient is worse off than she would have been had Surgeon performed the operation as she should have. We have lost the harm that Feinberg's condition identifies.¹⁴⁴

While Feinberg's condition will not do, the Safety Condition is satisfied: Patient is worse off in the actual world than she would have been in the close counterfactual world in which Surgeon performs the operation to the best of her ability. Accordingly, Patient can be attributed a right.

3.2.2 Suboptimal Supererogation

Consider the following case of suboptimal supererogation. As the names suggest, suboptimal obligatory benefit and suboptimal supererogation differ in that the benefit is obligatory in suboptimal obligatory benefit. Partly because of this, let me flag that it is more controversial that any wrong (let alone any *wronging*) is done in cases of suboptimal supererogation.

One/Two Arm. Beth is about to lose both of her arms. Ann can save Beth's left arm by pressing Button 1, though at a substantial cost to herself. Ann can save both of Beth's arms by pressing Button 2, though at a trivially greater cost to herself than pressing Button 1.¹⁴⁵

¹⁴⁴ Things might be more complicated if there is some undirected duty that Surgeon is under, a duty whose existence could break the circle (e.g., an undirected duty because of the Hippocratic oath). However, we could use a case without this noise. And, in any case, see the discussion of deontic normality from chapter 4, section 2.2. We might add the following complaint to Feinberg's condition: in a similar way to my objecting to building normality into harm, sure, Surgeon leaves Patient worse off than she *should* have been given Surgeon's undirected duty; yet, it is unclear why this matters to whether she has been *harmed* by Surgeon.

¹⁴⁵ (Pummer 2016; Horton 2017; Pummer 2019; Muñoz forthcoming). For discussions in different contexts, see: (Parfit 1982, 131; 2011b, 225; Kagan 1989, 16; Tadros 2011b, 162). *Cf.* (McMahan 2018; Sinclair 2018), though note these authors object only to the range of cases in which it is wrong to act in the suboptimal way. Though I stick with the following usage for ease of exposition, it might be misleading to say “suboptimal *supererogatory* act” since one might think an act cannot both be supererogatory and wrong.

Suppose that the cost to Ann is substantial enough to make it permissible for Ann *not* to act—Ann’s acting is supererogatory. We might nevertheless think that it would be wrong for Ann to save only Beth’s left arm. So, while it would *not* be wrong for Ann not to act, it *is* wrong for her to act in the suboptimal supererogatory way.

Since the verdict that it is wrong to act in the suboptimal way is somewhat controversial, let me offer two arguments for this claim. One argument runs as follows (Pummer 2016). Perhaps the basis of Ann’s moral option not to act lies in the costs, for Ann, of acting (Kagan 1989; Scheffler 1994). If the only way that you can save a drowning child is by subjecting yourself to some sufficiently large harm, intuitively it is not wrong for you not to save the child. But, if the costs are trivial, intuitively it is wrong for you not to save the child. Appealing to costs in a similar way explains the verdict of our case of suboptimal supererogation above. While there is a sufficiently great cost to ground Ann’s moral option not to act, there is no sufficiently great cost to ground Ann’s having the moral option to save only one, rather than both, of Beth’s arms.

A different argument for why Ann acts wrongly by saving only Beth’s left arm starts from the following idea (Horton 2017): if our actions are not reasonably justifiable to the people whom they might affect, those acts are wrong (Scanlon 1998; Parfit 2003). And, while Ann might be able to reasonably justify her not acting by appealing to the costs, she cannot reasonably appeal to costs to justify her saving *only* one, and not both, of Beth’s arms—the additional costs are too trivial. So, it is wrong to save only Beth’s left arm.

Now, while it has been argued in the literature that it is wrong to act in the suboptimal way, it seems intuitive that, if Ann were to save Beth’s left arm, Ann does not merely act wrongly but wrongs Beth. For example, if Ann is about to save Beth’s arm, Beth may demand of Ann that she not save only her left arm; if Ann saves only Beth’s left, she may demand of Ann an explanation of what Ann did, perhaps even an apology for saving only one arm when she could so easily have saved both. Appealing to the fact that Ann has

(Thanks to Joe Horton for this worry.) Instead, we could say, “the suboptimal act that otherwise would be supererogatory...” However, crucially, even if acting in the suboptimal way is not, strictly speaking, supererogatory, *acting* is still supererogatory in the strict sense, since it is both permissible not to act and permissible to act (namely, it is permissible to act in the optimal way).

wronged Beth straightforwardly makes sense of these demands. But, if we think *T* wrongs *X* only if *T* violates a duty she owes to *X*, this means Ann's duty not to save only Beth's left arm is owed to Beth.

However, this is hard to square on the canonical statement of the Interest Theory.¹⁴⁶ Beth is no worse off than she would have been had Ann not acted as she did—in fact, she is much better off. Since Beth is no worse off than she would have been, Beth's wellbeing is not of sufficient weight to place Ann under a duty not to save only her left arm. So, Beth cannot have a right against Ann that Ann not save only her left arm.

The Safety Condition makes sense of *One/Two Arm*. Beth's wellbeing is sufficiently weighty through comparison of the world in which Ann saves only Beth's left arm with the world in which Ann saves both of Beth's arms to place Ann under a duty not to save only her left arm.¹⁴⁷

(Suppose one does not accept the claim that it is wrong for Ann to save only Beth's left arm. I think one should still think it a virtue of the Safety Condition that it can make sense of the following conditional claim: if it is wrong to act in the suboptimal way, the wrong is directed (due to a directed duty not to act in the suboptimal way). The canonical statement of the Interest Theory cannot make this conditional claim.)

In this subsection, we have seen that the Safety Condition has extensional accuracy in two further cases of harmless wronging in addition to preemption and pure risk. This, along with the Safety Condition's extensional accuracy in reply to the Problem of Preemption, as well as the reasons that rights respond to modality introduced in subsection 3.1, gives

¹⁴⁶ Recall, even if we endorse Weak Correlativity, so think Ann could owe this duty to Beth without Beth holding a correlative right against her, the Interest Theory (Canonical) explains what it is for one to owe a duty to another (chapter 1, section 4.2).

¹⁴⁷ This verdict holds only if the world in which the actor performs the optimal act is sufficiently close to the world in which she performs the suboptimal act. Since these cases of suboptimal supererogation tend to be cases in which the optimal act is minimally costlier than the suboptimal act, this is a plausible assumption. But perhaps there are counterexamples in which the optimal act is performed only in worlds that are not sufficiently close for the Safety Condition. This would mean the Safety Condition does not explain suboptimal supererogation but accidentally gets the right verdict in cases like *One/Two Arm*.

us good reason to endorse the Safety Condition. (At least, good enough reason to stick with me for another few chapters.)

4. In Defence of Safety

In this section, I defend the Safety Condition against three objections. First, that the Safety Condition is too obscure to be of help. Second, that it does not offer a full solution to the Problem of Preemption. Third, that we ought to build Safety into our account of harm instead of rights. I also introduce a fourth problem: that it overgenerates rights. However, the full significance of this problem is apparent only once we see how the Safety Condition deals with the Problem of Pure Risk in chapter 8; so, I delay replying to this problem until it is at its most pointed.

4.1 *The Obscure Problem*

One might object that offering a set of conditions for determining whether a world is close enough for it to satisfy the Safety Condition is a hopeless task. Accordingly, the Safety Condition itself is hopeless. When discussing a similar problem with his view of counterfactuals, Lewis says: ‘It may be said that even if possible worlds are tolerable, still the notion of comparative overall similarity of worlds is hopelessly unclear, and so no fit foundation for the clarification of counterfactuals *or anything else*’ (Lewis 1973, 91, emphasis mine). Since the objection is, roughly, that the notion of closeness is too obscure to play any fundamental role in our theory of rights, let us call this the Obscure Problem.

There are several things to say in reply to this problem. The first is to note the Safety Condition has *principled* extensional accuracy where other solutions to the Problem of Harmless Wronging fail. It accounts for why Passenger has a right against Attendant denying her admittance onto the plane. And, in chapter 8, we see it accounts for why Target has a right against Shooter subjecting her to risk of harm. The modally *undemanding*, canonical version of the Interest Theory does not offer such an account. (And recall, even if one is not wedded to the Interest Theory, I argued harmless wrongdoing is still

problematic.)¹⁴⁸ This gives us evidence that we are closer to having the correct theory of rights with the Safety Condition than without it, even if we cannot offer a robust account of how to determine what worlds count as sufficiently close.¹⁴⁹

Second, the force of the objection itself can be called into question. One might think that delineating the set of close worlds is going to be too obscure because either it is “ill-understood” or it is “vague.” With Lewis, I think talk of the closeness of worlds is not ill-understood, but vague (Lewis 1973, 91–95). However, moral theory is already vague, so vagueness itself should not worry us. For example, most people think it is permissible to break a promise if there is enough good at stake—but how much good is sufficient? Williamson notes how with respect to the safety condition on knowledge that closeness will be vague because *knowledge* itself is vague: ‘If one believes p truly in a case α , one must avoid false belief in other cases sufficiently similar to α in order to count as reliable enough to know p in α . The vagueness in “sufficiently similar” matches the vagueness in “reliable”, and in “know”’ (Williamson 2000, 100). We could posit a similar suggestion here: since whether one’s wellbeing is sufficiently weighty to place others under a duty is underspecified and vague, which worlds count as close enough will be underspecified and vague. But it is not fair to say it is ill-understood. Above in section 2.2., we offered the beginnings of a fairly robust account of closeness.

Third, the Obscure Problem can be mitigated by distinguishing between comparative and quantitative closeness. We have been comparing degrees of closeness between worlds. Of both *Plane Crash* and *Hitmen*, I said that world 2 is closer to world 1 than world 3 is to world 1. But, I said, world 3 is still pretty close to world 1. Implicit in this is some notion of other worlds being less close than world 3 to world 1 (think of the worlds in which Passenger and Hitman₁ do not make it out of bed). However, this need not imply that closeness can be measured in some precise way. As Lewis puts it, we need only a *comparative* and not necessarily a *quantitative* concept of closeness: ‘One world is more similar than another to a third; but we need never say how much more, and the question how much more need

¹⁴⁸ Chapter 1, section 5, and chapter 3, section 4.

¹⁴⁹ In a similar context, Kagan says ‘insofar as it [his *modal personism*] can accommodate many of our deeply held intuitions [...] I think there may well be a great deal to be said for it’ (Kagan 2016, 18).

not make sense' (Lewis 1973, 50).¹⁵⁰ The worry that the Safety Condition is too obscure should be lessened by bearing this in mind—we need not be relying on some perfectly quantifiable account of closeness.

In discussion of modal safety elsewhere in the literature, most approaches to determining which worlds count as sufficiently close focus on identifying limiting cases at either end of the spectrum (identifying worlds that are, and are not, close and pointing to some mechanisms that explain this) (Williamson 2000, 123–28; Pritchard 2005, 145–78; Pettit 2012, 32). We might look to Williamson's remark that we can think of safety in the same way as we think of a child being safe if she is six feet from the edge of a cliff, but unsafe if she is six inches from the edge. Features of context, such as the severity of the harm that would befall the child were the event to occur, will play a role in determining how close a world needs to be in order for it to be too close for the child to be safe. Within the context of freedom (which, below, I suggest is modally demanding), Phillip Pettit suggests that 'this range of worlds is discernible only on an intuitive, context-sensitive basis' (Pettit 2012, 32). Nicolas Southwood goes as far as to say: 'It seems to me that to have an understanding of the relevant [modally demanding] value *just is*, in part, to have an understanding of the relevant range of circumstances' (Southwood 2015, 510).

Perhaps the preceding might only fuel one's scepticism about the Safety Condition. In order to dampen the foregoing worries, let us consider other areas of moral, political, and legal theory in which modal safety is required.

First, consider freedom. Suppose an agent is faced with a set of doors corresponding to their option set (Berlin 1969, xlvi). Perhaps freedom consists of the absence of interference with an agent's preferred option. This view will not do. It is implausible that agents are free in making decisions when, even though the door they choose is open, all other doors might have been locked. It is also objectionable that agents can make themselves

¹⁵⁰ Justin Snedegar offered a similar value-charged example. We might say, "Both £100 and a nice dinner are better than a cup of coffee, though £1,000 is much better." But how much better? It is unclear whether there even is an answer to this question. Even a value-monist, who thinks there *is* an answer to this question, must admit that it is not going to be an easy task figuring that out. Yet, does this obscurity lead us to be sceptical of Snedegar's statement?

(more) free by adapting their preferences in line with those options available to them (Berlin 1969, xxxix). Instead, for an agent to be free perhaps it must be that any door they might have pushed is open. Call this the Non-Interference View (Carter 1999; Kramer 2003). The Non-Interference View requires safety across an agent's option set.

Suppose that the set of doors are all open, yet a third-party has the ability to lock any of the doors. She decides not to. A third view of freedom, the Non-Domination View, says that non-interference (any door being open) is not sufficient for freedom. Instead, '[w]hat freedom ideally requires is not just that the doors be open but that there be no door-keeper who has the power of closing a door' (Pettit 2012, 66). We can see that the Non-Domination View also appeals to safety. Yet, unlike the Non-Interference View, it is not sufficient that we see how agents' options fare across worlds; we also consider additional worlds in which others might be differently disposed towards that agent.

Second, consider the friendship between two individuals, Homer and Barney. Plausibly, both whether Homer and Barney *are* friends and the value of their friendship requires that Homer thinks of and treats Barney in certain ways. He must treat Barney with concern and compassion. He must help Barney, even if that help is quite demanding. However, both their friendship and its value do not merely require that Homer think of and treat Barney in the relevant ways in the actual world. Rather, it requires that Homer would think of and treat Barney in certain ways in other worlds, were things to be different (Pettit 2015, 11–42). For example, both their friendship and its value require that Homer would treat Barney in the relevant ways even if he had less time to invest in their friendship, even if Barney became less funny, and so on. Both the friendship itself and its value requires safety across relevant worlds.

Now, as with the Safety Condition, both freedom (either as Non-Interference or Non-Domination) and friendship do not require safety across *all* worlds. For example, in order for an agent to be free in any morally meaningful sense, we need not require that her choices not be interfered with or dominated across all states of affairs. That would be *too* demanding.¹⁵¹ Similarly, Homer and Barney's friendship does not require that Homer

¹⁵¹ See Pettit: 'this range of worlds [...] does not include all possible worlds' (Pettit 2012, 32).

thinks of and treats Barney in the relevant ways across all circumstances. Homer need not help Barney move a dead body. As with the Safety Condition, we require safety only across some set of close worlds. So, how does this discussion relate to our sceptic who thinks we cannot delineate which worlds count as close enough for the Safety Condition? Well, if one is sceptical of the positive account of closeness developed across subsection 2.2 and the first half of this subsection, one is going to need to rethink many other areas of moral, legal, and political theory.¹⁵² Because of this, I am going to press on with the Safety Condition.

But before doing so, suppose my optimism about the robustness of the account of closeness as well as discovering which worlds are sufficiently close is misplaced. This need not mean modality is playing no role in rights; rather, it could mean only that modality does not help us determine which rights obtain. In the context of replying to a similar challenge for the safety condition on knowledge, Williamson says: ‘In many cases, someone with no idea of what knowledge is would be unable to determine whether safety obtained [...] One may have to decide whether safety obtains by first deciding whether knowledge obtains, rather than vice versa’ (Williamson 2009, 305). Williamson thinks safety offers a circular account of knowledge. Similarly, we could think the Safety Condition is a circular account of rights.

Now, one might think there is a problem with this circular account of determining whether a world is sufficiently close: it means we cannot learn anything about the nature of directed duties and rights by appeal to the Safety Condition.¹⁵³ However, this objection is misplaced. Even if we cannot determine whether one is owed a duty and holds a right by the lights of the Safety Condition alone, we ought still be interested in why it is necessary that whenever one is owed a duty, that duty is grounded in how one fares across close worlds. This is so even if we need a prior understanding of whether one is owed that duty to know whether the world is as close enough.

¹⁵² For other modally demanding goods, see (Raz 1986, 369–99; Pettit 2015; Southwood 2015; Kagan 2016; Lazar 2017).

¹⁵³ Those who are not worried about circularity will not be so worried about this problem (Cruft 2019).

4.2 *Very Preempted Harm*

In *Plane Crash*, the Safety Condition works by comparing the actual world with the close world in which Passenger boards the plane and it does not crash. In *Hitmen*, the Safety Condition works by comparing the actual world in which Hitman₁ shoots Victim with the close world in which Hitman₁ does not shoot Victim and Hitman₂ does not shoot Victim. In both of these cases, we might say, we are looking past the *closest* (counterfactual) world to some other close (counterfactual) world in which the preempting harm does not occur. However, here lies the recipe for a counterexample: make the harm *very* preempted.

Smokin' Aces. Hitman₁ is contracted to kill Victim. Unbeknownst to Hitman₁, there are one hundred other hitmen waiting in the wings. Each is ready to kill Victim if the previous hitman fails.

In *Smokin' Aces*, the Safety Condition has to compare

(SA World 1) Hitman₁ shoots Victim (Hitmen₂₋₁₀₁ were waiting in the wings),

with

(SA World 101) Hitman₁-Hitman₁₀₁ all do not shoot Victim.

Victim is worse off in (SA) world 1 than she is in (SA) world 101—(SA) world 101 is the closest world that allows us to say this. However, one might think (SA) world 101 is *not* a close enough world for the purposes of satisfying the Safety Condition. So, Victim will not have a right that Hitman₁ not shoot her. Like the Pro Tanto and the Normality Thesis, the Safety Condition is not a full solution to the Problem of Preemption.

Why is (SA) world 101 not a particularly close world? Think back to Lewis. We discover the closeness of two worlds at t by reference to the number and size of the violations of the laws of nature that would be required at t to render those worlds convergent after t . On Lewis's view, we would need 100 miracles to get us from (SA) world 101 to (SA) world 1. But perhaps that is too many miracles. (If 100 miracles leaves (SA) world 101 close enough to satisfy the Safety Condition, the objector could always add more hitmen.)

One might object that *Smokin' Aces*, and cases of Very Preempted Harm more generally, are somewhat *gimmicky* and so should be set aside. I am not sure they are. Think back to the example of harm that is normally preempted in the real world, the polluting factories. We supposed that, because of certain factors, these kinds of polluting factories pop up in close proximity to each other. On the supposition that one polluting factory is sufficient to harm on its own, any harm caused by a particular factory would normally be preempted. Salient for our purposes is that *many* of these factories may pop up in close proximity to each other; enough, perhaps, to make a world in which there is no other factory present (to avoid preemption) too far away to satisfy the Safety Condition.

There are at least four ways that we might address to this problem. First, we might relax the criteria for how close a world needs to be in order for it to satisfy the Safety Condition. We can quickly set this reply aside: however expansive we allow closeness to be, our objector could always add one more degree of preemption—she could always add one more hitman. (This would also make the Safety Condition more likely to *overgenerate* rights, an objection I introduce in the following subsection and engage with in chapter 8.)

Second, we might amend the Safety Condition. The element of the Safety Condition that cases of very preempted harm put pressure on is that X must be worse off than she would have been in at least one *close* world. We could revise this closeness element:

Interest Theory (Safety, Relevance Variant). For X to have a right against Y that $Y \Phi$, Y 's not Φ -ing must cause X to be worse off than she would have been in at least one relevant world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .

Since there is no stipulation that the world of comparison needs to be a close world, we need not worry that (SA) world 101 is far away from (SA) world 1. It turns only on whether we think (SA) world 101 is relevant, and it seems relevant.

However, there is an obvious complaint with the Relevance Variant: how do we work out which worlds are relevant? Despite the Relevance Variant faring better with the Problem of Very Preempted Harm than the standard Safety Condition, I am worried about whether there is a non-circular way to answer this question. So, while the Relevance

Variant may score very well extensionally, it has less explanatory power than the Safety Condition.¹⁵⁴

Third, we might lean on our theory of closeness to avoid the verdict that the world in which the harm is not preempted is not close. For example, because the differences between (SA) world 1 through to (SA) world 101 are *themselves* so similar, one could argue (SA) world 101 is *not* actually all that far from (SA) world 1. Call this the *Similarity Strategy*. Taking this strategy, we have not relaxed how close a world needs to be in order for it to be close enough to satisfy the Safety Condition (as suggested and set aside above). Rather, we have suggested that (SA) world 101 is not actually that far away, so will satisfy the Safety Condition. We might even say, how far away these worlds are tends towards some limit no matter how many hitmen we add.¹⁵⁵

There are at least two limitations with the Similarity Strategy. First, I have often motivated the Safety Condition with language such as,

It could easily have been the case that...

Even if the Similarity Strategy is a feature of the correct view of closeness, we might wonder whether it is tracking the locutions that I have been using to motivate safety—if not, the specific semantics we are leaning upon undermines (or at least, is not in keeping with) the motivational story behind the Safety Condition. In *Smokin' Aces*, could it easily have been the case that Hitman₁ shot Victim and no one else would have? I am not sure.

Second, the Similarity Strategy works because the many different preempting acts are relevantly similar, thereby diminishing their impact on how far away worlds are. But maybe this is just a fiat of *Smokin' Aces*. Perhaps one could devise a case in which all of the

¹⁵⁴ The Safety Condition answers this challenge: “Well, the worlds that are relevant are those that are close!” At the end of the preceding section, I said one could offer a circular account of determining which worlds are sufficiently close for the Safety Condition. The circularity of the Relevance Variant appears more problematic than that circularity: the circular account of closeness had an answer to the question why are those worlds relevant.

¹⁵⁵ This would imply the distance between the world with one hitman and the world with two hitmen is a lot greater than the distance between the world with 100 hitmen and the world with 101 hitmen (which, itself, is greater than the distance between the world with 1000 hitmen and 1001 hitmen).

different things that cause the harm to be preempted are of a different character, meaning that the Similarity Strategy will be of little help. For example, we could suppose that were Hitman₁ not to have killed Victim, a boulder would have fallen on her and, had the boulder not fallen, lightning would have struck her, and so on.

We have considered three ways of replying to the Problem of Very Preempted Harm. The first way, of expanding those worlds which count as close enough, did not help much. The second way, of moving to the Relevance Variant of the Safety Condition, weakens the explanatory power of the Safety Condition. And the third way, the Similarity Strategy, was also not wholly satisfying. A final option is to accept our problem's conclusion—when harm is preempted to a great degree, along many different dimensions, people do not have rights against that harm.¹⁵⁶ Let me make a few remarks so this bullet is easier to swallow.

First, above we considered only one feature of closeness (the Similarity Strategy) that helps solve the problem. We might hold out hope that there are other features of the semantics of closeness that will provide a solution to our problem.

Second, and more substantially, in refining the Interest Theory with the Safety Condition, we are attempting to offer a *reductive* account of rights (and directed duties)—an account that explains rights and, correlatively, what it is to owe a duty to another person by appealing to some other feature(s). The Interest Theory (with the Counterfactual Account) began with the idea that rights are difference makers. The Safety Condition adds, “rights are difference makers or *could-easily-have-been* difference makers.” The Problem of Very Preempted Harm is, “Well, what if the putative duty-bearer couldn't easily have been a difference maker?” Perhaps it should not be surprising that we need to accept some counterintuitiveness along this line.¹⁵⁷ When someone really is doomed to suffer some bad fate,

¹⁵⁶ The “along many different dimensions” is included to address the Similarity Strategy.

¹⁵⁷ Steiner says, comparing the Interest Theory to the Will Theory of rights: ‘theories of rights don't come cheap. Buying either one of them involves paying some price in the currency of counter-intuitiveness’ (2000, 298). More generally, consider the paradox of analysis.

perhaps they do not have rights against us that we not make that fate preempted to more of an extent.

What is more, I said that though this thesis is framed in terms of the Interest Theory, the Problem of Harmless Wronging poses a problem for other theories of rights. For example, most theories of rights say that, other things being equal, the stringency of a right corresponds to the harm that would befall its holder were that right not to be respected. What, then, is going to be the stringency of Victim's right against being harmed in *Smokin' Aces*? While the problem does damn the Safety Condition, other views are going to be in trouble too.

Third, though I suggested in chapter 4 that there are real-world examples in which it is normal for harm to be preempted (as well as suggesting in this subsection that there may be realistic cases in which harm may be preempted to a large extent along similar dimensions), we might wonder how prevalent cases are in which harm is preempted to a large extent *along distinct lines*. I did have trouble getting a case going above. This difficulty gives us evidence that the counterintuitiveness may be unfounded.¹⁵⁸

4.3 *Safety and Harm*

In chapter 4, we considered a response to the Problem of Preemption that insisted, intuitively people are harmed when harm is preempted. Any view of harm that says otherwise is mistaken. Once we have an account that arrives at the putatively correct verdict that people are harmed by preempted harm, the canonical statement of the Interest Theory arrives at the correct verdict that people have rights against preempted harm. But I argued we ought not go that way because the other accounts of harm we considered were found wanting. However, I have now suggested the Safety Condition is extensionally accurate (notwithstanding the complications of the Problem of Very Preempted Harm) in generating rights against preempted harm. Why not make a directed appeal to something like

¹⁵⁸ Theron Pummer has suggested climate change on a global scale might be an example. Climate change will cause large amounts of harm that is preempted along many degrees by lots of different types of cause.

the Safety Condition in our account of harm and benefit? This view looks something like this.

Safety Account of Harm and Benefit. Y harms X iff (and because) Y makes X worse off than she would have been in at least one close world in which Y does not act as Y did. Y benefits X iff (and because) Y makes X better off than she would have been in at least one close world in which Y does not act as Y did.

If we amended the Interest Theory (Canonical) with the Safety Condition, we say people are not harmed by preempted harm, though have rights against preempted harm. If we amend the Counterfactual Account with the Safety Account, we say people are harmed by preempted harm, and so have rights against preempted harm (on the canonical statement of the Interest Theory). These views are extensionally equivalent, though intensionally different.

The Safety Account of Harm and Benefit is similar to Tadros's *Complex Counterfactual View*, according to which, 'E harms X only if X is worse off than he would have been in a relevant possible world where E did not occur. Comparison with more than one possible world may be warranted in a single case, yielding different verdicts about harm and benefit' (Tadros 2016a, 177). Key to the Safety Account of Harm and Benefit, as with Tadros's Complex Counterfactual View, is that it is possible for X to be both harmed and benefited by a single act since there can be close worlds in which X is worse off than X would have been had Y not acted as she did as well as close worlds in which X is better off than X would have been had Y not acted as she did.

I would not be too worried if one takes all I say about the Safety Condition and builds it into their account of harm. However, let me offer two reasons why I prefer building safety into rights. Both of these reasons need to be taken on good faith since they reply on arguments not yet presented. However, it is worth flagging the Safety Condition's relevance to harm now.

First, in chapter 6, I argue that merely being subjected to risk is not itself harmful. And then in chapter 8, I show that the Safety Condition is satisfied in cases of pure risk, such as *Roulette*. If we build safety into our account of harm, this means being subjected to risk

is itself harmful. So, it is preferable to build the Safety Condition into our account of rights and say of our example of pure risk imposition, *Roulette*, though Shooter does not harm Target by playing roulette with her while she sleeps, she does violate her rights.

The second reason requires a little more introduction. Whereas the Problem of Very Preempted Harm suggests that the Safety Condition will undergenerate rights, the Safety Condition may also overgenerate rights. The Safety Condition requires only that there is one close world in which the right-holder fares sufficiently better through the duty-bearer acting as the duty requires. But there are many close worlds. Will not the Safety Condition be too easy to satisfy?

Take *Plane Crash*. There is a close world, through comparison with which, Passenger is made worse off by being denied admittance onto the plane. The Safety Condition is satisfied, so Passenger has a right against being denied admittance onto the plane.¹⁵⁹ However, there is also another close world, through comparison with which, Passenger is made *better off* by being denied admittance onto the plane: through comparison of world 1 and world 2. Does this imply that Passenger has a right against Attendant that Attendant deny Passenger admittance onto the plane? It might appear so—the Safety Condition *has* been satisfied. This is one example of the Problem of Overgeneration.

The Problem of Overgeneration gets worse after showing how the Safety Condition solves cases of pure risk. Currently, we are comparing the duty-bearer not acting as the duty requires with close worlds to the world in which the duty-bearer acts as the duty requires. By the end of chapter 8, we see the Safety Condition also requires that we look not only to the closest world in which the duty-bearer does not act as the duty dictates, but to other worlds close to that world. Because the Problem of Overgeneration is not yet at its meanest, I delay addressing it until chapter 8.

Relevant for our purposes at this point is that the Safety Account of Harm and Benefit will overgenerate harms and benefits. (So much is obvious from Tadros's saying: 'Comparison with more than one possible world may be warranted in a single case, yielding

¹⁵⁹ Again, subject to the other necessary and jointly sufficient conditions being satisfied.

different verdicts about harm and benefit'.) One of the ways that I suggest we resist the Problem of Overgeneration is by appealing to the other considerations on rights, including the Doctrine of Double Effect and the Doctrine of Doing and Allowing.¹⁶⁰ While one could try and appeal to these sorts of considerations to avoid the verdict that the Safety Account overgenerates harms and benefits, these agential considerations do not look like the sorts of considerations that affect whether one has been harmed. For example, if one accepts Doctrine of Double Effect, why would whether I intend the harm done to you affect if it is harmful or beneficial? But, it is plausible that it affects whether you hold a right against me. Because of this, it is better to build safety into rights and not harm so that we can resist the Problem of Overgeneration in these ways.

5. Conclusion

This chapter has introduced the Safety Condition as a solution to the Problem of Harmless Wronging through focusing on the Problem of Preemption. In section 3, I gave a taster of the reasons we have to endorse the Safety Condition in addition to its extensional accuracy. And in section 4, I defended the Safety Condition against three objections. First, that the Safety Condition's reliance on the closeness of worlds is too obscure to affect rights. Second, that the Safety Condition will undergenerate rights when harm is very preempted. And, third, that we ought build safety into our account of harm and not rights. In the next chapter, we turn to the Problem of Pure Risk.

¹⁶⁰ Introduced in chapter 3, section 3.3.

Part III: Risk

6. Risk and Harm

1. Introduction

In the previous chapter, I introduced the Safety Condition as a solution to the Problem of Harmless Wronging through focus on the Problem of Preemption. Let us turn to our other more specific version of the Problem of Harmless Wronging, the Problem of Pure Risk. Recall our example of pure risk imposition:

Roulette. Target is asleep. Her housemate, Shooter, comes into her room and plays Russian roulette with her. She pulls the trigger but, luckily, no bullet is fired. Shooter, content with having played a round of roulette, never plays roulette again.

Intuitively, Target has a right against Shooter playing roulette with her—a right that Shooter violates. However, since the risked harm that Shooter subjects Target to does not materialise, Shooter does not harm Target. Since Target is not harmed by the violation of her putative right, her wellbeing is not of sufficient weight to place Shooter under a duty not to play roulette with her. This means the necessary condition set for the ascription of a right on the canonical statement of the Interest Theory is not satisfied. So, Target has no right against Shooter that Shooter not play roulette with her.

There are at least two ways that we might object to the preceding. First, though the risked harm does not materialise in *Roulette*, one might nonetheless argue that Shooter harms Target. If Shooter harms Target merely by subjecting her to risk of harm, Target can have a right against Shooter that she not play roulette with her given the canonical statement of the Interest Theory. Second, though Shooter does not harm Target as things turn out, Shooter did not know this when she acted. If our directed duties and others correlative rights are determined not by the facts, given how things turn out, but instead by the beliefs of potential duty-bearers or the evidence available to them, Target's wellbeing may well be of sufficient weight to place Shooter under a duty not to play roulette with her.

This and the following chapter argue that neither of these ways of proceeding are successful: risk is not itself harmful and our duties and others' rights are determined by the facts. Then, in chapter 8, I show how the Safety Condition solves *Roulette*—though Shooter does

not harm Target, she could easily have; and, the Safety Condition normatively requires not only that we are not harmed, but that we could not easily have been harmed.

Returning to this chapter, some think that Shooter harms Target even though the risked harm does not materialise. They might further think that the moral significance of Target's action lies in this fact.¹⁶¹ This chapter answers two questions. First, is risk of harm itself harmful?¹⁶² (For brevity, is risk harmful?) Second, if risk is harmful, does risk being harmful solve the Problem of Pure Risk? I argue that risk is not harmful. If risk is not harmful, solving the Problem of Pure Risk cannot lie in it being harmful. For reasons that become clear below, some people might not follow me all the way to thinking that risk is *never* harmful. If this is correct, I argue, at the least, risk is harmful only in a small range of cases. If risk is harmful only in a subset of the much larger set of cases in which it looks like we have rights against risk, the solution to the Problem of Pure Risk cannot lie only in it being harmful.

Let me outline the structure of this chapter. We see that there are at least two arguments for the idea that risk is harmful. First (section 2), often people desire not to be subjected to risk. Perhaps people are harmed through being at risk because their desires are frustrated. Second (section 3), by exposing others to risk, one might frustrate their autonomy. Perhaps people are harmed through having their autonomy frustrated. I argue that both views are found wanting. In section 4, I argue against the view that risk is harmful, however the view is worked out.

Before beginning, a preliminary is required. In *Roulette*, the following morning, if Shooter were to explain to Target what happened, Target might reasonably become psychologically distressed. She might fear that Shooter will, again, break into her room and play roulette with her (despite Target's assurances to the contrary). She might buy a lock for her door, spend time at friends' houses, or even move. In all of these cases, Target might

¹⁶¹ (Finkelstein 2003; Adler 2003; Oberdiek 2009; 2012; 2017a; Lazar 2015; 2017; Placani 2017). With the exception of, perhaps, Lazar, all seem to imply that the moral significance of risk lies in its being harmful.

¹⁶² Among other reasons, determining whether risk is harmful is also of significance because the law needs an answer to this question. The paradigm case is *Hotson v East Berkshire Area Health Authority* [1987] AC 750 (HL). For discussion, see (Perry 1995, 2014, 45–48; Finkelstein 2003, 975–90).

be harmed by Shooter. But these harms (the psychological distress, fear, and disruption) are *downstream* from the risk imposition and, as *Roulette* makes salient, contingent. While of moral significance, they are not the subject of our discussion.

2. The Desire Solution

Often, people desire not to be subjected to risk of harm. Desires might play some role in harm. Perhaps people are harmed through being at risk because their desires are frustrated. Perhaps Target is harmed because she has a desire not to be the subject of risk, a desire that Shooter frustrates. Let us call this *the Desire Solution*.

There are several ways that the Desire Solution might work, depending on the account of wellbeing that is correct. Assume the Desire Theory of wellbeing, on which wellbeing consists only in the satisfaction of one's desires. In *Roulette*, if Target has a desire that Shooter not subject her to risk, Target is harmed by Shooter. This is because Shooter makes Target worse off than Target would have been, had she not played Russian roulette with Target. She makes Target worse off than Target would have been because she causes Target's desires to be more frustrated than they would have been, had she not acted as she did. This is one version of the Desire Solution.

The Desire *Solution* is different from the Desire *Theory*. The Desire Theory is a view about wellbeing. The Desire Solution is an answer to the question, "Is risk harmful?" One need not hold the Desire Theory to defend the Desire Solution. There are other views of wellbeing in addition to the Desire Theory that see the frustration of one's desires as bad for an individual (though, unlike the Desire Theory, they do not imply that it is the only thing that is bad for an individual). (For example, Hybrid Theories that subsume the Desire Theory will think the frustration of one's desires are bad for one.) On these views, risk will be harmful in virtue of one's desires not to be at risk being frustrated. Let us assume the Desire Theory. The objections considered below applies equally to these other views of wellbeing.

2.1 Actual and Informed Desires

Some try to force a dilemma upon the defender of the Desire Solution, both horns of which they take to be problematic (Adler 2003, 1251–53; Perry 2007, 200–201; Oberdiek 2012, 346–47). In this subsection I argue this dilemma can be dissolved. In the following subsection, I introduce a problem that is not so easy to avoid.

The Desire Theory says that what is good for an individual is the satisfaction of their desires. We might ask whether the desires that are relevant for the Desire Theory are an individual's *actual* or *informed* desires.¹⁶³ In *Cherry Pie*, suppose that I want some cherry pie and there is one in front of me. Unbeknownst to me, I have developed a severe allergy to cherries (Heathwood 2016, 139). If the Desire Theory cares about my actual desires (call this version of the Desire Theory the Actual Desire Theory), my life goes better, for me, if I get the pie than if I do not. This verdict seems odd. My life seems to go worse for me if I get the pie. (Below, we turn to a more sympathetic reading of this case.) More generally, as James Griffin writes, 'notoriously, we mistake our own interest' (Griffin 1986, 10).

In *Cherry Pie*, were I to be presented with all of the non-evaluative facts, I probably would not continue to desire a slice of pie. This might give us reason to think that the desires that are relevant for the Desire Theory are our *informed* desires. Call this the Informed Desire Theory.

Above, I said that some people try to force a dilemma upon defenders of the Desire Solution. On the first horn of the dilemma, assume we hold the Actual Desire Theory. Stephen Perry suggests, if one holds the Actual Desire Theory, 'people can prefer or disprefer almost anything, so there would seem to be nothing in principle to prevent someone from preferring a risky to a non-risky state of affairs' (Perry 2007, 200). I take the worry to be that this means risk is not necessarily harmful because someone may not desire not to be at risk in some particular circumstances. John Oberdiek rejects this horn of the dilemma because he thinks that the Actual Desire Theory is too implausible in general. He writes,

¹⁶³ E.g., (Sidgwick 1981, 109–15; Rawls 1999, 365–72; Griffin 1986, 10–20). By actual desires, I mean something closer to an individual's *hypothetical* desires—the desire she would have, were someone to ask her. See Sumner on the move from *revealed* to *hypothetical* desires (Sumner 1996, 119).

‘we often prefer, through ignorance or whatnot, what is in fact bad for us [and this] is decisive’ (Oberdiek 2012, 347).

On the second horn of the dilemma, assume we hold the Informed Desire Theory. That theory says that what is good for an individual is having the desires satisfied that she would have were she in possession of all the non-evaluative facts. But, it is pressed, from this ‘omniscient perspective’ of fully informed desires, ‘risk just disappears’ (Oberdiek 2012, 346). The idea is supposed to be, if one knows that some risk will materialise, there is no *risk* of harm, but harm in the straightforward sense; and if, on the contrary, one knows that the risk will *not* materialise, it will not be the case that one does not desire to be subjected to the putative risk. On either horn of the dilemma, then, the Desire Solution fails.

The Desire Solution can be defended on both fronts. Let us begin with the first horn of the dilemma. In response to Perry, suppose that it is not the case that someone desires that they not be subjected to some particular risk. Perry is correct in thinking that this means that *that* risk is not harmful to that particular person in those particular circumstances. But this is not so much an objection to the Desire Solution as statement of the Actual Desire Theory itself.¹⁶⁴ This fiat of the Desire Theory results from its *subjective* character as a theory of wellbeing: on the Desire Theory, ‘getting a good life has to do with one’s attitudes towards what one gets in life rather than the nature of those things themselves’ (Heathwood 2016, 135). Many people endorse the Desire Theory *because of* its subjective character.

What of Oberdiek’s rejection of the Actual Desire Theory because of its implausibility in reply to cases like *Cherry Pie*? We might not need to embrace the Informed Desire Theory to get the intuitively correct verdict in *Cherry Pie*. This is because the Actual Desire Theory is ambiguous between saying that it is good for me, all-things-considered, to eat the pie and that it is good for me, pro tanto, to eat the pie. If we embrace the second disambiguation, the Actual Desire Theory can maintain that it is good for me, pro tanto, to eat the cherry pie, but that it is not good for me all-things-considered. It might not be good for

¹⁶⁴ It is also the case that one may desire a risky state of affairs over a non-risky state of affairs, fully informed, so Perry’s objection applies equally to the Informed Desire Theory.

me all-things-considered because doing so will frustrate many of my other (pro tanto) desires. In favour of this line of thinking, compare the following two states of affairs: (S1) I do not have a desire to eat the cherry pie, eat it nonetheless, and have an allergic reaction; (S2) I do have a desire to eat the cherry pie, eat it, and have an allergic reaction. My life goes worse for me in S1 than S2. The pro tanto disambiguation can explain this.¹⁶⁵

Given the preceding two paragraphs, the Actual Desire Theory horn of the dilemma does not seem as worrying as some have thought. Let us move on to consider the second horn of the dilemma. One might respond to the second horn of the dilemma by suggesting that there is a distinction between, on the one hand, being fully informed of all of the non-evaluative facts in relation to the circumstances of a decision before deciding, and, on the other hand, being fully informed in relation to the outcome of the decision. We might put the point in the following way: agents can be more (or fully) informed about their choices from an ex ante perspective (before the fact) without becoming fully informed about their choices from an ex post perspective (after the fact, in knowledge of whether the risk has, or has not, materialised). And, risk disappears in the way that Oberdiek suggests only on the second disambiguation of informed desires.¹⁶⁶

For example, let us amend *Cherry Pie*. As above, suppose I want a slice of cherry pie and there is one in front of me. Unbeknownst to me, I have developed a severe allergy to

¹⁶⁵ See (Sumner 1996, 131; Heathwood 2005). The purpose of this paragraph has been to show that the Actual Desire Theory is not simply to be set aside. That said, I am sceptical of whether this pro tanto disambiguation will work in all cases, for similar reasons as those that lead me to be sceptical of the Pro Tanto Thesis as a full solution to the Problem of Preemption (chapter 4, section 2.1). The pro tanto disambiguation provides the intuitively correct verdict in *Cherry Pie* because the aspects of wellbeing at stake are fairly distinct—(i) enjoyment from eating pie versus (ii) pain from an allergic reaction. However, how plausible is this solution when the pro tanto aspects of one’s wellbeing that play off against each other in the all-things-considered judgement are more similar? Suppose that I desire state of affairs S3 over S4 because I think that S4 will be painful and S3 not painful. I am mistaken. S3 is painful and S4 would not have been. What pro tanto desire is satisfied by my choosing S3? What is more, the Informed Desire Theory *can* recognise the putative advantage of the pro tanto disambiguation of the Actual Desire Theory noted directly above in the text. The Informed Desire Theory can say that, even if I knew all of the facts, I would rather have an allergic reaction from a pie that I enjoyed than an allergic reaction from a pie that I did not enjoy (that is to say, fully informed, I would desire S2 over S1).

¹⁶⁶ This will need to be slightly refined for one might think, if determinism is true *and* one knew all of the ex ante facts, they will be able to figure out all of the ex post facts. Thanks to Antony Duff and Sandra Marshall for this point. We could go for a Reasonably Informed Ex Ante Desire Theory.

cherries. Unlike above, suppose that this allergy results in an allergic reaction fifty percent of the time. When the allergy does not manifest, I have no adverse reaction. The Informed Desire Theory might say that what is good for me should be decided from the informed *ex ante* perspective (given knowledge of my allergy and the propensity of its occurrence) *or* from the *ex post* perspective (given both knowledge of my allergy *and* knowledge of whether, in this instance, I will have an allergic reaction to the pie). Risk does not disappear on the *ex ante* perspective of the Informed Desire Theory.

We have been considering the following dilemma that threatens the Desire Solution. If one endorses the Actual Desire Theory, agents may desire a risky state of affairs over a non-risky state of affairs and so risk will not necessarily be harmful. However, this does not so much seem to be an objection to the Actual Desire Theory rather than a fiat of the Desire Theory's subjective character as a theory of wellbeing. If one endorses the Informed Desire Theory, one might think that risk just disappears insofar as the agent will be informed of whether or not the risk will manifest itself, and so would only desire not to be subject to risks that will materialise. While this might be true on the Informed Desire Theory (*Ex Post Perspective*), this is not true on the Informed Desire Theory (*Ex Ante Perspective*). Risk might itself be harmful on the Actual Desire Theory or on the Informed Desire Theory (*Ex Ante Perspective*).

2.2 *Derivative and Non-Derivative Desires*

We are trying to argue that risk of harm is itself harmful. If risk is itself harmful, we can easily accommodate why we have rights against risk of harm. Since it might be bad for us when our desires are frustrated, and we have desires not to be at risk of harm, risk might be harmful when our desires not to be at risk are frustrated.

Recall again that the Desire Theory says that what is good for an individual is the fulfilment of their desires. There are several choice-points for the Desire Theorist. We just considered whether the desires that are relevant for the Desire Theory are individuals' *actual* or *informed* desires. The choice-point that is important for our purposes in this subsection is whether the desires that are relevant are both individuals' *derivative* and *non-*

derivative desires or only their *non-derivative* desires.¹⁶⁷ Let us say a person's desire for something is non-derivative if she desires it for its own sake, and her desire for something is derivative if she desires it, but not for its own sake. On what we can call the Non-Derivative Desire Theory, what is good for an individual is only the fulfilment of their non-derivative desires. On the Unrestricted Desire Theory, what is good for an individual is both the fulfilment of their derivative and non-derivative desires.

Suppose that Dan wants it to snow in the mountains so that he can get some good skiing in. It does snow but, as it turns out, Dan does not end up going skiing (Heathwood 2016, 139). Does the satisfaction of Dan's desire that it snow in the mountains make his life go better for him, even though he does not end up going skiing? Dan's desire that it snow was only a derivative desire. The satisfaction of Dan's desire that it snows seems only to derive its value from his desire to get some good skiing in. Dan's desire that he get some good skiing in is a non-derivative desire (or, at the least, closer to a non-derivative desire). The Non-Derivative Desire Theory says Dan's life does not go better for him. The Unrestricted Desire Theory says his life does go better for him.

We have good reason to endorse the Non-Derivative Desire Theory. First, the Unrestricted Desire Theory is intuitively implausible. As Chris Heathwood suggests, 'intuitively, the fulfillment of [Dan's] desire that it snow was not in the end of any benefit to [Dan]' (2016, 139).

Second, compare the following two worlds.¹⁶⁸ In both worlds, I want to eat an apple. In both worlds, one apple remains on a tree. In world 1, I need only to reach out to get the apple. In world 2, I need a ladder to reach the apple. To satisfy my desire to get the apple, I will also desire a ladder, though this desire will be only derivative. Luckily, there will be a ladder nearby. If the Unrestricted Desire Theory is correct, my life goes better for me in world 2 than it does in world 1. This is because, in both worlds, I get to satisfy my desire to eat an apple. But in world 2, I get to satisfy an *additional* desire—my desire for a ladder.

¹⁶⁷ (Sidgwick 1981, 109; Heathwood 2005; 2016, 139). See (Parfit 1984, 117; 2011a, 58–59) for relevant discussion concerning the Desire Theory of *reasons*.

¹⁶⁸ The case is inspired by (Parfit 2011a, 59).

But it is implausible that my life goes better for me in world 2 than world 1 in virtue of the satisfaction of the additional derivative desire.

So, we have good reason to prefer the Non-Derivative Desire Theory. Returning to risk, the Desire Solution works in the following way: if the Desire Theory (of wellbeing) is correct, and if Target has a desire that Shooter not subject her to risk, Shooter harms Target. She does so by frustrating Target's desire not to be at risk. And, relating this to the broader problem under consideration in the thesis, the Problem of Pure Risk, if Shooter harms Target in *Roulette*, we can easily accommodate why Target has a right against Shooter subjecting her to risk of harm. However, Target's desire that she not be at risk is a derivative desire—it is derivative upon, at the least, her non-derivative desire that she not be harmed. And, if we ought to endorse the Non-Derivative Desire Theory, Target is not made worse off by the frustration of her desire not to be at risk. But this means that Target is not harmed by the frustration of that desire. So, the Desire Solution fails.

The preceding argument rests upon the claim that Target's desire that Shooter not subject her to risk is merely a derivative desire. This seems fairly obvious to me—why would one desire not to be at risk of harm *for its own sake*? But perhaps it is not as obvious to everyone.

One way to support this claim is to consider a case in which one holds the desire not to be at risk of harm, though not the desire not to be harmed. Suppose that Bloggs is indifferent between being harmed to some degree and not being harmed. Suppose also that Bloggs's indifference is rational—suppose, were she to be harmed, she would be compensated to the level that would leave her indifferent between not being harmed, and being harmed and compensated. Finally suppose that despite this indifference, Bloggs desires that she not be subject to risk of harm. Once we factor away any non-derivative reasons for which she might hold this desire not to be at risk, the desire looks very peculiar. (It is not that Bloggs desires not to be at risk because being so will make her anxious or on edge. The desire also cannot be explained away with recourse to Bloggs's being risk-adverse for her risk aversion will already be accommodated when determining the compensation owed.) The explanation for why this desire looks peculiar, I submit, is that it is a derivative desire without a non-derivative desire to latch onto.

One might object that showing how peculiar the desire not to be at risk would have to be in order for it to be non-derivative, as I have done, does *not* show that having a non-derivative desire not to be at risk is not possible. Rather, it shows only that it is peculiar. And, because of this, risk *can* sometimes be itself harmful given the Desire Solution—namely, when someone has a non-derivative desire not to be at risk of harm.

Perhaps this objection is onto something. However, I am inclined to think that these rare cases may be captured by other restrictions not discussed here on which desires get to count for the Desire Theory (for example, perhaps desires need to be sufficiently rational for their satisfaction or frustration to count towards one's wellbeing, and perhaps non-derivative desires not to be subject to risk are not sufficiently rational). But I do not develop this here. Instead, note that this objection shows only that the Desire Solution would work in these extremely peculiar, rare cases in which someone has the desire not to be at risk for a non-derivative reason. And so, even if this objection is correct, the solution to the Problem of Pure Risk cannot lie in it being harmful given the Desire Solution—this would mean we have rights against risk only in these rare cases.

I have argued that the Desire Solution fails because the desire that we not be at risk of harm is a derivative desire (derivative upon our non-derivative desire not to be harmed); and, the frustration of derivative desires does not *itself* constitute a diminishment to one's wellbeing and, thereby, a harm. In section 4, we consider an argument against any view on which risk itself is harmful; this objection tells against the Desire Solution.

3. The Autonomy Solution

3.1 Autonomy and Harm

There is another way that risk might be harmful. One reason why we might think the satisfaction of Dan's desire that it snows in the mountains does promote his wellbeing is that it gives him the *option* of getting some good skiing in. And, this is true even if he does not end up going skiing. More generally, we might think having the option to do things is valuable to us, even if we do not end up doing those things.

One reason why one may think this is if one believes having options is partly constitutive of autonomy. In particular, one might endorse:

Adequate Range. For an individual's choice to be autonomous, she must have an adequate range of valuable options (Raz 1986, 372).

Oberdiek motivates Adequate Range by saying that 'one is autonomous when one can plot and pursue one's own worthwhile path, and to do this, one needs to have access to a range of valuable options' (Oberdiek 2017a, 9). We return to Adequate Range below (section 3.2.1).

Returning to risk, suppose that autonomy plays some role in wellbeing—one's life goes better or worse, for them, when their autonomy is promoted or frustrated. Consider:

Two Doors. Chooser is faced with a choice. Let choosing between two doors, *A* and *B*, stand in for an autonomous choice between two different valuable things. Unbeknown to Chooser, Locker locks door *B* (he stops her from being able to do whatever valuable thing door *B* takes the place of). Chooser chooses door *A*.

Locker risks harm to Chooser. Straightforwardly, he risks it being the case that Chooser chooses door *B* and is unable to do whatever valuable thing door *B* takes the place of. If Adequate Range is correct, Locker frustrates Chooser's autonomy, even if Chooser is unaware of this. Locker frustrates Chooser's autonomy because he stops her from having an adequate range of valuable options to choose from. Because he frustrates Chooser's autonomy and because, we are assuming, Chooser's autonomy is partly constitutive of her wellbeing, he harms Chooser. More generally,

if *Y* subjects *X* to risk and, thereby, frustrates *X*'s autonomy, *Y* harms *X*.

Let us call this the *Autonomy Solution*. Given the Autonomy Solution, risk is sometimes harmful.¹⁶⁹ Given that risk is sometimes harmful, we can have rights against risk on the Interest Theory (Canonical).

The Autonomy Solution has support in the literature. Oberdiek suggests that, ‘while imposing risk does not involve material harm [...] it can nevertheless constitute a setback to a nonmaterial autonomy interest of a certain kind’ (Oberdiek 2012, 342). By subjecting someone to risk,

it effectively attaches sanctions to or normatively forecloses certain options that would otherwise be available to the individual, thereby narrowing the risked person’s set of worthwhile opportunities. Narrowing one’s open future diminishes one’s autonomy suitably understood, and it is in this that the moral significance and thus the potential impermissibility of pure risking lies. (Oberdiek 2012, 351–52)

Oberdiek offers us the following analogy. By laying a trap, though I may not experientially affect anyone, I do nonexperientially affect people: ‘This is because the trap takes away the option, or more accurately renders unacceptable the exercise of the option, of stepping where the trap has been set’ (Oberdiek 2012, 352). If enough traps were laid, one’s autonomy would be completely frustrated.

Seth Lazar holds a somewhat similar view. He says, ‘if others avoidably make us dependant on [luck] for our avoidance of wrongful harm, they harm us’ and that, when one subjects another to risk, this ‘contravenes [an] important interest [...] in *being secure*’ (Lazar 2015, 102, 2017, 7). He thinks this for at least two reasons: first, being secure serves ‘contingent benefits such as peace of mind, or the ability to plan for the future’; second, ‘the more you depend on luck, the less control you have over your life, and so the less autonomy you have’ (Lazar 2017, 8; 2015, 102).¹⁷⁰ The first of these reasons does not concern

¹⁶⁹ Ben Colburn has pointed out that it is possible that one can remove options from others but not make it the case that Adequate Range is not satisfied. This is possible if the individual whose options have been limited still has *an* adequate range of sufficiently valuable and diverse options. In cases of this sort, exposing others to risk will not be harmful. The defender of the Autonomy Solution could reply by revising Adequate Range and saying that, whenever one’s options are diminished, one’s autonomy is frustrated to *some* extent.

¹⁷⁰ See note 174 for a third reason.

pure risk, so can be set aside. The second reason appeals to autonomy, and so makes Lazar vulnerable to the objections raised against the Autonomy Solution below.¹⁷¹

Maria Ferretti holds another similar view. She thinks, ‘what is specifically morally problematic about risk’ consists in not ‘respecting people as moral agents’, which she cashes out in terms of a diminishment of overall freedom (Ferretti 2016, 262). While this account is not directly answering the questions of whether risk is harmful, it does fall foul of problems similar to those discussed below (sections 3.2.2-3).

3.2 Against the Autonomy Solution

This subsection raises three objections against the Autonomy Solution.

3.2.1 Theoretical Commitments

The Autonomy Solution relies upon substantive, controversial theoretical commitments concerning autonomy. First, one needs to accept Adequate Range. Second, one needs to think that autonomy is non-derivatively valuable. Let us take these components in turn.

Adequate Range says that, in order for one to enjoy an autonomous choice, one needs to have chosen from an adequate range of valuable options. Some views of autonomy deny this. On these views, whether a choice is autonomous depends only on how the decision was arrived at. Here is a toy view of that sort.

Serena Olsaretti thinks that a choice is *not* voluntary ‘if it was made because no other acceptable alternative was available’ (Olsaretti 1998, 54). Olsaretti continues that we should distinguish between ‘first-order desires and wishes, which is what we focus on when considering the voluntariness of actions, and autonomy as the second-order capacity to reflect critically over one’s first-order preferences and desires, and to decide which ones to act on’ (Olsaretti 1998, 73). This view of autonomy would say that a choice is autonomous only if, first, it was not made because no other acceptable alternative was available

¹⁷¹ I end up explaining risk’s moral significance by appeal to something close to security (in my terms, *safety*). However, I do not think that safety is itself partly constitutive of wellbeing. See chapter 5, section 4.3. Further, having security as partly constitutive of wellbeing makes the view vulnerable to the Magnitude of Harm Problem raised below (section 4).

and, second, because it was consistent with a second-order desire (that itself does not exist because no other acceptable alternatives are available).¹⁷² Given this, one does not actually need an adequate range of valuable options to choose from. Rather, one needs, first, not to have chosen because no other acceptable option was available and, second, one's choice needs to be consistent with one's second-order desires. There is no mention of one having an adequate range of options to choose from.¹⁷³ So, it is not obvious that everyone will endorse Adequate Range.

Now for the second restriction to the Autonomy Solution. Some people think that autonomy is *non-derivatively* valuable. These people think that autonomy is valuable in and of itself. Others think that autonomy is only *derivatively* valuable. Suppose that one is a Hedonist. It is hard to know what will maximise pleasure for others. Since individuals are best placed to work this out, one might value autonomy. But, on this view, autonomy's value is derivative upon hedonic value. It is consistent with this view that an individual's autonomy is frustrated and yet their wellbeing promoted—if the frustration of autonomy brings about more pleasure than would otherwise have been the case. On this view, the frustration of that individual's autonomy would not, in these circumstances, be harmful.

Returning to risk, if one thinks that autonomy is only derivatively valuable, the Autonomy Solution fails. This is because, in cases of pure risk imposition, we are supposing that the risk *does not* materialise. But this means that the frustration of the individual's autonomy does not affect the non-derivative value on which the value of autonomy depends.

The objection raised in this subsection is limited in scope (it was meant only to point out, rather than undermine, the two theoretical commitments) but raises two issues. First, the Autonomy Solution can be appealed to only by those who accept Adequate Range and think autonomy is non-derivatively valuable. Second, we might think that, other things being equal, an answer to the question “Is risk harmful?” is unattractive to the extent that

¹⁷² I say “only if” and not “if, and only if,” because there might be other necessary conditions on autonomy.

¹⁷³ While lacking an adequate range of valuable options will often stop one from making a choice because no other option was available, this is contingent. And, because we are considering *pure risk*, where one is unaware that one's options are limited, one's autonomy will not be frustrated in this way.

it relies upon controversial normative commitments—the Autonomy Solution relies upon two such commitments.

3.2.2 *Non-Autonomous Individuals*

The Autonomy Solution implies that risk is harmful only for people who are capable, at that time, of leading autonomous lives. This is because it is only those people who are capable of having their autonomy frustrated *in virtue of* having valuable options removed. This precludes, among others, agents with undeveloped, compromised, or damaged rational capacities (for example, very young children, the severely mentally disabled, and those suffering from the later stages of Alzheimer’s disease). We can draw two conclusions from this. First, the Autonomy Solution does not say that risk is harmful for those without autonomy. If the solution to the Problem of Pure Risk lies in risk’s being harmful, this means those without a capacity for autonomy do not have rights against risk of harm. So, the Autonomy Solution does not provide us with a full solution to the Problem of Pure Risk.

Second, we can posit the following argument. The Autonomy Solution leaves us with an asymmetry: exposing autonomous individuals to risk is itself harmful, whereas exposing agents with undeveloped, compromised, or damaged rational capacities to risk is not itself harmful. This asymmetry *itself* is intuitively implausible. This suggests the Autonomy Solution is implausible.

Let me say a little more about why the Autonomy Solution implies risk is not itself harmful for those without a capacity for autonomy. Suppose that Villain kills someone with severe mental disabilities. Villain does not frustrate this person’s autonomy for they have no autonomy for the Villain to frustrate. *A fortiori*, Villain does not frustrate this person’s autonomy by exposing them to *risk* of death. So, subjecting this person to risk is not itself harmful (given the Autonomy Solution).

The same holds if Villain subjects a baby to risk of death. If Villain were to kill the baby, we would not say she frustrates the baby’s autonomy (for, again, the baby has no autonomy for Villain to frustrate)—and, we certainly would not say that she frustrates the baby’s autonomy in virtue of denying her a valuable range of options to choose from. So, we should not say Villain frustrates the baby’s autonomy by exposing her to risk of death.

What the Villain may do were she to kill the baby is deny the baby the *opportunity* of having an autonomous life. However, if the risked harm does not materialise and so the baby lives, Villain does not deny the baby this opportunity. So, she has not interfered with the autonomous life the baby will *later* come to have. This is related to the objection raised in the following subsection.

3.2.3 Roulette and Autonomy

The Autonomy Solution works by showing that, when one subjects another to risk, one may frustrate the other's autonomy by stopping them from having an adequate range of valuable options to choose from, independently of whether or not the risk materialises. In *Two Doors*, Locker frustrates Chooser's autonomy by removing the option of going through door *B*. Thereby, he stops Chooser from having an adequate range of valuable options. But this means that the Autonomy Solution applies only to cases in which the risk affects the exercise of a potentially autonomous choice—and it applies only in virtue of stopping *that* choice from being autonomous. This analysis does not straightforwardly extend to *Roulette*. Target is asleep. She is not exercising any autonomous choices at the time at which the risk is imposed. And so, it is not obvious that Shooter's subjecting her to risk removes any valuable options from her in a way that undermines her autonomy. We can ask, what valuable option does Shooter remove? So, it is not obvious that Autonomy Solution solves the Problem of Pure Risk.

Now, Shooter risks it being the case that Target does not have “any future choice”, autonomous or otherwise. But that itself does not *frustrate* Target's autonomy—it merely *risks* frustrating her autonomy. To see this, consider the following case.

Two Doors (Two Choices). Chooser is faced with a choice. Let choosing between two doors, *A* and *B*, stand in for a potentially autonomous choice. After going through either door, Chooser will then be faced with another choice. Let choosing between A_1 and A_2 stand in for a potentially autonomous choice that she would face if she went through door *A*. Let choosing between B_1 and B_2 stand in for a potentially autonomous choice that she would face if she went through door *B*. Unbeknownst to Chooser, Locker locks door *B*. Chooser chooses door *A*.

Suppose that Chooser is now choosing between A_1 and A_2 . Is *that* choice autonomous? It seems so. Before going through the door, Chooser faced two potentially autonomous choices: Choice 1, between doors A and B , and Choice 2, between A_1 and A_2 or between B_1 and B_2 , depending on which door she went through. If Adequate Range is correct, Locker stops Choice 1 from being autonomous—he does so by stopping Chooser from having an adequate range of valuable options. However, he does not stop Choice 2 from being autonomous. He merely *risks* Chooser not being able to make that Choice. Similarly, in *Roulette*, though Shooter might stop Target’s choice of going to sleep from being autonomous for example, she does not stop any of Target’s future choices from being autonomous.

In support of the verdict that Choice 2 was autonomous, despite Choice 1 not being autonomous: if Choice 2 is rendered non-autonomous in virtue of Choice 1 not being autonomous, this implies a single non-autonomous choice can taint all future choices, rendering them non-autonomous. But this is implausible.

3.3 *Moving Beyond Autonomy*

The previous objection leads to a more general one. The Autonomy Solution says that risk is harmful *because*, when one is exposed to risk, one’s autonomy is frustrated. Accepting for the sake of argument that risk *is* harmful, is autonomy the correct sort of explanation *why* risk would itself be harmful? I am not convinced.

Let us conclude our discussion of the Autonomy Solution. We have seen that the Autonomy Solution can be appealed to only by those who accept both Adequate Range and that autonomy is non-derivatively valuable. These are both fairly controversial theses. More substantively, the Autonomy Solution leaves us with several asymmetries—some cases of risk are harmful, and some cases of risk are not harmful. But it is odd to think that exposing adults with autonomy to risk is (in some circumstances) harmful, while exposing individuals with undeveloped, compromised, or damaged rational capacities to risk is not harmful. It is also odd that Locker exposing Chooser to risk is harmful in *Two Doors* but that Shooter’s exposing Target to risk in *Roulette* is not harmful. I have also just suggested that explaining why risk is itself harmful by reference to autonomy does not seem wholly satisfying. In the following section, I argue that any view on which risk is harmful leads to

some intuitively implausible results concerning the magnitudes of harm. All of this gives us good reason to think that the Autonomy Solution fails to show that risk is harmful.

Some might object to the preceding along the following lines. First, they say, “While you are undecided on our two controversial theoretical commitments, I think that Adequate Range is correct, and that autonomy is non-derivatively valuable.” Second, they might object that the second and third objections do not necessarily impugn the Autonomy Solution as an answer to the question “Is risk itself harmful?” This is because these objections show only that there are *some* cases in which risk is not harmful. But this does not show that risk is not harmful in all cases. For those who say this, we can note a conservative conclusion: the Autonomy Solution does not provide us with a full solution to the Problem of Pure Risk.

4. The Magnitude of Harm Problem

We have seen that the Desire Solution fails. I have also argued that the Autonomy Solution fails (though I noted the more conservative conclusion that even if, against what I have argued, the Autonomy Solution says risk is harmful *in some cases*, it is too restrictive to offer us a full solution to the Problem of Pure Risk). But perhaps there might be other ways to claim that risk is harmful.¹⁷⁴ This section raises a general worry with the very idea that risk is harmful.

¹⁷⁴ Here are three further views not focused on in detail. (Placani 2017) thinks, in cases in which risk undermines peoples’ moral status as agents, risk is harmful in virtue of setting back agents’ dignity-interests. I think Placani’s account gets things backwards—subjecting others to risk is disrespectful *because* we are under directed duties not to subject others to risk of harm, so appealing to disrespect is not going to help us ground the right (see (Kagan 1989, 176) for a similar argument about the role of respect in moral theory). There are two further problems with Placani’s account. First, it can only accommodate risk of harm being itself harmful *when imposed by other agents*. This is because only other agents can disrespect us. However, things other than agents can harm us, such as trees and animals. Placani’s view leaves a strange asymmetry: whether risk is harmful depends on whether it was caused by an agent. For general discussion on neutrality between harms caused by agents and non-agents, see chapter 3, section 3.2 and (Bradley 2012, 394; Herington 2018, 182). More generally, it will not offer us a full solution to the Problem of Pure Risk because of this. Second, suppose I subject someone to risk of harm. Whether this is itself harmful will depend, on Placani’s account, on whether I am acting justifiably or not—since I do not disrespect a culpable aggressor

Commonly, we do not think that risk of harm is itself harmful. That much should be indicated by my saying “*risk of harm*” and not, “risk of *more harm*”. While this itself is perhaps not a decisive objection, it does lead to a problem with the very idea that risk is harmful. Suppose that Threatener imposes a 0.5 risk of harm, *h*, onto Victim₁, and a 0.25 risk of the same harm, *h*, onto Victim₂. For example, Threatener will electrocute each victim. All else is equal (specifically, each victim responds to Threatener’s electrocution in the same way). It seems clear that if risk is harmful, the greater the probability of some harm, *h*, occurring, the more harmful the risk is. This implies, if the risks do not materialise for either victim, Victim₁ is harmed more than Victim₂. Assuming that risk is harmful, this is intuitively plausible.

But if risk is itself harmful then, were both risks to materialise, Threatener harms Victim₁ more than she harms Victim₂. This is because both victims suffer the same harm (the electrocution) and Victim₁ also suffers a greater additional harm (having been at a greater risk) than Victim₂. This is intuitively implausible. Were one to be presented with the original case and told that both risks materialise, we would want to say that each victim has been harmed to the same extent. Call this the Magnitude of Harm Problem.¹⁷⁵

when I justifiably act in self-defence. This again leaves a strange asymmetry: wrongful risk of harm is itself harmful, but risk of harm that is not wrongful is not itself harmful.

Lazar offers a third reason for his view in addition to the two mentioned at the end of section 3.1. He says, ‘one’s security is often grounded in others’ positive dispositions towards one’ and ‘[e]ven if I need never draw on others’ concern, I am better off just by their having this positive disposition towards me’ (Lazar 2017, 8–9). It is unclear how others’ disposition towards me in itself affects my wellbeing. Perhaps others’ dispositions towards me may affect others’ valuable relationships with me. Even if this is correct, it leaves his view vulnerable to the two objections raised against Placani’s view.

(John 2011) thinks that risk of harm is itself harmful because it undermines agents’ capacity to form reasonable plans. I am largely sympathetic to Herington’s objection that one needs only to *believe* one is free from risk for one to enjoy the good of planning (Herington 2018, 194–98); one’s beliefs of this sort are not affected in cases of pure risk (see the discussion above in section 1 on such harms being downstream from the risk imposition). What is more, John’s focus on planning leaves his view vulnerable to objections similar to those raised against the Autonomy Solution (especially, sections 3.2 and 3.3): for example, babies do not enjoy the good of planning, so cannot be harmed in virtue of frustrating that good.

¹⁷⁵ Two notes. First, the Magnitude of Harm Problem relies on an intuition—that Victim₁ and Victim₂ are harmed to the same extent. Finkelstein thinks risk of harm is itself harmful and has this intuition in a similar case she considers (Finkelstein 2003, 990-1). One way to move beyond the intuition is to think about the problem’s implications. For example, if risk of harm is itself harmful, given that we tend to think the level of compensation that is owed is proportionate to the harm caused, this would imply that Victim₁ is owed

Put differently, we can distinguish between the risk harm (the harmfulness of being exposed to some risk), the outcome harm (the harmfulness of the materialisation of the risk), and the all-things-considered harm. Since the risk harm is greater for Victim₁ than Victim₂, and the outcome harm is the same for each victim, the all-things-considered harm must be greater for Victim₁ than Victim₂. But this seems implausible.

In support of the Magnitude of Harm Problem, recall how the Counterfactual Account of Harm allows events caused by non-agents to be harmful.¹⁷⁶ A tree falling in the wind can harm me just as much as a person's dropping a tree on me can. The Magnitude of Harm Problem seems even more counterintuitive when we think of non-agential harms. Suppose Victim₁ and Victim₂ are at different risks of being harmed by some natural event, for example, being struck by lightning. Now suppose both victims are struck. It is very implausible that Victim₁ is harmed more than Victim₂ insofar as she was subjected to greater risk than Victim₂.

Note, the Magnitude of Harm Problem is consistent with the verdict that Threatener *wrongs* Victim₁ to a greater extent than she wrongs Victim₂. Being able to say this simply requires offering an account of the moral significance of risk that does not rely on risks' being harmful.

One might reply to the Magnitude of Harm Problem by suggesting that, when we say Victim₁ and Victim₂ suffer the same harm, we implicitly disambiguate the harms they suffer. First, we have the different risk harms they suffer through being subjected to different risks of the same harm. Second, we have the *same* outcome harm that they suffer from the electrocution. When we say that both victims are harmed to the same extent, we are referring to the same outcome harm that they suffer from the electrocution. We are not referring to the different risk harms they suffer.

more in compensation than Victim₂. But that is odd. Second, here's another version of the Magnitude of Harm Problem: Victim₃ is exposed to probability, x , of harm h ; Victim₄ is exposed to a greater probability, y , of a smaller harm, g . If both risks materialise, there is a some set of values for our four variables that mean that the victims have been harmed to the same extent. That is implausible.

¹⁷⁶ Chapter 3, section 3.2.

This reply will not do. If the oddness of thinking that Victim₁ and Victim₂ suffer different harms was reducible to an ambiguity relating to which harm we were asking after, the counterintuitiveness that the Magnitude of Harm Problem identifies would vanish when we re-ask,

all-things-considered, is Victim₁ harmed more than Victim₂?

But the answer to this *unambiguous* question seems to be: “No, both victims are harmed equally.” The counterintuitiveness of saying that the victims are harmed to a different extent cannot be explained away by appealing to ambiguity.

Claire Finkelstein offers us a potential solution to the Magnitude of Harm Problem.¹⁷⁷ She says that ‘the disvalue of the risk harm is absorbed into the loss in welfare if the risk actually materializes’ (Finkelstein 2003, 993). So, in reply to our case above, if neither of Threatener’s risks materialise, both victims are harmed merely by being subject to risk. Victim₁ is harmed more than Victim₂. But, if both risks materialise, the different risk harms they suffer are absorbed into the outcome harms. And, the outcome harms are equal. Call this the *Absorption Solution*.

There are at least two problems with the Absorption Solution. First, it is very difficult to see what principled explanation might lie behind the Absorption Solution. *Why* does risk harm persist only when the risked harm does not materialise?¹⁷⁸ Second, I am very unsure of how the Absorption Solution is actually supposed to work.

Suppose the outcome harm of either victim being electrocuted is 100. And suppose, for illustrative purposes, we apply a crude form of expected utility theory to determine how

¹⁷⁷ Finkelstein is responding to two slightly different problems (Finkelstein 2003, 990-1). She thinks we should distinguish between intentionally harmful acts and risky harmful acts. This dichotomy is inconsistent with the idea that individuals can intentionally subject others to risk of harm. But this is clearly possible.

¹⁷⁸ One might reply, in cases in which the risked harm materialises, this shows the risk of harm is actually 1. So, both Victim₁ and Victim₂ are harmed by the same amount. Yet, on what basis do we get to say that the probability of harm is actually 1? I do not pursue this question here because whatever we say will imply, on grounds of parity, that if the risked harm does not materialise, the risk was actually 0; thereby, the risk harm will be nothing. So, the Problem of Pure Risk will not be solved by arguing the risk is harmful.

harmful the risk of being electrocuted is.¹⁷⁹ Victim₁ has a 0.5 chance of being electrocuted: she is harmed by 50 through being subjected to risk of being electrocuted. Victim₂ has a 0.25 chance of being electrocuted: she is harmed by 25 through being subjected to risk of being electrocuted. The Magnitude of Harm Problem says that, if both risks materialise, the all-things-considered harm for Victim₁ is 150 and for Victim₂ is 125. This is implausible.

The Absorption Solution says the all-things-considered harms are equal because the risk harm is absorbed into the outcome harm, and the outcome harms are equal. Yet, how can this be? If the risked harms do not materialise, Victim₁ is harmed by 50 and Victim₂ is harmed by 25. If the risked harms do materialise, both victims are harmed by 100, all-things-considered. But, this means the extent to which Victim₁ is worse off if the risked harm materialises than if it does not materialise is 50 and the extent to which Victim₂ is worse off if the risked harm materialises than if it does not materialise is 75. The extent to which either victim is worse off if the risked harm materialises than if it does not *just is* the outcome harm. So, if the Absorption Solution is correct, the outcome harms cannot be equal—the outcome harm for Victim₁ is 50 and for Victim₂ 75. But this is implausible. How can the outcome harms be different depending on the level of risk one was subjected to?¹⁸⁰

This section has shown that, if risk is harmful, then two agents who are subjected to different levels of risk of the same outcome harm are, if both risked harms materialise,

¹⁷⁹ We will multiply the probability of her being harmed by the harm that she will suffer, were the risked harm to materialise. All I say should be consistent with other ways of determining *how* harmful some risk of harm is.

¹⁸⁰ Perhaps one might suggest the outcome harms are 100 for each victim, even though Victim₁ is only worse off by 50 and Victim₂ is only worse off by 75 if the risked harms materialise. They might explain this by saying the risk harm *preempts* some of the outcome harm. While we are familiar with preempted harm from chapter 4, a simple example should be helpful. Suppose Ann shoots Beth. Had Ann not shot Beth, Annabelle would have shot Beth. We might think the outcome harm is like Ann's shooting Beth and the risk harm is like Annabelle's shooting Beth. I have argued that defenders of the Counterfactual Account of Harm should accept the intuitively problematic conclusion that people are not actually harmed by preempted harm. Luckily, we can still account for preempted harm's moral significance by appealing to the Safety Condition as suggested in the previous chapter. In any case, if a defender of the Absorption Solution were to go this way, they would owe us a compelling explanation of *why* risk harm preempts some of the harmfulness of the outcome harm.

harmed to different extents all-things-considered. This is counterintuitive. This was the Magnitude of Harm Problem. We then considered the Absorption Solution to this problem—when risked harms materialise, the risk harm is absorbed into the outcome harm. Yet, this was even more implausible for it implies, when all else is equal, outcome harms are different depending on the level of risk one was exposed to.

5. Conclusion

This chapter has answered two questions. First, is risk itself harmful? Second, if risk is harmful, does risk's being harmful solve the Problem of Pure Risk? I have answered "No" to both questions. In reply to whether risk is harmful, we have just seen that the Magnitude of Harm Problem implies that any view on which risk is harmful leads to implausible results. We have also seen that the Desire Solution fails because the desire that we not be at risk of harm is a derivative desire, and the setting back of derivative desires does not *itself* constitute a harm. And we have seen that the Autonomy Solution says that risk is harmful only if one accepts Adequate Range and if one values autonomy for non-derivative reasons. But, even then, the Autonomy Solution leaves us with implausible asymmetries: it says that exposing autonomous individuals to risk is, in some circumstances, harmful, while exposing individuals with undeveloped, compromised, or damaged rational capacities to risk is not harmful; and it says that Locker's exposing Chooser to risk is harmful in *Two Doors* but that Target's exposing Target to risk in *Roulette* is not harmful.

In reply to our second question (whether risk being harmful solves the Problem of Pure Risk), since risk is not harmful, risk being harmful cannot solve the Problem of Pure Risk. I did note a conservative conclusion. Even if a defender of the Autonomy Solution is willing to accept all the objections and restrictions raised against it, risk being harmful does not offer a full solution to the Problem of Pure Risk.

7. Facts, Evidence, or Beliefs?

1. Introduction

The previous chapter argued that risk of harm is not itself harmful. This means that in *Roulette* Shooter does not harm Target. Since the risked harm Shooter subjects Target to does not materialise, Target's wellbeing is not of sufficient weight to place Shooter under a duty not to play roulette with her. And this means that the necessary condition set for the ascription of a right on the canonical statement of the Interest Theory is not satisfied. So, Target has no right against Shooter that Shooter not play roulette with her.

Arguing that risk is harmful is one way to dissolve the Problem of Pure Risk. Let me reintroduce the other way we could go. Suppose Ann is morally conscientious. She is trying to work out whether Beth holds a right against her. She then begins to wonder whether it is because of, first, her *beliefs* that Beth holds a right against her. Or, second, because of the best available *evidence* that Beth holds a right against her. Or, third, the *facts*—how things actually are—that makes it that Beth holds a right against her. All of this is to say that Ann wonders whether the rights that others hold against her and, correlatively, the directed duties she is under are determined by her beliefs, the evidence, or the facts.

That was fairly abstract. Think of Shooter. We can suppose she thinks there is a one in six chance that she will kill Target. And we can suppose the best available evidence supports this. If rights do not depend on the facts but on duty-bearers' beliefs or the evidence available to them, Target's wellbeing will be of sufficient weight to place Shooter under a duty not to play roulette with her. The Problem of Pure Risk is solved.¹⁸¹

¹⁸¹ Moving from a fact-y view may also provide a solution to the Problem of Preemption: it certainly solves *Plane Crash* where, presumably, Attendant does not believe nor does the evidence support that she is benefiting Passenger by denying her admittance onto the plane. Whether this move from focusing on the facts is a full solution to the Problem of Harmless Wronging depends on what we think of a variant in which one performs some would-be harmful act in the knowledge it is preempted. I am tempted to think Hitman₁ violates Victim's rights even if she knows Hitman₂ will deliver a lethal shot if she does not shoot. Appealing to her beliefs or the evidence will not help us there.

This chapter argues that rights depend on the facts. So, the Problem of Pure Risk is not so easily solved. (That said, there will be a note of trepidation in that I leave an important question unanswered.) This paves the way for my showing how the Safety Condition solves the Problem of Pure Risk in the following final chapter. In section 2, I distinguish further between beliefs, evidence, and facts. In section 3 and 4, I argue against what I call the Belief- and Evidence-Relative Views. And in section 5, I consider two objections to what I call the Fact-Relative View.

Before beginning, three preliminaries. First, suppose I am wrong and that we ought to endorse the Belief- or Evidence-Relative View. This would provide us with a solution to the Problem of Pure Risk. Even still, the Safety Condition provides a solution to the Problem of Harmless Wronging whichever view is correct. This gives us reason to endorse the Safety Condition independently of whether the Fact-Relative View is correct.¹⁸²

Second, if one does not endorse the view that rights are grounded in one's beliefs, one will want to make use of *blameless wrongings*. *Y* blamelessly wrongs *X* iff *Y* infringes a directed duty owed to *X*, though is not blameworthy for doing so. We have good independent reason to think there are blameless wrongings. For example, when someone acts under duress, in extreme cases we tend to think that she acts wrongly, though is not blameworthy for doing so. And in less extreme cases, we tend to think that she is still blameworthy, though not as blameworthy as she would have been without the duress. This is sufficient to give us reason to think wrongdoing and blameworthiness come apart.¹⁸³

Third, most people think that rights obtain in virtue of something. Interest Theorists think that rights hold in virtue of their holders' wellbeing. Will Theorists think that rights hold in virtue of right-holders' autonomy. Even theories that have a non-instrumental character tend to say that rights obtain in virtue of something. For example, Frances Kamm says

¹⁸² My hunch is that the strength of this reason is inversely proportionate to our confidence in either the Belief- or Evidence-Relative View. For example, suppose we are very sure of the Evidence-Relative View. Then, that the Safety Condition solves the Problem of Pure Risk whichever view is correct does not really matter. Now suppose we think the Evidence-Relative View is correct, but we are not very sure. Then, we have good reason to endorse the Safety Condition.

¹⁸³ See (Zimmerman 1997) for defence of the view that one can be blameworthy for acting in the right way.

that rights obtain in virtue of their reflecting our inviolability.¹⁸⁴ Whether rights depend on the facts, evidence, or beliefs is of interest whichever of these views we endorse. The Fact-, Evidence-, and Belief-Relative Views are written so as to be agnostic on which view about the nature of rights is correct (though obviously enough wellbeing matters for us).¹⁸⁵

2. The Distinction

In *Case 1 (Belief-Relative)*,

Doctor believes, against the best evidence available to her, that giving treatment will kill Patient. Doctor gives the treatment to Patient, and Patient lives.

In *Case 1*, in some sense, Doctor does not act wrongly. She does what the evidence tells her to do. And she saves Patient when, suppose, Patient would otherwise have died. However, there is some sense in which Doctor acts wrongly. She does what she believes will kill Patient.

In *Case 2 (Evidence-Relative)*,

Doctor believes, against the best evidence available to her, that giving treatment will save Patient's life. Doctor gives the treatment to Patient, and Patient lives.

In *Case 2*, in some sense, Doctor does not act wrongly. She does what she believes will save Patient. And, as it turns out, she saves Patient when Patient otherwise would have

¹⁸⁴ '[F]undamental human rights [...] are not concerned with protecting a person's interests, but with expressing his nature as a being of a certain sort, one whose interests are worth protecting' (Kamm 2007, 253).

¹⁸⁵ Suppose we endorse the Will Theory and that rights depend on potential duty-bearers' beliefs. Then, X will have a right against Y that $Y \not\Phi$, iff, relative to Y 's beliefs, X has the power to alter Y 's duty to Φ . Suppose we endorse Kamm's picture of the nature of rights and that rights depend on the evidence. Then, X 's right will depend on whether the evidence supports that X is inviolable.

died. However, there is some sense in which Doctor acts wrongly. She does not do what the best available evidence tells her to do.

And in *Case 3 (Fact-Relative)*,

Doctor believes, on the best evidence available to her, that giving treatment will save Patient's life. Doctor gives the treatment to Patient and Patient dies.¹⁸⁶

In *Case 3*, in some sense, Doctor does not act wrongly. She does what she *believes* will save Patient. And she does what, on the best available *evidence* to her, will save Patient. But, as it turns out, she *in fact* kills Patient (suppose that Patient would have survived were it not for the treatment). In *some* other sense, then, she acts wrongly.¹⁸⁷

In *Case 1*, Doctor acts wrongly in the belief-relative sense, but not in any other sense. In *Case 2*, Doctor acts wrongly in the evidence-relative sense, but not in any other sense. And in *Case 3*, Doctor acts wrongly in the fact-relative sense, but not in any other sense.

Before seeing how this tripartite distinction applies to rights, two clarifications. First, so far, I have only presented three cases that support the idea that we can distinguish between these three senses of wrong. I have not said anything about which of these senses we ought to be concerned with.¹⁸⁸

Second, people often distinguish between the *objective* and the *subjective*. In reply to *Case 1*, they say that Doctor does not act wrongly in the objective sense but does act wrongly in the subjective sense. And in reply to *Case 3*, they say that Doctor does not act wrongly in the subjective sense but does act wrongly in the objective sense. While this distinction is good in as far as it goes, it does not enable us to account for the wrong of *Case 2*. There

¹⁸⁶ For more on this tripartite distinction, see (Parfit 2011a, 151–53; Tadros 2011b, 217–20).

¹⁸⁷ For support for the sense in which one can act wrongly in the fact-relative sense, see section 5.1 and (Parfit 2011a, 151–53; Thomson 1990, 79–104).

¹⁸⁸ Some try to argue for only one perspective (Thomson 1990, 79–104). Others do not deny the existence of all kinds of perspectives, but suggest that one is fundamental (Parfit 2011a, 161–62) or that one is that with which we are *really* interested (Zimmerman 2014, 18–24). Others, still, are pluralists about which perspective is important (Smith 2018).

seems to be a sense in which Doctor acts wrongly, but not by reference to the subjective or objective senses. Appealing to the evidence-relative sense fills this gap.¹⁸⁹

3. Beliefs

We might hold the

Belief-Relative View. X 's right that $Y \Phi$ depends on Y 's beliefs.¹⁹⁰

The idea behind the Belief-Relative View is supposed to be that Y has certain beliefs about what will happen in the world were she to Φ and were she not to Φ .¹⁹¹ If, relative to her beliefs, she would be under a duty to Φ owed to X , she *is* under a duty to Φ , correlating with X holding a right against her. Here is another way to think about it: if, were those beliefs to be true, she *would* be under a duty to Φ owed to X , then she is under a duty to Φ correlating with X holding a right that $Y \Phi$. For example, I have a set of beliefs about what would happen were I to stab you—you would begin to profusely bleed, and so on. Were those beliefs to be true, whatever our view of rights (but especially on the Interest Theory), you would have a right that I not stab you. So, on the Belief-Relative View, you have a right that I not stab you. This holds irrespective of what is likely to happen were I to stab you or of what will actually happen.

The way I have formulated the Belief-Relative View has it that the beliefs that are relevant for rights are those of the potential duty-bearers. The reason for this is that, when all else

¹⁸⁹ This gap is often filled by the “reasonable person”. The reasonable person is *one* specification of the evidence-relative view (the evidence set that the reasonable person would have available to them that they are able to respond to) (Oberdiek 2017, 49). Others use *Prospectivism* to fill the gap between the belief-relative/subjective and the fact-relative/objective, where what is *prospectively* best is that which gives the best ‘prospect of achieving what is of value in the situation [...] and of avoiding what is of disvalue’ (Zimmerman 2014, 32).

¹⁹⁰ So, an Interest Theorists of this flavour would say something like, for X to have a right against Y that $Y \Phi$, X 's interests must be of sufficient weight, relative to Y 's beliefs, to place Y to be under a duty to Φ .

¹⁹¹ If one endorses the Belief-Relative View and thinks there are both directed and undirected duties, it is likely one will think undirected duties depend on one's beliefs, too. Yet, it is possible that one might hold a fact-relative account of undirected duties, but think whether you owe something to someone, correlating with their holding a right against you, is a matter of your evidence.

is equal, the right in question requires something of the potential duty-bearer; so, it should be *their* beliefs that are relevant. We could specify the view in other ways, for example, that the beliefs that are relevant are those of potential right-holders.

Here is a much less plausible way that the Belief-Relative View could work:

Moral Belief-Relative View. X 's right that $Y \Phi$ depends on Y 's believing: " X has a right that I Φ ."

The Moral Belief-Relative View is much less plausible than the Belief-Relative View. It is deeply implausible that one party holds a right against another party only when the second party *thinks* that the first party holds a right against her.

An initial problem with the Belief-Relative View is that, however one motivates the Belief-Relative View, it is unclear why the Moral Belief-Relative View should not be correct in place of the Belief-Relative View. Put differently, what reason could be given in favour of the Belief-Relative View that does not speak, to a greater extent, in favour of the Moral Belief-Relative View? For example, on the Evidence- and Fact-Relative Views (to be defined below), X can hold a right against Y , even when Y has no beliefs about how her actions might affect X . Perhaps one might motivate the Belief-Relative View by saying that it is unfair that Y can owe something to X when she is unaware of the features of the world that place her under that duty. But if this is true, why is it not also unfair that X can hold a right against Y when she is unaware that X holds a right against her?

A second problem with the Belief-Relative View is that it will implausibly overgenerate rights. This is because Y might mistakenly believe something about the world that would mean X holds a right against Y were those mistaken beliefs to be true. But it might be intuitively implausible that X holds this right against Y . Consider,

Voodoo. Believer wants to cause Victim great pain. Believer believes in voodoo. Unbeknownst to Victim, Believer creates and then stabs a voodoo doll of Victim.

I find it implausible that Believer infringes Victim's rights. But, were Believer's beliefs about the world to be correct, Victim would have a right that Believer not stab a voodoo

doll of him. So, on the Belief-Relative View, Victim has a right against Believer that Believer not stab a voodoo doll of him.

Perhaps others might think that Victim *does* have a right that Believer not stab a voodoo doll of him, a right that Believer violates.¹⁹² However, the Belief-Relative View pulls in the opposite direction too: it implausibly *undergenerates* rights. And, even if one is not worried about the overgenerations, it is harder to ignore the undergenerations. The Belief-Relative View undergenerates rights because *Y* might mistakenly believe something about the world that would mean that *X* does not have a right against *Y* were those mistaken beliefs to be true. But it might be highly implausible that *X* does not hold such a right against *Y*. Consider,

Real Gun. Non-Believer has a gun in front of her. She has good evidence that the gun is real, though she has failed to avail herself of that evidence. She fires the gun at Victim, thinking that it is a toy gun.

Non-Believer violates Victim's rights. Some might want to say that she is not blameworthy for doing so, but she violates Victim's rights nonetheless. However, firing toy guns is not the sort of thing that others have rights against you that you not do. Non-Believer believes she is about to fire a toy gun. Given the Belief-Relative View, Victim does not have a right that Non-Believer not fire the gun at her.¹⁹³

So, the Belief-Relative View implausibly undergenerates rights. The reason underlying why the Belief-Relative View implausibly undergenerates rights gives us a more general, fourth reason not to endorse the view: the Belief-Relative View makes what rights *I* have

¹⁹² I must admit, my intuitions are less clear on whether Believer has *wronged* Victim.

¹⁹³ One might think this case is under-described. Suppose there is an envelope on top of the gun saying, "Herein contains whether or not this is a real gun." After reading this, Non-Believer believes she may find out whether the gun is real by opening the envelope but decides not to. Given the high stakes of the case and Non-Believer's beliefs about what is contained in the envelope, one might think Non-Believer owes it to Victim to open the envelope, and Victim has a correlative right against Non-Believer that she do so. Even if this is correct, this would imply only that Victim has a right that Non-Believer open the envelope, and not a right that Non-Believer not fire the gun at her. But Non-Believer clearly has the second of these rights regardless of whether she has the first. And, in any case, we might not always be able to trace back to such a belief.

depend upon others' beliefs about the world. But that is implausible. What rights I have should not depend on others' beliefs. That fails to pay attention to the importance that *I* have—an importance that should be reflected by the rights that I hold.

4. Evidence

The natural remedy for the Belief-Relative View's tendency to undergenerate rights is to say, "Well, Non-Believer *should've* known that the gun was real—there was good evidence available to her!" We might thus hold the

Evidence-Relative View. *X's* right that $\mathcal{Y} \Phi$ depends on the best available evidence to \mathcal{Y} .¹⁹⁴

In its most general form, the Evidence-Relative View will say that *X's* right that $\mathcal{Y} \Phi$ depends on the evidence. We can then specify different evidence-relative sets to have different disambiguations of the view. I have gone with the "best available" specification as it seems most intuitive.

We can get at what is meant by evidence by thinking about *Real Gun*.¹⁹⁵ Suppose that the gun looks very real, there are signs all around saying "WARNING: Live Firearms", and so on. Non-Believer has good evidence available to her that the gun is real, even though she never avails herself of this evidence and never forms the belief that the gun is real. Similarly, when a doctor receives some blood test results, she has evidence available to her as to whether her patient has this-or-that condition before she opens the results.

This points to another salient feature of evidence: whether some piece of evidence is available to \mathcal{Y} , and so whether it impacts what rights others hold, also depends on whether \mathcal{Y}

¹⁹⁴ So, an Interest Theorist of this flavour would say, for *X* to have a right against \mathcal{Y} that $\mathcal{Y} \Phi$, *X's* interest must be of sufficient weight, relative to the best available evidence to \mathcal{Y} , to place \mathcal{Y} under a duty to Φ .

¹⁹⁵ What is meant by evidence here is not what is tended to be meant by epistemologists. For example, Williamson sees one's evidence as the totality of propositions one knows. Conee and Feldman see one's evidence as all of one's current mental states (Williamson 2000; Conee and Feldman 2004). It is hard to square these accounts of evidence with how the moral theorist wants to use it.

can respond to that evidence. For example, while the doctor may have good evidence available to her in the form of the test results, the patient does not—the patient simply could not avail herself of that evidence.

So, we have a working understanding of what is meant by the best evidence available to the duty-bearer. While there is more to be said, this should be sufficient for our purposes. A final feature of the Evidence-Relative View is that what rights people have depends on the evidence *to potential duty-bearers*. I take it the thought behind this is again that rights place demands on the correlative duty-bearer. And so, it is the best available evidence *to them* that determines whether others hold rights against them.¹⁹⁶

With that in place, let us assess the Evidence-Relative View. First, when objecting to the Belief-Relative View, I said

The Belief-Relative View makes what rights *I* have depend upon others' beliefs about the world. But that is implausible. What rights *I* have should not depend on others' beliefs. That fails to pay attention to the importance that *I* have, an importance that should be reflected by the rights that *I* have.

We might be sceptical that questions about the evidence others possess and are able to respond to determines what rights *I* hold. Again, we might think this does not pay enough attention to the significance that rights reflect about their holder.

I am unsure of whether this objection will move those who feel the force of the Evidence-Relative View.¹⁹⁷ So, let us move onto our second objection. Because the Evidence-

¹⁹⁶ Recall the similar remarks I made about the Belief-Relative View. These sorts of considerations are most explicit in (Scanlon 2008; Quong 2015; Oberdiek 2017).

¹⁹⁷ Saying that, Zimmerman 'acknowledges the force of this objection' (Zimmerman 2014, 117–18). It is interesting that, if we focus on directed duties rather than their correlative claims, this objection is a lot less powerful (i.e., it does *not* seem implausible that what duties *I* owe others depends on the evidence available to me—in fact, that sounds plausible). Perhaps this gives us some reason to doubt correlativity. It also takes us into interesting questions of whether we ought to see rights or duties as prior in justificatory terms (Raz 1986, 170–71; Waldron 1985, 14; Kramer 2000, 39). I am inclined to see rights as prior, so am tempted to focus more on the implausibility of my rights depending on the evidence available to others, and not the plausibility of others' duties depending on the evidence available to them.

Relative View says that what rights we have depends on the best available evidence to the duty-bearer, this means when new evidence comes into existence, the rights we have changes. This verdict is odd. Consider

Day's End. [Resident] always comes home at 9:00 pm., and the first thing he does is to flip the light switch in his hallway. He did so this evening. [Resident] flipping the switch caused a circuit to close. By virtue of an extraordinary series of coincidences, unpredictable in advance by anybody, the circuit's closing caused a release of electricity (a small lightning flash) in [Neighbour's] house next door. Unluckily, [Neighbour] was in the path and was therefore badly burnt. (Thomson 1990, 229)

Given the Evidence-Relative View, when Resident harms Neighbour, Resident does not infringe Neighbour's rights. Some find this implausible. I am on the fence. (It is unlikely that one would want to say that Resident is blameworthy for doing so—but this is consistent with Resident infringing Neighbour's rights. Again, there are blameless wrongs.) Instead of focusing on this feature of the case, let us focus on our present challenge of what happens when new evidence becomes available to potential correlative duty-bearers.

Suppose that some evidence becomes available to the electrical board about what is going to happen when Resident flips his light switch. They call Resident to warn him. While on a first-order level I am unsure of whether Neighbour has a right that Resident not flip the switch, I am more sure that it would be weird for the electrical board to say, "Look Resident, we're aware of some evidence that you're not. Were we to make this evidence available to you, it will be true to say that Neighbour has a right that you not flip that light switch; but it will not be true to say that he has the right if we don't make that evidence available to you." It seems more natural for them to say, "We've become aware of some evidence about whether Neighbour has a right that you not flip that switch." But if the Evidence-Relative View is correct, Neighbour only gets a right that Resident not flip the switch *once* the evidence becomes available to Resident. That seems implausible.¹⁹⁸

Might one object that, given the electrical board is aware of the problem, this means that there is evidence available to Resident that means he is under a duty not to flip the switch,

¹⁹⁸ Thomson uses a similar argument to this against the belief-relative ought (Thomson 1990, 223).

correlating with Neighbour holding a right that he not flip the switch? On the one hand, I am not sure that the evidence is *available* to Resident. Is there evidence available to you whether your spouse has lied to you on the grounds they could tell you? This is odd. In any case, on the other hand, this only pushes the point back—when the evidence became available to the electrical board, new rights came into existence. And this is what I find odd.

T. M. Scanlon says, ‘[i]n the original example, the injury to [Neighbour] was said to be due to “an extraordinary series of coincidences, unpredictable in advance by anybody.” In the modified example, [the electrical board] knows about this effect and could easily tell [Resident]. So the situation is quite different’ (Scanlon 2008, 51). Similarly, Jonathan Quong says, ‘[t]he initial description of the example stipulates that the harm [Resident’s] flipping of the switch will cause is unpredictable. But if this statement is true, then the situation is unpredictable in advance by anybody’ (Quong 2015, 251). These two remarks are confused. Something can be unpredictable at t_1 but become predictable at t_2 (for example, because evidence has become available at t_2 that was not previously available at t_1). And in such cases, the Evidence-Relative View will say a right *comes into existence* between t_1 and t_2 . But it seems more natural to say that people *find out* (or, *get better evidence about*) what rights hold. This speaks in favour of the Fact-Relative View.

So far, we have two reasons to be sceptical of the Evidence-Relative View: first, that it is implausible that our rights depend on the evidence available to others and others’ ability to respond to evidence; second, that it implies when new evidence comes into existence, new rights come into existence rather than us gaining better evidence about what rights exist. Let us move onto a third problem with the Evidence-Relative View (one that we might think is the most worrying): the Evidence-Relative View counterintuitively under-generates rights. Consider

Duped Soldiers. A group of young soldiers are successfully fooled by a totalitarian regime into believing that the regime is good and just, and is under repeated attacks from their evil neighbours, the Gloops. The regime’s misinformation campaign is subtle and absolutely convincing: the soldiers are justified in believing what they are told by the regime. Once the misinformation campaign is complete, these Duped Soldiers are given orders to attack and destroy a Gloop village on the border, which, they are told, is really a Gloop terrorist camp plotting a major attack. In fact, everything the regime has said is a lie, and the Gloop village

contains only innocent civilians. The Duped Soldiers prepare to shell the village and are about to (unknowingly) kill all the innocent civilians in it. A peacekeeping force from a neutral third country patrols the border and could avert the attack, but only by killing the Duped Soldiers. (Quong 2015, 261)

By posing an unjustified threat to others, individuals can make themselves liable to be harmed. To say that an individual is liable to be harmed is to say that harming them would not wrong them nor would it violate their rights, and so they would not be justified in defending themselves.¹⁹⁹ In *Duped Soldiers*, the best available evidence to the soldiers says that the Gloop villagers are liable to be attacked, so have no rights against being attacked. This means, given the Evidence-Relative View, the soldiers (1) do not violate the villagers' rights, (2) do not wrong the villagers, nor are they (3) liable to be harmed themselves in defence of the Gloop villagers (for example, by the peacekeeping force). All three of these verdicts are deeply implausible.

Quong agrees that these verdicts are 'unacceptable' (Quong 2015, 261). However, he does not think his version of the Evidence-Relative View is committed to them.²⁰⁰ He begins by separating the following two questions:

- (i) Under what conditions does a person have a claim[-right] not to be harmed by a particular type of act performed by another person?
- (ii) Has some particular person, *A*, waived, transferred or forfeited this claim[-right] not to be harmed by another person, *B*?

(Quong 2015, 261)

When answering the first question, Quong thinks we should appeal to evidence-relativity (for our purposes, the Evidence-Relative View). But when answering the second question, we should appeal to 'what *A* has actually done, and not on *B*'s evidence about what *A* has done' (Quong 2015, 261).

¹⁹⁹ Some might add that, for an individual to be liable to some harm, they must have forfeited rights that they previously held against that harm (McMahan 2005, 386; Frowe 2014, 3); cf. (Tadros 2016b). Some might think that it does not necessarily follow from *T*'s being liable to *X* defending herself that *T* is not permitted to defend herself—one might think liabilities can be symmetrical. We return to this below.

²⁰⁰ Zimmerman accepts these implications in individual cases of self-defence (Zimmerman 2008, 97–117).

(Quong thinks that the fact-relative view of forfeiture and waiver ‘grants the right-holder a more effective degree of control over the right, something that is typically of central importance in the justification of the right’ (Quong 2015, 262). I take it that this is meant to explain why the fact-relative view of forfeiture is not ad hoc.²⁰¹ However, it is not clear why something like this is not true of claim-rights against being harmed in general, which are determined from the evidence-relative perspective for Quong: the Fact-Relative View gives the right-holder a more effective degree of protection over her wellbeing, something that is typically of central importance in the justification of that right.)

Quong’s distinction between (i) claims against being harmed and (ii) whether one is liable to be harmed is unstable.

- (1) Recall from chapter 1, if X has a claim that Y not Φ , Y owes X a duty not to Φ . If you have a claim against me that I not hit you, I owe you a duty not to hit you.
- (2) And recall from a few paragraphs above, if X is liable to be harmed by Y , Y does not owe X a duty not to harm her (say, by Φ -ing). If you are liable to me hitting you, I do not owe you a duty not to hit you.
- (3) If Y is not under a duty not to Φ , X has no claim that Y not Φ . If I am not under a duty not to hit you, you have no claim against me that I not hit you.
- (4) So, whether X is liable to be harmed by Y affects whether X has a claim against Y . Whether you are liable to my hitting you affects whether you have a claim against me that I not hit you.

Here is the problem: in (1), we began by asking whether X has a claim that Y not Φ . According to Quong, that is determined by the evidence available to Y . In (2), we asked what happens if X is liable to be harmed by Y . According to Quong, that is determined by what X has in fact done. But from (2), which was determined by what X has in fact done, we arrived at the verdict that (3) X has no claim against Y . But, whether or not X

²⁰¹ This would leave Quong’s view unmotivated when it comes to those for whom control has no value. See discussion of the Incapacity Undergeneration from chapter 2, section 4.

has a claim against \mathcal{I} was supposed to be determined by the evidence available to \mathcal{I} , and not by what X has in fact done. So, something has gone wrong—you cannot separate claim-rights from liabilities to be harmed.

To put this in simpler terms, according to Quong, whether you have a claim that I not hit you depends on the evidence available to me. But, whether you have made yourself liable to be harmed by me (for example, whether you have made yourself liable to my hitting you) depends on what you have in fact done. But, whether you have made yourself liable to be harmed by me, which depends on what you have in fact done, affects whether you have a claim that I not hit you. Yet, whether you have a claim that I not hit you was supposed to be determined by the evidence available to me, and not by what you have in fact done.

Given the correlativity of claims against being harmed and the absence of a liability to be harmed, here is what I think a defender of the Evidence-Relative View is forced to say about *Duped Soldiers*: given the best available evidence to the soldiers, the villagers have no claim against the soldiers that they not be harmed since the villagers have (given the evidence) made themselves liable to be harmed. And, given the best available evidence to the villagers, the soldiers have no claim against the villagers not to be harmed since the soldiers have (both given the evidence and the facts) made themselves liable to be harmed. And so, what we actually have is both parties having a liberty-right to harm the other party, much like a boxing match. Yet, that is deeply implausible. One way to see this is to think of what a third-party may do if they see the events unfold. It is unclear what might explain why the third-party ought to intervene on behalf of the villagers—after all, both parties have a liberty right to try and harm the other party.²⁰²

Suppose the villagers knew that the soldiers had been duped. Then, things might get worse for the Evidence-Relative View. What might justify the villager's liberty-right to defend themselves? They cannot point to their right not to be killed as a justification for using defensive force for, by hypothesis, they have no right against the soldiers not to be attacked *and* they know this. If one thinks others become liable to be harmed only when they subject

²⁰² (Frowe 2015, 271–74) raises further problems with Quong's attempt to deal with *Days End*.

others to *unjustified* threats of harm, the villagers are not permitted to defend themselves. This is even more implausible than the verdict that each party has a liberty-right to defend themselves.²⁰³

5. Facts

I have argued we have three reasons to be sceptical of the Evidence-Relative View. First, the rights people hold should not depend upon the evidence available to others; second, when new evidence becomes available, we gain evidence *of* what rights already exist rather than new rights coming into existence; third, the Evidence-Relative View implausibly undergenerates rights.²⁰⁴ Given these problems and those with the Belief-Relative View, we ought to look elsewhere. Namely, to the

Fact-Relative View. X 's right that $Y \Phi$ depends on the facts.

The Fact-Relative View should not need much explaining. Instead, I would like to consider two problems with it.

4.1 *Fact-Relative Oughts*

The first problem begins with the very plausible idea that there must be some sort of connection between what rights others hold against us and what we ought all-things-

²⁰³ This will follow on accounts of liability defended by, e.g., (McMahan 2005; Fabre 2012; Frowe 2014). One might posit an agent-relative prerogative to defend oneself (Quong 2009; 2016; Fabre 2012, 61), yet this move would still struggle to explain why a third-party may intervene on behalf of the villagers. Obviously enough, it also requires we endorse agent-relative prerogatives to do great amounts of harm, which might strike one as implausible (Frowe, n.d.).

²⁰⁴ Helen Frowe has suggested that one way to respond to these problems from the perspective of the Evidence-Relative View is to reject the correlativity between rights and directed duties. Then we can say, *my* rights are determined by the evidence available to me, but *your* duties owed to me are determined by the evidence available to you. So, my rights do not change as new evidence comes available to *you* (though, my rights change as new evidence becomes available to me). And the villagers will have rights against the soldiers harming them. However, abandoning correlativity comes with costs. For example, even if the villagers have rights against being harmed by the soldiers, the soldiers are under no duty not to attack the villagers, so have liberty-rights to attack the villagers. Further, it is difficult to see how a third-party could prioritise saving the villagers (both parties are acting within their rights). Finally, there are more general costs of abandoning correlativity, namely that it strips rights of their relationality; this will be problematic if we think rights' relationality explains what is distinctive (and perhaps indispensable) of rights.

considered to do.²⁰⁵ But now suppose one could show there is no fact-relative ought—no sense of what we ought to do relative to the facts. How, then, does the Fact-Relative View relate to what we ought to do?²⁰⁶

Happily, we have good reason to think there is a fact-relative ought. Consider

Fever. Bloggs’s baby has a fever. The best available evidence to Bloggs says that it would be best to starve the baby. (Feed a cold, starve a fever.)
Bloggs starves the baby, and the baby dies.

Bloggs’s belief- and evidence-relative “oughts” say that she ought to have starved the baby. Now suppose that, at the hospital after the case has unfolded, a doctor says to Bloggs, “I realise that you are not at fault, but you really ought not to have starved the baby. You ought to have fed it aspirin dissolved in apple juice.” This seems perfectly natural. However, this “ought” can only be made sense of if the doctor is referring to what Bloggs ought to have done in the fact-relative sense (Thomson 1990, 172–73). This gives us reason to think there is a fact-relative sense of ought.

Tadros doubts that there is a fact-relative ought.²⁰⁷ Yet, he agrees that the doctor’s comment seems natural and that, ‘[w]hen Bloggs is given the information, we say, [s]he finds out what [s]he ought to have done’ (Tadros 2011b, 223). How, then, to account for this ought without appeal to fact-relativity?

Tadros suggests that, were Bloggs to have had better evidence, it would have been the case that she ought to have fed the baby aspirin dissolved in apple juice. Given this, when we say that Bloggs ought to have fed the baby aspirin dissolved in apple juice, we are

²⁰⁵ In chapter 1 I said we ought to be interested in the connection between what rights obtain and what duty-bearers are required to do all-things-considered. However, in this section I am going to speak of “ought”. This is because the felicity of some of the examples changes when we move from oughts to requirements. In conversation, Justin Snedegar has suggested the felicity of the examples changes, despite requirements being what we are really interested in, because of the different pragmatics of requirement from ought.

²⁰⁶ Note, the objection raised in this section is going to be especially worrying if we endorse the All-Things-Considered View on which, if X has a claim-right against Y that $Y \Phi$, Y is required [entailing Y ought] to Φ .

²⁰⁷ Though I do not have space to discuss it here, see (Tadros 2011b, 220–24).

comparing, on the one hand, what Bloggs ought to have done relative to the evidence that she had available to her at the time with, on the other hand, ‘what Bloggs ought to have done relative to some better set of evidence’ (Tadros 2011b, 223). More generally,

when we say that something is [fact-relative wrong] but not [evidence-relative wrong] what we really do is to contrast the epistemic circumstances of the person with some superior epistemic circumstances that might have been available to [them]. So fact relativity is better understood as superior epistemic relativity (Tadros 2011b, 224).

I am not convinced. First, a dialectical worry. Consider what sorts of reasons one might have for thinking there is an evidence-, but no fact-relative ought. Perhaps one finds it implausible that agents *ought* to do things if the evidence that they ought to do those things is not available to them. What we might call the *first-person* evidence-relative ought responds to this reason: it says what agents ought to do is that which the best available evidence, *to them*, that *they* can respond to, tells them to do. However, once we move away from the first-person evidence-relative ought to superior evidence sets—as Tadros suggests—we undercut the very reason we have for preferring the evidence-relative over the fact-relative ought. It is no longer that the evidence is accessible to the person the ought requires action of. (We could go further than merely pointing out this dialectical worry: once we have abandoned relativising ought to what the agent can actually respond to, why stop short of the facts?)

Second, some fact-relative oughts do not look like superior evidence-relative oughts. Let us amend *Fever*. Suppose that, luckily, Blogg’s baby survives *despite* Bloggs having starved the baby when better evidence would have told Bloggs to feed the baby aspirin. This is because Bloggs’s baby has a rare, unpredictable condition: sometimes, when the baby presents with feverish symptoms, starving the baby saves it, whereas feeding it aspirin would kill it; but, more often than not, feeding the baby aspirin saves the baby.

Now, it seems natural for the doctor to say something along the following lines: “In case this happens again, you ought to feed the baby aspirin dissolved in apple juice. I guess you were lucky, though, for you ought not to have done that this time, for that would have killed the baby—you ought to have starved the baby, as you did.” Here, doctor uses two conflicting senses of ought. The second ought is the fact-relative ought. Tadros says that the fact-relative ought should be understood as a superior evidence-relative ought. I find

it incredibly difficult to see how *this* fact-relativity ‘is better understood as superior epistemic relativity’. Getting lucky, as Bloggs has done, cannot constitute a superior evidence-set.

(If one wants to dig their heels in and hold onto Tadros’s strategy, we end up with the following dialectical point—an incredibly important dialectical point for the purposes of this thesis. To make sense of what appears to be the fact-relative ought, we are saying there are superior evidence-relative oughts. Sometimes, as in our amended *Fever*, this superior epistemic ought is determined by what, in fact, turns out to happen. If the evidence-relative ought can be determined in this way, the problem made salient by harmless wrongdoing reemerges—the superior epistemic ought tracks what turns out in our cases of harmless wrongdoing. But this means one *cannot* solve the Problem of Harmless Wronging by appealing to evidence-relativity, for evidence-relativity has been made so expansive so as to respond to what, in fact, happens. For example, in *Plane Crash*, relative to some superior epistemic ought, what Attendant ought to do is deny Passenger admittance on the plane. In *Roulette*, if harm is all we care about then, relative to some superior epistemic ought, it is not the case that Shooter ought not to play roulette with Target.)

So, we have good reason to think there is a fact-relative ought. Given this, we can say that rights contribute to, at the least, the fact-relative ought.

4.2 *Jackson Cases*

The second problem with the Fact-Relative View is that one might think it gets the wrong verdict in some cases.²⁰⁸ Consider,

Jackson’s Case. All the evidence at Doctor’s disposal indicates, in keeping with the facts, that giving Patient drug A would cure her partially and giving her no drug would render her permanently incurable. However, the evidence leaves it completely open whether it is giving her drug B or drug

²⁰⁸ Another way to put the problem is that, even if the rights correspond to what one ought to do in the fact-relative sense, *Jackson’s Case* shows that what one ought to do in the fact-relative sense is not particularly important; and so, if rights correspond to the fact-relative ought, rights are not particularly important.

C that would cure her completely, and whether it is giving her drug B or drug C that would kill her.²⁰⁹

Suppose that it is drug B that would completely cure Patient, and so drug C that would kill Patient. Given the Fact-Relative View, Patient will have a right that Doctor give her drug B. Yet, we might find this counterintuitive. In fact, we might think that Patient has a right, given all the available evidence, that Doctor give her drug A (as well as rights that Doctor *not* give her either drugs B or C).

Given that, on the Evidence-Relative View, X 's right that $Y \Phi$ depends on the best available evidence to Y , Patient will have a right that Doctor give her drug A (as well as rights that Doctor not give her drugs B and C). We can say that this holds because giving drug A is the *expectably* best option. So, it seems that cases like *Jackson's Case* speak in favour of the Evidence-Relative View, and against the Fact-Relative View.

There are a few options available to the defender of the Fact-Relative View. First, we might argue that the Fact-Relative View can recognise that Patient has a right to drug A. I make some remarks to this effect in the conclusion to this chapter. However, I also raise doubts with this way of proceeding. So, here I would like to examine other ways we may go.

Second, we might bite the bullet and suggest that, though the Fact-Relative View says that Patient has a right to drug B, this is not as counterintuitive as one might think. For example, we might say it is just unclear which treatment Patient has a right to. We will presumably have a second-best ought to deal with these sorts of case. However, one might press, do we have second-best rights that adhere to this second-best ought? Further, what is the relationship between these second-best rights and the Evidence-Relative View?

I am going to take a third option and suggest that the Evidence-Relative View is susceptible to structurally analogous counterexamples like *Jackson's Case*. Thereby, cases like

²⁰⁹ This case is taken from (Zimmerman 2014, 30) with some slight alterations, originating in (Jackson 1991, 462–63). See the *Miners* case also, discussed in (Parfit 2011a, 159–61; Tadros 2011b, 222). Tadros says he is not sure who invented the much-discussed *Miners* case; since there is a fact-relative ought, he really ought to know.

Jackson's Case give us no reason to prefer the Evidence-Relative View over the Fact-Relative View. Since I have suggested we have reason to prefer the Fact-Relative View over the Evidence-Relative View in the previous section, this means we have most reason to endorse the Fact-Relative View.

Before that, note that even if I am wrong, and so if cases like *Jackson's Case* do give us reason to prefer the Evidence-Relative View, we still have countervailing reason to prefer the Fact-Relative View from the previous subsection. On the balance of reasons, I prefer the Fact-Relative View. The argument I make below that *Jackson's Case* gives us *no* reason to prefer the Evidence-Relative View simply makes this verdict more secure.

The problem underlying the Fact-Relative View can be stated as follows.²¹⁰ We have two standards of what *X* has a right to, Standard 1 and Standard 2. We have at least three acts that the potential duty-bearer could perform. We then have a case in which the potential duty-bearer is unsure of what *X* has a right to on Standard 1, but knows that it has no chance of being act 1. And yet, because the other options could go so badly, we think that the verdict of Standard 2 is correct, and that *X* has a right to act 1. In *Jackson's Case*, we have Standard 1 (Fact-Relative View) and Standard 2 (Evidence-Relative View). Doctor is unsure whether Patient has a right to drugs B or C on the Fact-Relative View, but knows both could go badly. So, we think Patient has a right to drug A, as the Evidence-Relative View suggests.

Yet, we can run cases of this structure against the Evidence-Relative View. This is because the best available evidence to duty-bearers—that which fixes what rights obtain on the Evidence-Relative View—is not always going to be luminous: we are not always going to know what the Evidence-Relative View says that we have a right to. And when we do not know, there are some cases where we ought to adhere to some more risk adverse standard than the Evidence-Relative View. Consider,

Smith's Case. Patient has some condition. Doctor has three available options: drugs D, E, or F. Suppose that Doctor has to give some treatment.

²¹⁰ This formulation of the problem owes a lot to (Smith 2011, 5–7; Littlejohn 2009, 238–41).

Doctor knows that giving each of them may cure Patient. However, Doctor also knows that giving drug D may leave Patient badly off (e.g., with a bad rash). Giving drug E may leave Patient extremely badly off (e.g., with the loss of a foot). And, giving drug F may leave Patient incredibly badly off (e.g., with the loss of a hand). There is no evidence available to Doctor as to the likelihood of each treatment's success or failure. Doctor calls Colleague to ask her what to do. Colleague says that, "Drug D is not best. As to which of E or F is best..." and then the phone goes dead. Doctor tries to call back, but there's no luck. Doctor needs to give a treatment now.²¹¹

Suppose Colleague is both Doctor's epistemic superior and a very reliable testifier. Doctor has no evidence of the respective probabilities of the treatments, so the evidence does not support any of D, E, or F. However, she has excellent evidence that drug D is not best. So, Doctor knows that Patient will not have a right that she give her drug D. It is just that Doctor does not know whether it is drug E or F that is best.²¹² And yet, because either treatment could go so badly, intuitively Patient has a right to drug D—this is true, even though Doctor knows, relative to the best available evidence to her, *that* option has no chance of being the one Patient has a right to on the Evidence-Relative View.

(Here is another example (Littlejohn 2009, 238–41). Suppose there is evidence available to you as to what treatment is best, though you are not sure whether you will respond to that evidence correctly due to other pressures. Suppose you know that D is not best but are not *sure* whether E or F is best—were you to have a little more time, you could figure

²¹¹ This case is based on (Smith 2011, 5), with some details changed. For discussion, see (Zimmerman 2014, 69–76).

²¹² What *does* Patient have a right to in *Smith's Case* given the Evidence-Relative View? With a further assumption, the Evidence-Relative View arrives at the verdict that Patient has a right that Doctor give her drug E. This is because the Evidence-Relative View needs to say something about what rights obtain when potential duty-bearers have no probabilities (or insufficiently robust probabilities) concerning what to do. For example, suppose that Colleague never answered the phone in the first place. If Doctor has to prescribe some drug, Patient has a right that Doctor give her drug D, and rights against Doctor giving her drugs E and F. When faced with no probabilities, Zimmerman plausibly suggests that we ought to apply a principle of indifference, on which all of the options should be assigned equal probabilities. Given that Doctor has good evidence that drug D is not best, drug D is off the table. We assign equal arbitrary probabilities to the remaining options, and so drug E comes out as best, and so the drug Patient has a right to.

it out with more confidence. Knowing that if you get it wrong, things could go disastrously, you ought to take the conservative option that you know has no chance of being best relative to your evidence.)

Just as *Jackson's Case* presents us with an example in which the Fact-Relative View tells us that Patient has a right that Doctor give her drug B where, intuitively, Patient has a right that Doctor give her drug A, *Smith's Case* presents us with an example in which the Evidence-Relative View tells us that Patient has a right that Doctor give her drug E, where, intuitively, Patient has a right that Doctor give her drug D. If *Jackson's Case* and *Smith's Case* are analogous in all morally relevant ways, *Jackson's Case* gives us no reason to prefer the Evidence-Relative View over the Fact-Relative View.

Now, the defender of the Evidence-Relative View cannot appeal to just any old solution to *Smith's Case*. This is because the defender of the Fact-Relative View could then use that solution to solve *Jackson's Case*. What the defender of the Evidence-Relative View needs is a solution *only* to *Smith's Case*. They need to find a disanalogy between *Jackson's Case* and *Smith's Case* and draw on that disanalogy to solve *Smith's Case*.

The most obvious disanalogy between the cases is that, in *Smith's Case*, Doctor receives evidence that D is not best through testimony, whereas there is no testimony in *Jackson's Case*. Further, the testimony is not of probabilities on which she can base her judgement. Rather, it is testimony that drug D is not best. A defender of the Evidence-Relative View could say that testimony of what is best does not give you evidence of which option is best. So, Colleague's testimony that drug D is not best does not give *Doctor* evidence that drug D is not best.

However, this move has unacceptable consequences. Let us amend *Smith's Case* such that the phone does not cut out, and Colleague says "Drug D will not be best. As to whether to give either drug E or F, drug E is best." Colleague is Doctor's epistemic superior. She is a reliable testifier. It is very implausible that, on the Evidence-Relative View, this does

not make it the case that Patient has a right that Doctor give her drug E. So, testimony of what option is best must give Doctor evidence.²¹³

To conclude this subsection, *Jackson's Case* gives us no reason to prefer the Evidence-Relative View over the Fact-Relative View, as the Evidence-Relative View is susceptible to structurally analogous counterexamples. The task for the defender of the Evidence-Relative View is to find a solution to *Smith's Case* that both (i) does not lead to implausible results in other cases and (ii) is not available to a defender of the Fact-Relative View. I am not sure such an answer is available. Again, even if I am mistaken and *Jackson's Case* does speak in favour of the Evidence-Relative View, we have countervailing reasons to be sceptical of the Evidence-Relative View.

5. Conclusion

This chapter has questioned whether rights depend on duty-bearers' beliefs, the evidence available to them, or the facts. If rights do not depend on the facts, we have our solution to the Problem of Pure Risk. This is because, when it is intuitively plausible that others have rights that we not subject them to risk, often our beliefs and the evidence supports others' wellbeing being sufficiently weighty to place us under duties. I have argued that rights depend on the facts, and so we cannot solve the Problem of Pure Risk in this way.

Against the Belief-Relative View, I argued that it is unclear why the Moral Belief-Relative View would not be true in place of the Belief-Relative View, but that we have good reason to reject the Moral Belief-Relative View. Further, the Belief Relative View implausibly

²¹³ Zimmerman makes a similar move to solve *Smith's Case*, distinguishing between 'evidence *available* to someone and the evidence of which that person in fact *avails* himself of' (Zimmerman 2014, 72). Whereas he used to formulate his view in terms of evidence available to agents as we have been doing (Zimmerman 2008), he has revised his view so that it responds only to evidence that agents have availed themselves of. Yet, first, Doctor *has* availed herself of the evidence available to her in terms of Colleague's testimony. Notwithstanding this, second, Zimmerman's view is going to get our revised version of *Smith's Case* wrong. Third, it gets *Real Gun* wrong too, for Non-Believer has not availed herself of the evidence as to whether the gun is real.

over and undergenerates rights. Finally, more generally, I argued it is implausible that our rights depend on others' beliefs about the world.

Against the Evidence-Relative View, I argued that it is implausible that our rights depends on the evidence available to others and others' ability to respond to evidence. I also argued that the Evidence-Relative View implies, when new evidence becomes available, new rights come into existence rather than us gaining better evidence about which rights obtain; but this seems mistaken. And, finally, I argued that the Evidence-Relative View implausibly undergenerates rights.

All of this gives us good reason to endorse the Fact-Relative View. Yet, there remains the question what to say about *Jackson's* and *Smith's Case*. What I hope to have shown above is that *Jackson's Case* gives us no reason to prefer the Evidence-Relative View over the Fact-Relative View.

During discussion of whether the fact-relative ought can reach the verdict that Doctor ought give drug A in *Jackson's Case*, Littlejohn considers what Doctor may think to herself. He says, '*if I know I don't know whether it is drug [B] or C that is best and know that guessing could be disastrous, I ought to give drug A*' (Littlejohn 2009, 238). He builds on this, saying we might actually have a conditional, fact-relative ought to give drug A. On this view, we can see facts about uncertainty entering the picture. Similar to this, we might think that Patient has a conditional right to drug A, conditional on Doctor not knowing which of drugs B and C will fully cure Patient. However, I am unsure whether we are going to reconcile deficient evidence in cases like *Jackson's Case* with the deficient evidence of the soldiers in *Duped Soldiers*. For example, might not the soldiers think to themselves, "Given that I don't know that those people over the border whom we might attack are innocent, and given I know that not attacking could be disastrous (after all, I've great evidence they are terrorists, planning a terrible attack), we ought to attack them." I think further discussion of cases like *Real Gun*, *Duped Soldiers*, and *Jackson's Case* may point to a gap between what rights obtain and what one ought to do, where that ought has a more practical, decision-theoretic flavour. This is something I look to pursue moving forward. Before that, let us see whether we can hold rights against risk on a Fact-Relative View given the Safety Condition.

8. Safer Still

1. Introduction

This third part of the thesis focuses on the Problem of Pure Risk. Here again is our primary example:

Roulette. Target is asleep. Her housemate, Shooter, comes into her room and decides to play Russian roulette with her. Luckily, no bullet is fired. Shooter, content with having had a round of roulette, will never play roulette again.

In chapter 6, I argued risk of harm is not itself harmful. This means Shooter does not harm Target by exposing her to risk of harm. Since, nonetheless, intuitively Shooter violates Target's rights and wrongs Target, *Roulette* is an example of harmless wrongdoing. In chapter 7, I argued that rights depend on the facts, and not on duty-bearers' beliefs or the evidence available to them. So, though Shooter does not believe nor does the evidence decisively support that she will not harm Target, this does not affect Target's rights.

Since Target is not harmed by the violation of Shooter's putative duty not to play roulette with her, and since rights are determined by the facts, Target's wellbeing is not of sufficient weight to place Shooter under a duty not to play roulette with her. Given the canonical formulation of the Interest Theory, this means Target does not hold a right against Shooter that Shooter not play roulette with her. This is the Problem of Pure Risk. Like the Problem of Preemption, it is a specific version of the Problem of Harmless Wronging for the Interest Theory.

In chapter 5, I introduced the Safety Condition as a reply to the Problem of Harmless Wronging, focusing on the Problem of Preemption. The Safety Condition looks beyond what happens in the actual world to close possible worlds to ensure that people could not easily have been harmed. In this chapter, I further make the case for the Safety Condition. In section 2, I show how the Safety Condition solves the Problem of Pure Risk. With this, we see that the Safety Condition is a unified solution to the Problem of Harmless Wronging. In section 3, I say more about why rights respond to modality as the Safety Condition requires. In section 4, I defend the Safety Condition against some further objections.

2. The Safety Condition and Pure Risk

2.1 Making Safety Safer

Showing how the Safety Condition solves *Roulette* requires looking back to preemption. The Problem of Preemption arises because the right-holder is no worse off due to the violation of her right than she would otherwise have been. This is because a different event would have made the right-holder worse off to an equal or greater extent. The Safety Condition appeals to some close counterfactual world by reference to which the right-holder *is* left worse off:

For X to have a right against Y that $Y \Phi$, Y 's not Φ -ing must cause X to be worse off than she would have been in at least one close world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .

Think again of *Hitmen*, in which Hitman₂ follows Hitman₁ on every job she has, ready to complete a hit on Victim were Hitman₁ not to. In that case, rather than comparing the following two worlds:

HM World 1. Hitman₁ shoots Victim. (Hitman₂ was waiting in the wings.)

HM World 2. Hitman₁ does not shoot Victim. Hitman₂ shoots Victim.

the Safety Condition has us compare HM world 1 with

HM World 3. Hitman₁ does not shoot Victim. Hitman₂ does not shoot Victim.

Victim is sufficiently worse off in HM world 1 than she is in HM world 3 to place Hitman₁ under a duty not to shoot. This duty correlates with Victim's having a right that Hitman₁ not shoot.

When it comes to pure risk imposition, something slightly different is occurring. Like Victim in *Hitmen*, Target is no worse off than she would have been had the correlative duty been respected. But, unlike in cases of preemption (where, put most naturally, Victim is no worse off than she would have been had Hitman₁ not shot because she would have

been made *equally* worse off by Hitman₂), this is because the risked harm does not materialise. In *Roulette*, we are comparing:

R World 1. Shooter plays Russian Roulette. No bullet is fired.

R World 3. Shooter does not play Russian Roulette.

Target is no worse off in R world 1 than she is in R world 3. However, as with preemption, there is an extremely close world by reference to which Target is made worse off because of Shooter playing roulette with her—the world in which there *is* a bullet in the chamber when she pulls the trigger:

R World 2. Shooter plays Russian Roulette. Shooter fatally shoots Target.

Two quick clarifications. First, a terminological matter. The orderings of the worlds correspond to closeness with the actual world. HM world 2 is closer than HM world 3 to HM world 1, the actual world. When it comes to *Roulette*, R world 2 is closer than R world 3 to R world 1, the actual world. That is to say, it requires less of a departure from reality (less Lewisian miracles) to have a bullet in the chamber when Shooter pulls the trigger, as in R world 2, than to have Shooter not play roulette, as in R world 3.

Second, in *Plane Crash* and *Hitmen*, \mathcal{Y} Φ s in both world 2 and 3 (she acts as the duty dictates). Attendant does not deny Passenger admittance onto the plane in both world 2 and 3. Hitman₁ does not shoot Victim in both world 2 and 3. This is different from *Roulette*. \mathcal{Y} does not Φ in world 2 (she fires the gun) but does Φ in world 3 (does not fire the gun). All of this is returned to below, but, to anticipate that discussion, this difference is explained by the way in which preemption and pure risk differ as cases of harmless wrongdoing.

Returning to the Problem of Pure Risk, unlike when dealing with preemption, in which we compare world 1 with world 3 in order to locate Passenger's and Victim's being worse off as a result of the duty's being violated, in cases of pure risk imposition we need to compare world 2 with world 3. Target is worse off in R world 2 than she is in R world 3. The extent to which she is worse off is sufficient to place Shooter under a duty not to make Target the subject of her risky behaviour. The Safety Condition can accommodate this verdict. It requires recognising R world 2 as our focus when \mathcal{Y} fails to act as the duty

dictates instead of the actual world, R world 1. That is, it requires recognising R world 2 as the *world of evaluation* (to borrow some jargon from chapter 5) rather than the actual world, R world 1.²¹⁴ To make this especially explicit, we can amend the Safety Condition:

For X to have a right against Y that $Y \Phi$, Y 's not Φ -ing must cause X to be worse off in the world of evaluation than she would have been in at least one close world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .

Target is worse off in R world 2 because of Shooter playing roulette with her than she is in R world 3. The extent to which she is worse off is sufficient to place Shooter under a duty not to play roulette with her. The Safety Condition is satisfied.

In cases of preemption, our focus when Y does not act as the duty dictates (our world of evaluation) is world 1, the actual world. In *Roulette* (and cases of pure risk imposition more generally), our focus is world 2 (a world close to the actual world). It is in keeping with the Safety Condition for our focus on what happens when Y fails to act as the duty dictates to be worlds close to the actual world, rather than only the actual world itself. This is because we could easily have been in those close worlds. This is just to say that it could easily have been that there was a bullet in the chamber when Shooter pulls the trigger. While different events occur between R world 1 and R world 2 (Victim is not shot and Victim is shot), Shooter *acts* in the same way in both worlds.²¹⁵ And this is in keeping with the Safety Condition: we are unsafe to the extent that our wellbeing is not safely protected; as Shooter could easily have shot Target, Target's wellbeing was not safely protected; by allowing us to focus on worlds in which Shooter acts in the same way close to the actual

²¹⁴ This is discussed below, but to preempt that discussion: the world(s) of evaluation is the world(s) in which Y does not act as the would-be duty dictates and the world(s) of comparison is the worlds in which Y does act as the would-be duty dictates.

²¹⁵ Might one object that Shooter cannot have acted in *exactly* the same way in both worlds since something different happens in either world (viz., there is, and is not, a bullet in the chamber)? While there is something to this objection, it would force us to say that Shooter *acts* differently in the world in which she pulls the trigger, and there is a bullet in the chamber, from the world in which she pulls the trigger, and there is no bullet in the chamber. This seems strange. Whether there is a bullet in the chamber seems too far beyond Shooter's control for it to properly be described as an act of her own. Notwithstanding this, we could say: those close worlds in which Shooter acts, in the relevant sense, in the same way.

world, rather than only the actual world, Target's wellbeing is safely protected. So far so good.

At this stage, one might wonder how we determine which world to focus on as our world of evaluation when \mathcal{Y} does not act as the duty dictates. Why focus on the actual world in cases of preemption but then focus on worlds close to the actual world in cases of pure risk? Yet, we do not need to determine which of these worlds to uniquely focus on. While the Safety Condition speaks of *the* world of evaluation, this need not imply there is only one world of evaluation. Rather, there are multiple worlds of evaluation—those close worlds in which the potential duty-bearer performs the act that she may be under a duty not to perform. (This includes the actual world for the actual world is close to itself.) In *Roulette*, worlds of evaluation are close worlds in which Shooter puts a bullet in the cylinder of the gun, spins the cylinder, and pulls the trigger. In *Plane Crash*, worlds of evaluation are close worlds in which Attendant denies Passenger admittance onto the plane. In *Hitmen*, worlds of evaluation are close worlds in which Hitman₁ shoots Victim. Again, this is in keeping with the Safety Condition: it ensures Target's wellbeing is robustly protected. We do not need to determine which world to focus on when \mathcal{Y} does not act as the duty dictates since the Safety Condition looks to *all* close worlds in which the duty-bearer acts in the same way.

We could just as easily make another amendment to the Safety Condition to make explicit that there is not only one world of evaluation:

For X to have a right against \mathcal{Y} that $\mathcal{Y} \Phi$, there must be some set of close worlds in which \mathcal{Y} 's Φ -ing makes X worse off, and the extent to which she is worse off must be of sufficient weight to place \mathcal{Y} under a duty to Φ .

2.2 Preemption, Pure Risk, and Safety

Now that we have seen how the Safety Condition deals with cases of pure risk, we can better appreciate how the Safety Condition works. It makes for a two-part comparison. We have worlds of evaluation and worlds of comparison. Worlds of evaluation are worlds in which the potential correlative duty-bearer performs the action that she may be under a duty not to perform. Worlds of comparison are worlds in which the potential correlative duty-bearer does *not* perform this action. Whereas the canonical statement of the Interest

Theory focuses only on the closest worlds in which the duty is and is not respected, the Safety Condition also looks to other worlds (sufficiently) close to the closest worlds in which the duty is and is not respected.²¹⁶ And, if the potential right-holder is worse off in a world of evaluation than she is in a world of comparison, and if the extent to which she is worse off is of sufficient weight to place the potential correlative duty-bearer under a duty, the Safety Condition is satisfied, and she holds a right.²¹⁷

Above we saw that, while the world of evaluation in *Plane Crash* and *Hitmen* is world 1 (the actual world), by contrast the world of evaluation in *Roulette* is world 2 (a world close to the actual world). This is because cases of preemption and cases of pure risk involve different types of harmless wronging.

As a point of comparison, consider a mundane case in which one's rights are violated: suppose Threatener punches Innocent. Here, we compare the actual world, in which Threatener punches Innocent, with the closest counterfactual world in which Threatener does not punch Innocent. Our world of evaluation is the closest world (to the actual world) in which Threatener punches Innocent: namely, the actual world. Our world of comparison is the closest world (to the world in which Threatener punches Innocent) in which Threatener does not punch Innocent. Since Innocent is sufficiently worse off in the world in which Threatener punches her, than in the world in which Threatener does not punch

²¹⁶ To be a little more precise, worlds of evaluation are worlds (1) in which the potential correlative duty-bearer performs the action that she may be under a duty not to perform and (2) that are close to the closest world in which she acts in this way. Similarly, worlds of comparison for the Safety Condition are worlds (1) in which the potential correlative duty-bearer acts as the would-be duty requires and (2) that are close to the closest world in which she acts in this way.

²¹⁷ Again, subject to our satisfying the other necessary and sufficient conditions on rights. As a matter of pragmatics, we could begin with the closest worlds and see if the potential right-holder is sufficiently worse off for the Safety Condition to be satisfied. In effect, we begin with the canonical statement of the Interest Theory. If not, we move from the closest worlds to the next closest. We keep doing this until either the Safety Condition is satisfied or we have run out of sufficiently close worlds. A different way to go would be to begin with the closest worlds of evaluation and comparison that leave our potential right-holder sufficiently worse off to place our potential duty-bearer under a duty. We then look to whether these worlds of evaluation and comparison are sufficiently close. If they are not, we know the Safety Condition is not satisfied.

her, Threatener is under a duty not to punch Innocent correlative to Innocent having a right that Threatener not punch her.

In cases of preemption, we compare how our potential right-holder actually fares not with how she fares in the closest counterfactual world in which the potential duty is respected, but with a different counterfactual world. In cases of pure risk, we do use the closest counterfactual world in which the potential duty is respected. However, we do not use the actual world as our world of evaluation, but some world close to the actual world. So, preemption and pure risk differ from the mundane case in symmetrical ways.²¹⁸ This is laid out in the following table:

	World of Evaluation	World of Comparison
<i>Mundane Case</i>	Actual world.	Closest counterfactual world in which duty is respected.
<i>Preemption</i>	Actual world.	Not closest counterfactual world in which duty is respected, but a world close to this world in which the duty is respected.
<i>Pure Risk</i>	Not actual world, but a world close to this world in which the duty is infringed.	Closest counterfactual world in which the duty is respected.

(In this table I mean the world of evaluation and world of comparison that enables the Safety Condition to be satisfied. Strictly speaking, as I said in the opening paragraph to

²¹⁸ We could have a case of preempted pure risk. Suppose that, were Shooter not to have played roulette with Target, Shooter’s friend would have. And suppose that, were Shooter’s friend to have played roulette with Target, no bullet would have been fired. There, our world of evaluation would not be the actual world, nor would our world of comparison be the closest counterfactual world in which the potential duty is respected. Rather, we would compare the close world in which Shooter fires a bullet when she pulls the trigger to the world in which neither she nor her friend plays roulette with Target.

this subsection, worlds of evaluation are all those close worlds in which the potential correlative duty-bearer performs the action that she may be under a duty not to perform. So, in *Roulette* for example, the actual world *is* a world of evaluation; it is just that, with that as the world of evaluation, the Safety Condition is not satisfied.)

To be precise when expounding the Safety Condition, I have used some jargon. Here is a more natural gloss on the Safety Condition. For someone to hold a right against us that we not perform some action, we look to whether our performing that action could easily leave them sufficiently worse off to place us under a duty.

3. In Favour of Safety

Here is where we are at. The canonical statement of the Interest Theory falls foul of the Problem of Harmless Wronging. Further, I argued that seemingly plausible ways of responding to the problem considered in chapters 4, 6, and 7 are not wholly satisfactory. The Safety Condition solves the Problem of Harmless Wronging—it correctly generates rights in cases of preemption and pure risk while being sympathetic to the Interest Theory’s welfarist grounds of rights.²¹⁹ What is more, the Safety Condition offers us a unified account of *why* people are attributed rights across these two different types of case of harmless wrongdoing: the right-holder could easily be left worse off by the non-performance of the action required by the duty. The Safety Condition’s *principled* extensional accuracy is the primary virtue of the account that I would like to stress in this thesis.

In section 4, I respond to some objections to the Safety Condition. Before that, in this section I explain why rights respond to modality in the that way the Safety Condition prescribes. I suggest the explanations considered in subsections 3.1-2 are not found wholly satisfying. In subsection 3.3, I return to the ideas introduced in chapter 5 that rights respond to modality because this removes an objectionable form of luck from rights and makes duty-bearers sensitive to others’ wellbeing.

²¹⁹ Possibly notwithstanding some cases of Very Preempted Harm (chapter 5, section 4.2).

3.1 Protection from Harm

Endorsing the Safety Condition would have instrumental benefits. This might explain why rights respond to modality in the way the Safety Condition prescribes. First, the Safety Condition does a better job protecting right-holders' wellbeing than the canonical statement of the Interest Theory. As Lazar argues regarding a similar condition: 'If I am secure [if my wellbeing is robustly protected], then I escape harms that might otherwise have befallen me' (Lazar 2015, 12).²²⁰ To determine whether this is true, we need to compare the world in which we endorse the Safety Condition to the world in which we do not endorse the Safety Condition. If right-holders are harmed less in the world with the Safety Condition than the world without the Safety Condition, it better protects right-holders' wellbeing. And, we have reason to think this is the case. In the world without the Safety Condition, people may commit acts they think to be harmless that end up causing harm—acts that they otherwise would not have performed were others to have rights against them that they not perform those acts. For example, people may subject others to risks they think to be harmless that end up materialising into harm.²²¹

Second, that one's wellbeing is securely protected provides other instrumental benefits. For example, as said in the introduction to chapter 6, Target may become distressed, fearful, and her life disrupted were she to find out about Shooter's escapades. If we know our wellbeing is robustly protected because of the Safety Condition, we might feel generally secure from these burdens. Further, some think that knowing our wellbeing is robustly protected is necessary for us to enjoy other benefits, such as the benefit of planning our future (John 2011; Lazar 2017, 8).²²²

²²⁰ Lazar is arguing for what he calls 'Risky Killing': 'Other things equal, when A's Φ ing kills an innocent person, her action is objectively worse the higher the probability when she Φ ed that her action would kill an innocent person' (Lazar 2017, 5; 2015). Lazar motivates Risky Killing on the grounds that it makes our wellbeing more secure. For problems with Lazar's particular formulation of the view, see section 3.2 of chapter 6, note 174, and below in the text.

²²¹ For similar discussion on what we can call the *instrumental* version of the harm principle, see (Tadros 2011a; Edwards 2014). Roughly, on the instrumental version of the harm principle, criminalisation of some action, Φ , is permissible only if criminalising Φ -ing would prevent harm to people other than the actor.

²²² See chapter 6, note 17 for discussion.

However, there are at least three problems. First, if the Safety Condition was grounded in these instrumental considerations, it is difficult to see why when it comes to *harmless* wronging the duty-bearer would owe her duty to the right-holder—why she would wrong the right-holder. For, by hypothesis, the duty-bearer has not frustrated any of these goods. While generally Target may be better off if others do not subject her to risk of harm, in this particular case Shooter has not made her worse off by subjecting her to risk of harm. Second, it is not the existence of the Safety Condition that secures these instrumental benefits, but people *believing* their rights are robustly protected. But that is compatible with the Safety Condition being false as a matter of moral theory. And third, we might think these are just the wrong sorts of reasons for why we have rights in these cases of harmless wronging. So one might say: “Look, you just don’t need to work out whether having rights against having roulette played with you reduces harm in general to work out whether you have that right. Just look at the act!”²²³

Let us look elsewhere for reasons to endorse the Safety Condition.

3.2 *Freedom and Autonomy*

Recall, I argued that the two dominant views of freedom are modally demanding.²²⁴ The liberal Non-Interference View says that the extent to which an agent is free depends not only on whether she is subject to interference, but also on whether she would have been subject to interference in other possible worlds had she chosen differently. The republican Non-Domination View goes further, suggesting that the extent to which an agent is free depends both on whether our agent would have been subject to interference had she chosen differently and on additional possible worlds in which others are disposed differently towards her. Further, I said some views of autonomy (those that endorse Adequate Range) are modally demanding. For example, Oberdiek says, ‘one is autonomous when one can

²²³ For analogous complaints about instrumental versions of the harm principle, see (Ripstein 2006, 218; Duff 2007, 129; Stewart 2010, 30; Tadros 2011a, 50); *cf.* (Edwards 2014, 276).

²²⁴ Chapter 5, subsection 4.1.

plot and pursue one's own worthwhile path, and to do this, one needs to have access to a range of valuable options' (Oberdiek 2017, 9).²²⁵

Because the Safety Condition secures protection from harm across nearby worlds, it secures non-interference for people across nearby worlds. Thereby, the Safety Condition secures right-holders' freedom and, according to Adequate Range, satisfies a condition partly constitutive of autonomy. Put differently, satisfying the Safety Condition is partly constitutive of securing freedom and autonomy for right-holders.

Lazar takes this route.²²⁶ Taking these two considerations together, he says that 'I am better off if my freedom from wrongful harm does not avoidably depend on luck' (Lazar 2015, 12). He thinks this is because luck is antithetical to control, control is partly constitutive of autonomy, and '[b]eing autonomous is non-instrumentally valuable; most of us aspire to this ideal for its own sake, not because of other goods it brings us' (Lazar 2015, 12). More recently, he says 'it is non-instrumentally worse to avoid wrongful harm merely through good luck, than to do so robustly, because the more you depend on luck, the less control you have over your life, and so the less autonomous you are' (Lazar 2017, 8).

However, there are at least two reasons to be sceptical that this offers a full explanation of why to endorse the Safety Condition. First, recall there are some people for whom autonomy has no value. For example, very young children, the severely mentally disabled, and those suffering from the later stages of Alzheimer's disease. If we have reason to endorse the Safety Condition insofar as it secures autonomy, why would it matter for those for whom autonomy has no value that their wellbeing be robustly protected?²²⁷ Things are a little less clear when it comes to freedom. But I am sceptical that freedom has non-instrumental value for those for whom autonomy has no value. So I am sceptical that

²²⁵ Chapter 6, subsection 3.1.

²²⁶ Among those discussed in the previous subsection and in chapter 6, note 174.

²²⁷ See chapter 2, section 4, and chapter 6, subsection 3.2.2. Of course, one could say those for whom autonomy has no value have rights against harm, though no rights against harmless wrongs. But I find this implausible.

interfering with the freedom of those for whom autonomy has no value is bad for those people when that interference causes no harm (as in cases of harmless wrongdoing).

Second, at the beginning of chapter 3, I noted that the Will Theory says rights are exclusively grounded in the control they endow their holders with. Yet, I objected that this fails to pay attention to why at least some rights are important—in virtue of the harm that would befall one were those rights to be violated. Yet, if we say the Safety Condition is important in virtue of the fact that it secures freedom and autonomy for people, is not the Safety Condition making a similar mistake?

3.3 Sensitivity

In the previous two subsections, I considered two types of reason for why rights respond to modality in the way the Safety Condition prescribes. I suggested neither type of reason was wholly compelling. Instead, I think our best reason for endorsing the Safety Condition to support its extensional accuracy comes from the fact that the Safety Condition removes an objectionable form of luck from rights and formally requires that we as duty-bearers are sensitive to others' wellbeing. This idea was first introduced in chapter 5.

If the canonical statement of the Interest Theory is correct, duty-bearers can through sheer luck respect others' rights. For example, given that canonical statement, through sheer luck Attendant does not violate Passenger's rights. It could easily have been that Attendant did harm Passenger, and so would have violated her rights. Similarly, through sheer luck—because there is no bullet in the chamber when Shooter pulls the trigger—Shooter does not violate Target's rights. However, it could easily have been that there was a bullet in the chamber, and so Shooter would have violated Target's rights. The Safety Condition removes this form of luck from rights.

At this stage, some may wonder why it matters that rights do not depend on luck in this way. This is where our second reason comes in: the Safety Condition makes duty-bearers sensitive to others' wellbeing. (If rights did depend on luck in the way the Safety Condition rules out, duty-bearers would not be sensitive to others' wellbeing.)

When X holds a right against Y that $Y \Phi$, Y is not merely under a duty to Φ but owes that duty to X . Were Y not to Φ , Y would not merely act wrongly, but would wrong X . I

suggested that a theory of rights and directed duties should explain why this is. So far, fairly uncontroversial. Something else fairly uncontroversial is that many people think rights are a particular form of reason.²²⁸ If the Safety Condition is correct, one has reasons of this form not to treat others in certain ways in virtue of how one could easily affect them. For example, Shooter has a reason not to play roulette with Target *because* she could easily leave Target very badly off. That we do not merely have a reason not to harm others but ensure that we could not easily harm others is a plausible upshot of the importance of others' wellbeing. This explains why Shooter owes it to Target not to play roulette with her—why she would not merely act wrongly but would wrong Target.

This idea that rights ought to respond to modality in the way the Safety Condition prescribes inasmuch as it makes us sensitive to others' wellbeing is supported by the plausible thought that we ought to be diligent when it comes to how our actions affect others. For example, Tom Dougherty suggests:

If your actions might harm [others] or infringe [others'] rights, then you owe it to [them] to gather evidence about whether your actions do so [...] This is not just to say that someone who had failed to carry out these investigations would have a bad character. It is also to say that these investigations are required of them. (Dougherty 2018, 100)

The Safety Condition goes further than this, building it into our reasons as duty-bearers that we are diligent as to how our actions could easily affect others.²²⁹

²²⁸ Some people think they are reasons that are demandable, other things being equal (Cruft 2019, 33). Suppose you have most reason to go to the pub but decide not to. It is inappropriate for me to demand of you that you go to the pub. I may try and reason with you: explain that the pub is great fun, that we can have a nice drink. But compare that to a scenario in which I know you have promised someone that you will meet them at the pub. Then, I can demand of you: "You need to go to the pub—you made a promise." If the duty is directed, I think this adds that those reasons are, other things being equal, remainder inducing. When you fail to satisfy reasons corresponding to directed duties, when all else is equal, you owe an apology to the person to whom you owed the duty. The person to whom you owed the duty has special standing to blame and resent you. Some add that the reasons that correspond to directed duties have something like exclusionary weight. When I owe it to you to meet you at the pub, I have to exclude from my deliberation other reasons I may have not to go to the pub; for example, that it is cold out.

²²⁹ Dougherty attempts to motivate this duty of due diligence by suggesting, when *T* does not act with due diligence as regards *X*, she subjects *T* to a dignitary harm regardless of whether she otherwise harms *X* or violates *X*'s primary rights (Dougherty 2018, 100). As suggested earlier (chapter 6, note 174), this dignitary

At this stage, two questions present themselves. First, is there an analogous condition to the Safety Condition for undirected duties? Second, does it matter that duty-bearers' reasons are sensitive to right-holders' wellbeing only because we are epistemically limited?

On the one hand, if there is an analogous condition to the Safety Condition for undirected duties, this does not speak against the Safety Condition; in fact, on grounds of parity, it gives us more reason to endorse the Safety Condition. On the other hand, I am tempted by the thought that it is only in virtue of directed duties and rights' relationality that the Safety Condition is true. It is because one *owes* their duty to the right-holder that one is formally required to ensure that one could not easily fail to respect the grounds of the duty (in the Interest Theory's case, wellbeing). While one might well display bad character or not live up to full virtue by luckily respecting their undirected duties, one does not act wrongly; this can be explained by saying there is no one to whom one owes their undirected duty.

Let us turn to our second question of whether it is important that duty-bearers are sensitive to right-holders' wellbeing only because we are epistemically limited agents. Again, it is not obvious which way to go here. So, I again offer an argument in the alternative. On the one hand, it seems plausible that Hitman₁ owes it to Victim not to shoot her *even if* she knows Hitman₂ will shoot whatever happens. This can be explained by the modal fact that there are close possible worlds by reference to which Hitman₁ does harm Victim.²³⁰

What to say of a full knowledge variant of *Roulette*? It is hard to even conceptualise such a case—can one play roulette with someone else if one *knows* one will not shoot? Perhaps this gives us reason to think that, on the other hand, if we know we will not harm someone with our action then they do not have a right against us, even if we could easily harm them. Recall that I suggested it was implausible in *Voodoo* that Victim had a right that

harm is an upshot of our owing others duties (of due diligence in this case), so cannot be used to ground the duty.

²³⁰ In section 4.1, we examine how the stringency of rights is determined by the Safety Condition; those considerations may help when thinking about this case. Another thing that is worth bearing in mind is, if the duty-bearer knows she will not harm the right-holder and were to inform the right-holder of this, the right-holder may have no independent reason to refuse to permitting the duty-bearer performing the action they are under a duty not to perform.

Believer not stab a voodoo doll of them.²³¹ On the Interest Theorists' picture, rights obtain in virtue of the difference they would make to one's wellbeing. Believer could not make a difference to Victim's wellbeing in this case. Perhaps this is relevantly like cases in which one knows they will not harm others, though could (in the modal sense) harm them.

If we go this way and suggest that the Safety Condition mattering depends on our being epistemically limited agents, this is not to end up endorsing an Evidence-Relative View of the sort I rejected in the previous chapter. To see why, let us see how the Safety Condition fares with the three objections raised against the Evidence-Relative View. First, the most worrying of the problems: unlike the Evidence-Relative View, the Safety Condition does not undergenerate rights when duty-bearers' evidence sets are poor. Evidence affects whether others hold rights *only if* that evidence is sufficiently good for the duty-bearer to know they will not harm the right-holder.

Things are less clear when it comes to the other two problems with the Evidence-Relative View. The second objection was that the Evidence-Relative View says what rights *I* have depends on the evidence available to others and others' ability to respond to evidence; and this fails to pay attention to the importance of me, the right-holder. The third objection was that, on the Evidence-Relative View, what rights we hold changes when the evidence changes; but it is more plausible to think that we gain more evidence of what rights obtain. On this second way of understanding the Safety Condition, strictly speaking rights do depend on the evidence available to others and do change when better evidence becomes available. But evidence is playing a different sort of role on this picture. Evidence only affects rights when would-be duty-bearers have sufficiently good evidence about the modal space to *know* they will not harm the would-be right-holder. Since the Interest Theorists think rights are difference makers, perhaps it is unobjectionable that rights depend on evidence in this way. (But again, I must admit that I am not sure which way to go. If the second way of thinking about the Safety Condition does end up inheriting the problems of the Evidence-Relative View, this gives us reason to endorse the Safety Condition whatever the duty-bearers' knowledge of modal space.)

²³¹ Chapter 7, section 3.

4. In Defence of Safety (Again)

In the previous section, I completed my explanation of why it matters that our rights respond to modality in the way the Safety Condition prescribes. I suggested it matters because this removes an objectionable form of luck from rights and it makes us sensitive to others' wellbeing. In this section, I conclude our discussion of the Safety Condition by defending it against three objections. The first and second objection concern how we determine right-violations and right-ascriptions on the Safety Condition. The third objection suggests the Safety Condition overgenerates rights.

4.1 Violations

In *Roulette*, the Safety Condition says that Shooter violates Target's rights *because of* the comparison between the world in which Shooter does not play roulette with the world in which Shooter does play roulette and kills Target. However, does this not imply that Target's right to life has been violated? That is strange—no bullet was fired, and Target was left unharmed in bed. How, then, is her right to life violated?²³²

Here is one reply: if one has to look beyond the comparison between the actual world with the closest counterfactual world, we have a different type of right violation. This would imply Target's right to life has not been violated because the world of evaluation is not the actual world. Rather, say, her right that her life be robustly protected has been violated. However, going this way implies in *Hitmen* that Victim's right to life has not been violated, but her right that her life be robustly protected. This is because, in *Hitmen*, we do not look to the closest counterfactual world in which the duty is respected, but to a world close to that. This looks problematic to me. Victim has been killed by Hitman₁. It looks like her right to life has been violated.

Because of this, we should say, if the world of evaluation is the actual world, one's right against the actual conduct has been violated: since the world of evaluation in *Hitmen* is the actual world, Victim's right to life is violated by Hitman₁. But, if the world of evaluation is not the actual world, but some world close to that world, one's right against that conduct

²³² Thanks to Bodhi Melnitzer for pressing this objection.

has not itself been violated—rather, one’s right that one be safely protected from that conduct has been violated. Since the world of evaluation in *Roulette* is not the actual world, this implies that Beth’s right that her life not be robustly protected has been violated.

There is a subtler version of the problem under consideration. The Safety Condition says that Shooter violates Target’s rights *because of* the comparison between the world in which Shooter does not play roulette with the world in which Shooter does play roulette and kills Target. This implies Shooter’s right is as stringent as it would be were it certain (or more probable) that there was a bullet in the chamber. Similarly, the Safety Condition says Hitman₁ violates Victim’s rights *because of* the comparison between the world in which no one shoots Victim and Hitman₁ shoots Victim. This implies Victim’s right is as stringent as it would be were the harm to Victim not to be preempted. One might find both of these things implausible.

One option is to accept these conclusions. However, they are in tension with the following two asymmetries that I find plausible. First, other things being equal, one should stop someone who is certain to fatally shoot someone else from doing so, rather than stop Shooter from subjecting Target from a one in six chance of death. Second, other things being equal, one should stop someone from fatally shooting someone else rather than stop Hitman₁ from shooting Victim given that the harm is preempted. Yet, if Target’s and Victim’s rights are just as stringent as in harmful variants, what explains these asymmetries?

To accommodate this, we should suggest that the stringency of one’s rights depends not only on the comparison between the world of evaluation and the world of comparison, but on other modal considerations. There are several ways we could go here: we could think the stringency of one’s rights depends on whether we need to move from the closest worlds of evaluation and comparison to other worlds, on the number of worlds of evaluation and comparison that satisfy the Safety Condition, on the distance of those worlds from the closest worlds, and so on. And in *Hitmen*, since Hitman₁ could easily harm Victim, Victim has a right against Hitman₁; but, other things being equal, that right is less stringent than it would have been were the harm not preempted.

At this stage, it is worth noting that some people think there might be different measures of stringency of a duty: one duty could be stronger than another ‘in the *cost-requiring sense* if we would be morally required to bear greater burdens, if that were necessary, to fulfil this duty’ but weaker ‘in the *conflict-of-duty sense* if this duty would outweigh the other when these duties conflict’ (Parfit 2017, 369; Kamm 1985). Consider the following example, based on a case from (Kamm, n.d.):

Personal Lifeguard. Lifeguard is Client’s personal lifeguard. Suppose Lifeguard would be morally required to sacrifice their life for Client, were the need to arise. A wave threatens Client and 3 other people. Lifeguard can only save either Client or the 3.²³³

It is plausible that Lifeguard is not required to save the 3. An explanation for this is that Lifeguard can lean on the costs of saving the 3 to justify her not saving them. Yet, Lifeguard cannot lean on those costs to justify failing to save Client. So, her duty to save Client is more stringent in the cost-requiring sense than her duty to save the 3. Yet, it is also plausible that Lifeguard *may* save the 3 rather than Client. So, her duty to save the 3 is equally stringent in the conflict-of-duty sense. And so, while the duty to save Client is more stringent in the cost-requiring sense, the duty to save the 3 is equally stringent in the conflict-requiring sense.

Returning to harmless wrongdoing, since rights correlate with directed duties, it is open for us to disambiguate the stringency of rights in these two ways: we have the conflict-of-rights sense and the cost-requiring sense of the stringency of a right. To make sense of the two asymmetries mentioned above, we need only say the right not to be harmed is more stringent in the conflict-of-rights sense than the right not to be harmed when that harm is preempted and more stringent than the right not to be subjected to pure risk. And, we can explain this by appealing to the modal considerations mentioned above.

²³³ Thanks to Kamm for allowing me to use this case. She has it that Lifeguard can save either Client or 100 people, and in private communication has said she would deny that Lifeguard may fail to save Client if there were *only* 3 lives on the line.

But this leaves open whether the right not to be harmed is more stringent in the cost-requiring sense. And we could go either way on this. For what it is worth, my intuition is that it is easier to defeat a duty not to harm when that harm is preempted or less certain than when it is not. If we do go this way, and think duties not to perform harmless wrongs are less stringent in both the conflict-of-rights sense and cost-requiring sense, this also leaves open that they might differ on how much less stringent. So, rights against harmless wrongs might be a little less stringent in the cost-requiring sense but much less stringent in the conflict-of-rights sense.²³⁴

I began this subsection by arguing that Target's right to life is not violated in *Roulette*. Rather, her right that her life not be robustly protected is violated. But this does not commit us to saying Victim's right to life has not been violated in *Hitmen*. I then suggested we can (and ought to) accommodate the fact that, intuitively, rights may be less stringent when the violation of those rights would be harmless. However, by distinguishing between different senses of stringency, I made space for the possibility that those rights are *equally* stringent in some senses but not in others.

4.2 *Non-Violations*

All of the examples considered so far have been rights that have actually been violated. But, how does the Safety Condition work when ascribing rights in more mundane cases where one's rights are not actually violated?

Here is how one might think it works. Suppose we have Ann and Beth. Beth is going about her day as a respectable member of the moral community. Does Ann have a right that Beth not hit her? Yes. But one might think the Safety Condition implies otherwise. After all, we can suppose that Beth is so well disposed towards Ann that there is no close world in which she harms Ann. Since there is no close world (to the actual world) in which

²³⁴ It has surprised me how varied people's intuitions have been on these questions. Let us begin with the conflict-of-rights sense. Some people have said they think we ought to toss a coin to decide between both (i) harming someone or harming someone else when that harm is preempted and (ii) shooting someone with certainty or playing roulette with them. Some have reported only one of these intuitions. Some have denied both intuitions. There have been similarly varying intuitions when it comes to the cost-requiring sense of the stringency of a right.

Beth harms Ann, this means there is no close world by reference to which Ann has a sufficiently weighty interest to place Beth under a duty not to harm her. So, Ann has no right that Beth not harm her. This is an alarming result.²³⁵

The Safety Condition does not imply that we do not have rights against harm unless there is a close world in which we are harmed by the violation of our rights. The Safety Condition was wrongly applied above. Here is how things should work. First, we go to our world of evaluation. In *Plane Crash* and *Hitmen*, this is the actual world. In *Roulette*, this is a world close to the actual world in which Shooter puts a bullet into the cylinder, pulls the trigger, and shoots Target. Key to the identification of the world of evaluation is that it is the closest world(s) to the actual world (again, possibly inclusive of the actual world itself) in which the potential duty-bearer does not act as the potential duty dictates. In our three cases, where rights have actually been violated, this is either the actual world or some world close to the actual world. But, in cases in which one's rights are not actually violated, the world of evaluation cannot be the actual world—if it was, the right would be violated! Rather, it is going to be the closest world, to the actual world, in which the duty is being violated. For example, in our case of law and morality abiding Ann and Beth, we go to the closest world, to the actual world, in which Beth hits Ann. We now have our world of evaluation. This world may itself be quite far away from the actual world.

We need then to compare our world of evaluation, in which the potential duty is violated, with our world of comparison. The world of comparison is the closest world in which the potential duty is respected.²³⁶ We see if the potential right holder is worse off to enough

²³⁵ See the relevant discussion from note 78, chapter 3.

²³⁶ There are two ways we can go here that will have extensional differences in some cases. First, we can go to (1) the closest world, *to the actual world*, in which the potential duty is respected. In cases in which one's rights are not being violated this will always be the actual world. (Why? Because the closest world to the actual world is the actual world, and the duty is being respected in the actual world. See (Lewis 1973, 29–30) for discussion.) Second, we can go to (2) the closest world *to the world in which the potential duty is being violated* in which the duty is respected. Often, (1) and (2) will be the same. But they need not be. Suppose that Colin is in Canada and I am in St Andrews. We are wondering whether Colin has a right against me that I not hit him. The closest world in which I violate the potential duty not to hit him is one in which I am in Canada, and hit him (of course, it could be the world in which Colin is in St Andrews and I hit him). Now, the closest world, to the world in which I am in Canada and hit Colin, in which I do not hit Colin is one in which I am in Canada, but do not hit Colin. It is *not* the actual world, in which I am in St Andrews and do not hit him. But the closest world to the actual world in which I do not hit Colin is the actual world since,

of an extent in the world of evaluation than she is in the world of comparison to place the potential duty-bearer under the duty. So, we go to the closest world, to the world in which Beth hits Ann, in which Beth does not hit Ann—we see if Ann is worse off to enough of an extent to place Beth under a duty not to hit her.

We then iterate this process, looking at comparisons between our world of evaluation and other close worlds of comparison. We also see how other worlds of evaluations fare.²³⁷ This is why, even if there is no close world to the actual world in which Beth harms Ann, Ann still has a right that Beth not harm her.

4.3 Overgenerations

At the end of chapter 5, I introduced a problem that I did not address. Whereas the Problem of Very Preempted Harm questioned whether the Safety Condition will undergenerate rights, one might also worry it will overgenerate rights. In chapter 5, I delayed dealing with the problem because it was not yet at its worst. In that chapter, we focused on how the Safety Condition looks to worlds other than the closest counterfactual world in which the duty is respected. The worry was that there are many close worlds. We can now see why the problem has gotten worse following the discussion in this chapter: not only are there many close worlds of comparison, but there are also many close worlds of evaluation—close worlds in which the potential duty-bearer does not act as the duty-dictates. All the Safety Condition requires is that there is some close world of evaluation in which the right-holder fares sufficiently worse than in some close world of comparison. Is this not going to be too easy to satisfy?

Before turning to specific examples in which it seems the Safety Condition overgenerates, in subsection 4.3.1 I argue the Problem of Overgeneration is not that, because the Safety Condition overgenerates rights, this leads to a proliferation of rights, and proliferations of

by hypothesis, I do not hit Colin. So, (1) and (2) come apart. Since the world in which I am in St Andrews and do not hit Colin is not close to the world in which I am in Canada and hit Colin, this implies Colin does not have a right that I not hit him. So, we should go with way (2). (Offhand, some people have reported they do not find it odd that Colin does not have a right that I not hit him if I am in St Andrews, and he is in Canada; presumably, they think Colin would have a right were we together.)

²³⁷ See note 217 on how this is probably a matter of pragmatics.

rights are themselves problematic.²³⁸ I also argue the problem is not that this overgeneration leads to conflicts of rights, and conflicts are themselves problematic. Rather, the problem is that the Safety Condition may lead to counterintuitive rights. This we tackle in subsections 4.3.2-3.

4.3.1 *Why Proliferations and Conflicts Are Not Themselves Worrying*

One reason why we should not worry about proliferations of rights *themselves* is that most theories of rights already imply the existence of many, many rights. For example, Thomson worries if we have rights against risk of harm, this leads to a proliferation of rights (Thomson 1990, 244–46). But she endorses what she calls *The Means Principle for Rights*, on which ‘*if* (i) X has a claim against Y that Y not do beta, and (ii) if Y does alpha then he or she will thereby do beta, *then* X has a claim against Y that Y not do alpha’ (Thomson 1990, 157). As Zimmerman remarks, ‘it follows [from this] that I have a right [...] that you not stick a knife one inch into my back, a right that you not stick a knife two inches into my back, indeed an infinite number of such rights’ (Zimmerman 2008, 82). One might deny The Means Principle for Rights. But it is my hunch that there will be some transmission principle of this sort that leads to many, many rights.²³⁹

Notwithstanding the above, suppose one’s theory of rights is objected to because it leads to a proliferation of rights. One replies by listing every right that obtains on one’s theory. For every right, one asks “Is this right *alone* implausible?” Suppose the answer is “Yes.” Well, the problem is not that the view leads to a proliferation of rights but that the theory of rights has overgenerated, leading to an implausible right. Now suppose the answer is “No.” And suppose that we are delivered the same answer for every right. What, then, is the objection meant to be? “Your theory generates many rights, none of which individually are implausible, but which together are implausible.” This objection seems odd.²⁴⁰

²³⁸ See, e.g., (Sumner 1987; Wellman 1999).

²³⁹ For example, see the discussion in chapter 1, section 4.3, where I argued we can speak about rights at differing levels of specificity.

²⁴⁰ Perhaps the worry is that rights, as a normative phenomenon, are especially important and their proliferation devalues that importance. So, even if no particular right is implausible, the combination of all those rights is implausible. This is an interesting argument. I am inclined to think, while rights themselves are an

Perhaps the worry is that if there is a proliferation of rights, this leads to *conflicts* of rights. And, it is conflicts of rights that are implausible. But it is not obvious that a conflict of rights is worrying, either. After all, we are assuming the

Pro Tanto View. If X has a claim-right against Y that $Y \Phi$, Y has a pro tanto duty to Φ .²⁴¹

Given this view, X can both have a right against Y that $Y \Phi$ and Z , for example, have a right against Y that Y not Φ . This can be the case even though X 's and Z 's rights conflict. The Pro Tanto View mandates only that Y has a pro tanto duty to Φ and a pro tanto duty not to Φ . This does not mean she is both required all-things-considered to Φ and not to Φ .

But I did not spend too long defending the Pro Tanto View. So, suppose one endorses the

All-Things-Considered View. If X has a claim-right against Y that $Y \Phi$, Y is required to Φ .

Then, if we have a conflict of rights, we might be in trouble. We are especially in trouble if we do not admit moral dilemmas (where Y is required all-things-considered both to Φ and not to Φ) or if we have an example of a conflict that does *not* look like a moral dilemma. (Our examples of Ben needing to cut across Ann's land to save his life and Bystander needing to turn the trolley to save our sufficiently large number of people are two examples that do not look like moral dilemmas.)

But defenders of the All-Things-Considered View already have mechanisms in place to deal with putative conflicts: they think the content of one of the conflicting rights *specifies* that the right does not obtain in those circumstances.²⁴² Moral Specificationists think that X only has a claim-right that Y not *unjustifiably* or *wrongfully* Φ . Factual Specificationists

important normative phenomenon, this does not stop some rights from being fairly trivial. Those rights have the important normative upshots that rights do (discussed in chapter 1, section 1), but they are not very weighty in that instance.

²⁴¹ Chapter 1, section 4.4.

²⁴² Chapter 1, section 4.4.

think that X only has a claim-right that $Y \Phi$, unless: list of exemptive clauses. And if this machinery is already in place, we can use it to dissolve the potential conflicts the Safety Condition generates.

So, if the Safety Condition overgenerates rights this is not problematic because it will lead to a proliferation of rights, nor that it will lead to conflicts of rights. It will be problematic because of the implausible nature of the overgenerated rights themselves. So what do these implausible rights look like?

4.3.2 Ruling Out Worlds

There is a close world, through comparison with which, Passenger is made worse off by being denied admittance onto the plane. The Safety Condition is satisfied, so Passenger has a right against being denied admittance onto the plane.²⁴³ However, there is also a close world, through comparison with which, Passenger is made *better off* by being denied admittance onto the plane. Namely, through comparison of

(*World 1*) Attendant denies Passenger admittance onto the plane. The plane takes off and crashes killing everybody on board.

And

(*World 2*) Attendant does not deny Passenger admittance onto the plane. The plane takes off and crashes killing everybody on board.

Does this imply that Passenger has a right against Attendant that Attendant *deny* Passenger admittance onto the plane? It appears that though—the Safety Condition has been satisfied. This is one example of what we can call the Problem of Overgeneration.

Consider another example.

A&E. An unconscious Patient comes into A&E with a burst aneurysm. There is no chance that the aneurysm will stop bleeding spontaneously

²⁴³ Again, subject to our satisfying the other necessary but insufficient condition on right-ascriptions.

and so, without treatment, Patient will surely die. Surgeon can operate.
The surgery is very serious, though all-things-considered beneficial.

While the case stipulates that there is no nomologically possible world in which Patient recovers without treatment (by this I mean, a possible world with our laws of nature), there might nonetheless be a metaphysically close world in which the aneurysm stops bleeding. On the Lewisian view, all that would be required is a small miracle to stop the aneurysm from bleeding. Through comparison with this world, Patient would be worse off as a result of the operation: she will unnecessarily have gone through a very serious operation. Does this imply that Patient has a right against Surgeon performing the operation *because* there is a metaphysically close, though nomologically impossible, world in which the bleeding would stop?

I think the Problem of Overgeneration can be resisted. In what follows, I do so on several fronts. The remainder of this subsection is taken up with a first general strategy of ruling out worlds that would otherwise allow the Safety Condition to overgenerate rights by appealing to other features of closeness. Then, in the following subsection, a second general strategy makes appeal to other considerations on rights that rules out further overgenerations.

First, we could move again to the

Interest Theory (Safety, Relevance Variant). For X to have a right against Y that $Y \Phi$, Y 's not Φ -ing must cause X to be worse off than she would have been in at least one relevant world, and the difference in X 's wellbeing must be of sufficient weight to place Y under a duty to Φ .

Since there is no stipulation that close worlds are relevant, we need not worry that world 2 is very close in *Plane Crash* since we can say world 2 is not relevant. Similarly, we need not worry about the metaphysically close world in which Patient's burst aneurysm stops bleeding. The Problem of Overgeneration is solved. However, the problem from chapter

5 with the Relevance Variant has not miraculously solved itself: how do we determine which worlds are relevant?²⁴⁴

Second, we may introduce what we can call the *Realism Condition*, according to which only those worlds that are nomologically possible count as close for rights.²⁴⁵ In *A&E*, the metaphysically possible world in which the aneurysm stops bleeding is not close, so the Safety Condition will not be satisfied. Building the Realism Condition into our view of closeness says that for a world to be close it must be nomologically possible. Alternatively, the Realism Condition could be a separate necessary but insufficient condition on rights. This admits that merely metaphysically possible worlds *are* close, though are not relevant for rights.

The Realism Condition takes us only so far. It is of little use in *Plane Crash*. Another feature of closeness that helps us deal with cases like *Plane Crash* is that the closeness relation is what we can call nonreciprocal—roughly, just because w_1 is close to w_2 when w_1 is our focus does not entail that w_2 is close to w_1 when w_2 is our focus.²⁴⁶ For example, though if the plane crashes there *is* a close world in which the plane does not crash, it is often *not* the case that, if the plane does not crash, there is a close world in which the plane does crash. Roughly, this is because a lot needs to go wrong for a plane *to* crash—meaning the world in which it does crash is far away. But, it does not take a lot for a plane *not* to crash—meaning that the world in which the plane does not crash is close.²⁴⁷ While this does not

²⁴⁴ Though, the considerations extended in the following subsection help on this front.

²⁴⁵ This condition is based upon (McMahan 2002, 133–36).

²⁴⁶ For discussion, see (Lewis 1973, 50–52). He says ‘it can happen that j is more similar than k to i in the respects of comparison that are important at i ; k is more similar than i to j in the respects of comparison that are important at j ; yet i is more similar to j to k in the respects of comparison that are important at k ’ (Lewis 1973, 51). There, Lewis is emphasising that different facts about a world could determine which worlds are close, and those facts could differ across worlds. For example, colour could be very important at i , meaning j is closer than k to i because j is more similar in colour to i than k is to i . But colour could be moderately important at j and not particularly important at k . The point I am emphasising is slightly different, though Lewis’s point could also help us out with some examples of the Problem of Overgeneration.

²⁴⁷ To be incredibly crude, suppose that ten separate things need to go wrong for a plane to crash. Suppose the plane does crash (call this “Crash World”). There is a close world in which the plane does not crash for one of these things does not go wrong (call this “Almost Crash World”). Now, Crash World is close to Almost Crash World—it requires only one miracle to take us between those worlds. But, among the many possible worlds in which the plane does not crash are those in which none of the ten things go wrong (call

explain why Passenger does not have a right that *she* be denied admittance onto the plane (since in that case, there is a close world in which the plane crashes), it does explain why, generally, *we* do not have rights to be denied admittance onto planes—usually, there is no close world in which the plane crashes. The Problem of Overgeneration is not as worrying as it might have seemed.

So far, the general strategy has been to rule out worlds that would otherwise allow the Safety Condition to overgenerate rights by appealing to other features of closeness. However, none of these considerations (save for the Relevance Variant) explain why Passenger does not have a right that she be denied admittance onto the plane. A final way to resist the Problem of Overgeneration that adheres to this general strategy is to rethink our view of closeness. In the remainder of this subsection, I introduce a normalcy view of closeness. However, this move is only illustrative—I do not mean to imply we *should* move to normalcy (in fact, I introduce a problem with it).

Instead of assuming a Lewisian view of similarity, let us build on a recent account of normalcy (Smith 2016, 38–45). According to that view, we discover the closeness of some worlds at some time *t* through the call for explanation that is required if we to move from one world to the other. On this view of closeness, in *Plane Crash* world 2 is not a particularly close world to world 1. This is because the plane crashing would call for explanation. Whereas, world 3—the world in which Passenger is not denied admittance onto the plane and the plane lands at its destination—would be very close to our world of evaluation. The plane landing at its destination would not call for explanation. Given this, we can explain why Passenger does not have a right that she *be* denied admittance onto the plane.

However, this view of closeness causes problems in other cases.²⁴⁸ For example, in *Hitmen*, HM world 3 is far away from HM world 1 on this understanding of closeness because it *would* call for explanation. “What happened there? Why didn’t Hitman₂ shoot after Hitman₁ didn’t? They’re always following each other around.” So, this normalcy view of

these “Couldn’t Crash Worlds”). Crash World is not particularly close to any of Couldn’t Crash Worlds—ten miracles would be needed to get us between those worlds.

²⁴⁸ This problem was discussed in chapter 4, note 114.

closeness does not deal with all cases. However, this discussion has offered us an example of how we might hope to stem the Problem of Overgeneration by amending our view of closeness. We could be hopeful that there is an account of closeness out there that will solve all our problems. However, in lieu of having that account to hand, let us go back to assuming our Lewisian view of closeness.

4.3.3 Other Considerations

I have argued the Problem of Overgeneration is less severe than first thought. But, save moving to the Relevance Variant or Normalcy view of closeness, I have not been able to explain why Passenger does not have a right that she be denied admittance onto the plane in a satisfying way. To explain this, we need to turn to another feature of the Safety Condition—that it is a necessary but insufficient condition on rights. This means that X might not have a right to Φ even if there is a close world in which Y 's Φ -ing leaves X much better off. In this subsection we turn to the other necessary conditions and considerations that explain why Passenger does not have a right that Attendant deny her admittance onto the plane.

First, when introducing the Interest Theory (Canonical) in chapter 3, I said it is silent on the relation between the duty bearer's not acting as the duty requires and the harm (or failure to benefit) to the right-holder. In formulating the Safety Condition, I did take a stand on this.

For X to have a right against Y that Y Φ , Y 's not Φ -ing must cause X to be worse off in the world of evaluation...

That Y needs to *cause* X to be worse off seems to get the cases right: Attendant causes Passenger to be worse off through being denied admittance onto the plane were the plane not to have crashed; Hitman₁ causes Victim to be worse off were Hitman₂ not present; and, Shooter causes Target to be worse off were the risk to materialise. But, were Attendant to allow Passenger onto the plane, we would not say that she *causes* Passenger to

be worse off.²⁴⁹ Given this, the Safety Condition is not satisfied, so Passenger will not have a right that Attendant deny her admittance onto the plane.

Now, there may be problems with this causal restriction on the relationship between T 's Φ -ing and how X fares as there might be cases in which T 's not Φ -ing does not cause X to be worse off. Without a theory of causation on the table, we are not in a position to fully assess this issue. Due to space, let me introduce one last set of considerations that will also limit the Problem of Overgeneration.

Recall also from chapter 3 that Interest Theorists might accept any combination of the following non-exhaustive list of moral asymmetries:

Doctrine of Foreseeability. Other things being equal, it is harder to justify T 's harming X when the harm is (or reasonably should have been) foreseen by T than when the harm is not (or reasonably need not have been) foreseen by T .

Doctrine of Double Effect (DDE). Other things being equal, it is harder to justify T intending to harm X than it is to justify T merely foreseeing that her action will harm X .

Doctrine of Doing/Allowing (DDA). Other things being equal, it is harder to justify T doing harm to X than it is to justify T allowing harm to X .

In making DDE internal to rights, for example, we would say that, other things being equal, if T intends the harm that would befall X were she to violate the putative duty, it is more likely to warrant a right than if that harm were merely foreseen. This means that a harm to X caused by T might be sufficient to warrant a right if that harm is intended by T , though it might be insufficient if it was an unintended but foreseen consequence of her action. The same story could be told for Foreseeability and DDA.

²⁴⁹ This is not because omissions are not causal. While there is some disagreement on this, it is generally taken as a desiderata of a theory of causation that it can make room for omissions (Lewis 1986a, 189–93; Paul and Hall 2013, 173–214). If our theory of causality does not permit causation through omission, one will need to rethink the causation strategy developed here as it would otherwise rule out positive rights.

While all three conditions are *ceteris paribus* and stated fairly generally, endorsing any combination of them helps resist the Problem of Overgeneration. For example, suppose that Attendant were not to deny Passenger admittance onto the plane and the plane were to crash. Given the way that the case is stipulated, it is unlikely that Attendant foresees, intends, or does (that is, her action counts as a *doing* in the senses relevant DDA) the harm that would befall Passenger when the plane crashes. This helps explain why Passenger fails to have a right that Attendant deny her admittance onto the plane. (That is, notwithstanding my suggestion above that the Safety Condition is not satisfied in this variant in any case because Attendant would not *cause* Passenger to be worse off were she to allow her onto the plane).

Across these two subsections, I have offered a range of reasons why rights will not overgenerate given the Safety Condition. We may move to the Relevance Variant. We can endorse the Realism Condition. I have noted that closeness is non-reciprocal and suggested we may rethink our view of closeness altogether. And I have raised other considerations that may fail to be satisfied on rights.

5. Conclusion

Across chapters 5 and 8, I have introduced the Safety Condition as a solution to the Problem of Harmless Wronging for the Interest Theory. A natural gloss on the Safety Condition is, for someone to hold a right against us that we not perform some action, we look to whether our performing that action could easily leave them sufficiently worse off to place us under a duty. I have argued that the Safety Condition fares extensionally very well. I further suggested that the reason why rights respond to modality in the way the Safety Condition prescribes is that this removes an objectionable form of luck from rights and makes duty-bearers sensitive to others' wellbeing: we are formally required by rights on this picture not only not to harm people, but to ensure that we could not easily harm them.

There are some points that deserve more thought moving forward. First, we need to determine what to say about *Jackson's Case*, while being mindful of accommodating *Real Gun* and *Duped Soldiers*. Perhaps what further analysis will show is that the relationship between

rights and all-things-considered requirements is more complicated than one might have thought. Second, I laid out a lot of the logical space when it comes to determining the stringency of rights on the Safety Condition; it would be helpful to examine further what to say here to decide between the options. Third, it would be interesting to determine whether there is a condition analogous to the Safety Condition for undirected duties; if not, this will neatly tie into rights' relationality. Fourth, it would good to ask more about the relationship between modal facts and our knowledge of those facts—does what could easily have been the case matter when we know that this would not have been the case?

But these are tasks for another day.

Bibliography

- Adler, Matthew. 2003. 'Risk, Death and Harm: The Normative Foundations of Risk Regulation'. *Minnesota Law Review*, 1293–1445.
- Arneson, Richard J. 2001. 'Against Rights'. *Philosophical Issues* 11: 172–201.
- Bedi, Sonu. 2009. *Rejecting Rights*. New York: Cambridge University Press.
- Bennett, Jonathan. 1995. *The Act Itself*. Oxford: Clarendon Press.
- Berker, Selim. 2017. 'The Unity of Grounding'. *Mind* 127 (507): 729–77.
- Berlin, Isaiah. 1969. *Four Essays on Liberty*. Oxford: Oxford University Press.
- Boonin, David. 2014. *The Non-Identity Problem and the Ethics of Future People*. Oxford: Oxford University Press.
- Bowen, Joseph. Forthcoming. 'Addressing the Addressive Theory of Rights'. *Journal of Applied Philosophy*.
- . n.d. 'The Interest Theory of Rights at the Margins: Posthumous Rights'. In *Festschrift for Matthew H. Kramer*, edited by Visa A. J. Kurki and Mark McBride. Oxford University Press.
- Bradley, Ben. 2009. *Well-Being and Death*. Oxford University Press.
- . 2012. 'Doing Away with Harm'. *Philosophy and Phenomenological Research* 85 (2): 390–412.
- Carter, Ian. 1999. *A Measure of Freedom*. Oxford: Oxford University Press.
- Conee, Earl, and Richard Feldman. 2004. *Evidentialism: Essays in Epistemology*. Oxford, New York: Oxford University Press.
- Cornell, Nicolas. 2015. 'Wrongs, Rights, and Third Parties'. *Philosophy & Public Affairs* 43 (2): 109–43.
- Crisp, Roger. 2006. 'Hedonism Reconsidered'. *Philosophy and Phenomenological Research* 73 (3): 619–45.
- Cruft, Rowan. 2004. 'Rights: Beyond Interest Theory and Will Theory?' *Law and Philosophy* 23 (4): 347–97.
- . 2010. 'On the Non-Instrumental Value of Basic Rights'. *Journal of Moral Philosophy* 7 (4): 441–61.
- . 2013a. 'Introduction (Symposium: Rights and the Direction of Duties)'. *Ethics* 123 (2): 195–201.
- . 2013b. 'Why Is It Disrespectful to Violate Rights?' *Proceedings of the Aristotelian Society* 113 (2): 201–24.
- . 2017. 'The Circularity of the Interest and Will Theories of Rights'. In *New Essays on the Nature of Rights*, edited by Mark McBride. London: Hart Publishing.
- . 2019. *Human Rights, Ownership, and the Individual*. Oxford: Oxford University Press.
- Cruft, Rowan, Matthew Liao, and Massimo Renzo. 2015. 'Philosophical Foundations of Human Rights: An Overview'. In *Philosophical Foundations of Human Rights*, edited by Rowan Cruft, Matthew Liao, and Massimo Renzo, 1–42. Oxford: Oxford University Press.
- Darwall, Stephen. 2013. 'Bipolar Obligation'. In *Morality, Authority, and Law*, edited by Stephen Darwall, 20–39. Oxford: Oxford University Press.
- Dougherty, Tom. 2018. 'Affirmative Consent and Due Diligence'. *Philosophy & Public Affairs* 46 (1): 90–112.
- Duff, Antony. 2007. *Answering for Crime: Responsibility and Liability in the Criminal Law*. Oxford: Hart.

- Dworkin, Ronald. 1978. *Taking Rights Seriously*. Cambridge, Mass: Harvard University Press.
- Edwards, James. 2014. 'Harm Principles'. *Legal Theory* 20 (4): 253–85.
- Fabre, Cécile. 2000. *Social Rights Under the Constitution: Government and the Decent Life*. Oxford: Oxford University Press.
- . 2008. 'Posthumous Rights'. In *The Legacy of H.L.A. Hart: Legal, Political, and Moral Philosophy*, edited by Matthew H. Kramer, Claire Grant, Ben Colburn, and Antony Hatzistavrou. Oxford: Oxford University Press.
- . 2012. *Cosmopolitan War*. Oxford: Oxford University Press.
- Feinberg, Joel. 1980. 'The Nature and Value of Rights'. In *Rights, Justice, and the Bounds of Liberty*, edited by Joel Feinberg, 159–84. Princeton, NJ: Princeton University Press.
- . 1984. *Moral Limits of the Criminal Law. Volume 1, Harm to Others*. Oxford: Oxford University Press.
- . 1986. 'Wrongful Life and the Counterfactual Element in Harming'. *Social Philosophy and Policy* 4 (1): 145–78.
- . 1988. *Moral Limits of the Criminal Law. Volume 4, Harmless Wrongdoing*. Oxford: Oxford University Press.
- Feit, Neil. 2015. 'Plural Harm'. *Philosophy and Phenomenological Research* 90 (2): 361–88.
- . 2017. 'Harming by Failing to Benefit'. *Ethical Theory and Moral Practice*, September. <https://doi.org/10.1007/s10677-017-9838-6>.
- Ferretti, Maria P. 2016. 'Risk Imposition and Freedom'. *Politics, Philosophy & Economics* 15 (3): 261–79.
- Fine, Kit. 1975. 'Critical Notice of Lewis, Counterfactuals'. *Mind* 84 (335): 451–58.
- Finkelstein, Claire. 2003. 'Is Risk a Harm?' *University of Pennsylvania Law Review* 151 (3): 963–1001.
- Flanigan, Jessica. 2019. 'Duty and Enforcement'. *Journal of Political Philosophy* 27 (3): 341–62.
- Foot, Philippa. 1967. 'The Problem of Abortion and the Doctrine of Double Effect'. *Oxford Review* 5: 5–15.
- . 2002. 'Killing and Letting Die'. In *Moral Dilemmas*. Oxford: Oxford University Press.
- Frowe, Helen. 2014. *Defensive Killing*. Oxford: Oxford University Press.
- . 2015. 'Claim Rights, Duties, and Lesser-Evil Justifications'. *Aristotelian Society Supplementary Volume* 89 (1): 267–85.
- . 2019. 'If You'll Be My Bodyguard: Agreements to Save and the Duty to Minimize Harm'. *Ethics* 129 (2): 204–29.
- . n.d. 'The Duty to Save and the Duty to Minimise Harm'.
- Gardner, Molly. 2017. 'When Good Things Happen to Harmed People'. *Ethical Theory and Moral Practice*, October, 1–16.
- Gilbert, Margaret. 2018. *Rights and Demands: A Foundational Inquiry*. Oxford: Oxford University Press.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: Oxford University Press.
- . 2008. *On Human Rights*. Oxford: Oxford University Press.
- Hanna, Nathan. 2016. 'Harm: Omission, Preemption, Freedom'. *Philosophy and Phenomenological Research* 93 (2): 251–73.
- Hanser, Matthew. 1999. 'Killing, Letting Die and Preventing People From Being Saved'. *Utilitas* 11 (3): 277–95.

- . 2008. ‘The Metaphysics of Harm’. *Philosophy and Phenomenological Research* 77 (2): 421–50.
- . 2009. ‘Harming and Procreating’. In *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, edited by Melinda A. Roberts and David T. Wasserman, 179–99. Dordrecht: Springer Netherlands.
- . 2012. ‘Wrongness of Killing and the Badness of Death’. In *The Oxford Handbook of Philosophy of Death*, edited by Ben Bradley, Fred Feldman, and Jens Johansson. Oxford: Oxford University Press.
- Harman, Elizabeth. 2004. ‘Can We Harm and Benefit In Creating?’ *Philosophical Perspectives* 18 (1): 89–113.
- . 2009. ‘Harming as Causing Harm’. In *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, edited by Melinda A. Roberts and David T. Wasserman, 137–54. International Library of Ethics, Law, and the New Medicine. Dordrecht: Springer Netherlands.
- Hart, H. L. A. 1955. ‘Are There Any Natural Rights?’ *The Philosophical Review* 64 (2): 175–191.
- . 1982. *Essays on Bentham: Studies in Jurisprudence and Political Theory*. Oxford: Clarendon.
- Heathwood, Chris. 2005. ‘The Problem of Defective Desires’. *Australasian Journal of Philosophy* 83 (4): 487–504.
- . 2016. ‘Desire-Fulfillment Theory’. In *The Routledge Handbook of Philosophy of Well-Being*, edited by Guy Fletcher, 135–47. London: Routledge.
- Herington, Jonathan. 2018. ‘The Contribution of Security to Well-Being’. *Journal of Ethics and Social Philosophy* 14 (3).
- Hohfeld, Wesley Newcomb. 1919. *Fundamental Legal Conceptions as Applied in Judicial Reasoning*. New Haven, CT: Yale University Press.
- Horton, Joe. 2017. ‘The All or Nothing Problem’. *The Journal of Philosophy* 114 (2): 94–104.
- Hurka, Thomas. 1993. *Perfectionism*. Oxford: Oxford University Press.
- Jackson, Frank. 1991. ‘Decision-Theoretic Consequentialism and the Nearest and Dearest Objection’. *Ethics* 101 (3): 461–82.
- James, Susan. 2003. ‘Rights as Enforceable Claims’. *Proceedings of the Aristotelian Society* 103: 133–47.
- John, Stephen. 2011. ‘Security, Knowledge and Well-Being’. *Journal of Moral Philosophy* 8 (1): 68–91.
- Kagan, Shelly. 1989. *The Limits of Morality*. Oxford: Oxford University Press.
- . 2009. ‘Well-Being as Enjoying the Good’. *Philosophical Perspectives* 23 (1): 253–272.
- . 2011. ‘Do I Make a Difference?’ *Philosophy & Public Affairs* 39 (2): 105–41.
- . 2014. ‘An Introduction to Ill-Being’. In *Oxford Studies in Normative Ethics, Volume 4*, edited by Mark Timmons, 261–88. Oxford: Oxford University Press.
- . 2016. ‘What’s Wrong with Speciesism?’ *Journal of Applied Philosophy* 33 (1): 1–21.
- Kamm, F. M. 1985. ‘Supererogation and Obligation’. *The Journal of Philosophy* 82 (3): 118–38.
- . 2007. *Intricate Ethics*. Oxford: Oxford University Press.
- . n.d. ‘Duties That Become Supererogatory or Forbidden?’ In *Festschrift for Derek Parfit, Volume 1*, edited by Jeff McMahan. Oxford University Press.
- Klocksiam, Justin. 2012. ‘A Defense of the Counterfactual Comparative Account of Harm’. *American Philosophical Quarterly* 49 (4): 285–300.

- Kramer, Matthew H. 2000. 'Rights Without Trimmings'. In *A Debate Over Rights: Philosophical Enquiries*, edited by Matthew H. Kramer, N. E. Simmonds, and Hillel Steiner, 7–112. Oxford: Oxford University Press.
- . 2001. 'Getting Rights Right'. In *Rights, Wrongs and Responsibilities*, edited by Matthew H. Kramer, 28–95. Basingstoke: Palgrave.
- . 2003. *The Quality of Freedom*. Oxford: Oxford University Press.
- . 2010. 'Refining the Interest Theory of Rights'. *The American Journal of Jurisprudence* 55 (1): 31–39.
- . 2013. 'Some Doubts about Alternatives to the Interest Theory of Rights'. *Ethics* 123 (2): 245–63.
- . 2017. 'In Defence of the Interest Theory of Right-Holding: Rejoinders to Leif Wenar on Rights'. In *New Essays on the Nature of Rights*, edited by Mark McBride, 49–83. London: Hart Publishing.
- Kramer, Matthew H., and Hillel Steiner. 2007. 'Theories of Rights: Is There a Third Way?' *Oxford Journal of Legal Studies* 27 (2): 281–310.
- Lazar, Seth. 2015. 'Risky Killing and the Ethics of War'. *Ethics* 126 (1): 91–117.
- . 2017. 'Risky Killing: How Risks Worsen Violations of Objective Rights'. *Journal of Moral Philosophy*.
- Lewis, David K. 1973. *Counterfactuals*. Oxford: Basil Blackwell.
- . 1986a. 'Causation'. In *Philosophical Papers. Vol. 2*, 159–72. Oxford: Oxford University Press.
- . 1986b. 'Counterfactual Dependence and Time's Arrow'. In *Philosophical Papers. Vol. 2*, 32–51. Oxford: Oxford University Press.
- Liao, S. Matthew. 2010. 'Agency and Human Rights'. *Journal of Applied Philosophy* 27 (1): 15–25.
- Littlejohn, Clayton. 2009. 'Critical Notice of Michael Zimmerman's, Living with Uncertainty'. *Philosophical Books* 50 (4): 235–247.
- MacCormick, Neil. 1977. 'Rights in Legislation'. In *Law, Morality, and Society: Essays in Honour of H. L. A. Hart*, edited by P. M. S. Hacker and Joseph Raz, 189–209. Oxford: Clarendon Press.
- . 1982. 'Children's Rights: A Test-Case for Legal Rights'. In *Legal Right and Social Democracy: Essays in Legal and Political Philosophy*, 154–66. Oxford: Clarendon Press.
- May, Simon Căbulea. 2012. 'Moral Status and the Direction of Duties'. *Ethics* 123 (1): 113–28.
- . 2017. 'Desires, Interests and Claim-Rights'. In *New Essays on the Nature of Rights*, edited by Mark McBride. London: Hart Publishing.
- McMahan, Jeff. 1993. 'Killing, Letting Die, and Withdrawing Aid'. *Ethics* 103 (2): 250–79.
- . 1994. 'Revising the Doctrine of Double Effect'. *Journal of Applied Philosophy* 11 (2): 201–12.
- . 2002. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.
- . 2005. 'The Basis of Moral Liability to Defensive Killing'. *Philosophical Issues* 15 (1): 386–405.
- . 2018. 'Doing Good and Doing the Best'. In *The Ethics of Giving*, edited by Paul Woodruff. Oxford: Oxford University Press.
- Mill, John Stuart. 1969. 'Utilitarianism'. In *The Collected Works of John Stuart Mill*, X:203–60. Toronto: University of Toronto Press.

- Muñoz, Daniel. forthcoming. 'Three Paradoxes of Supererogation'. *Noûs*.
- Nelkin, Dana Kay, and Samuel C. Rickless. 2014. 'Three Cheers for Double Effect'. *Philosophy and Phenomenological Research* 89 (1): 125–58.
- Norcross, Alastair. 2005. 'Harming In Context'. *Philosophical Studies* 123 (1–2): 149–73.
- Nozick, Robert. 1974. *Anarchy, State and Utopia*. Oxford: Basil Blackwell.
- Oberdiek, John. 2004. 'Lost in Moral Space: On the Infringing/Violating Distinction and Its Place in the Theory of Rights'. *Law and Philosophy* 23 (4): 325–46.
- . 2008. 'Specifying Rights Out of Necessity'. *Oxford Journal of Legal Studies* 28 (1): 127–46.
- . 2009. 'Towards a Right Against Risking'. *Law and Philosophy* 28 (4): 367–392.
- . 2012. 'The Moral Significance of Risking'. *Legal Theory* 18 (3): 339–56.
- . 2017. *Imposing Risk: A Normative Framework*. Oxford: Oxford University Press.
- Olsaretti, Serena. 1998. 'Freedom, Force and Choice: Against the Rights-Based Definition of Voluntariness'. *Journal of Political Philosophy* 6 (1): 53–78.
- Owens, David. 2012. *Shaping the Normative Landscape*. Oxford: Oxford University Press.
- Parfit, Derek. 1982. 'Future Generations: Further Problems'. *Philosophy & Public Affairs* 11 (2): 113–172.
- . 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- . 2003. 'Justifiability to Each Person'. *Ratio* 16 (4): 368–390.
- . 2011a. *On What Matters: Volume One*. Oxford: Oxford University Press.
- . 2011b. *On What Matters. Volume Two*. Oxford Scholarship Online. Oxford: Oxford University Press.
- . 2017. *On What Matters: Volume Three*. Oxford: Oxford University Press.
- Paul, L. A., and Ned Hall. 2013. *Causation: A User's Guide*. Oxford, New York: Oxford University Press.
- Perry, Stephen. 1995. 'Risk, Harm and Responsibility'. In *Philosophical Foundations of Tort Law*. Oxford: Clarendon Press.
- . 2007. 'Risk, Harm, Interests, and Rights'. In *Risk: Philosophical Perspectives*. Chippenham: Routledge.
- Perry, Stephen R. 2003. 'Harm, History, and Counterfactuals'. *San Diego L. Review* 40: 1283–1313.
- . 2014. 'Torts, Rights, and Risk'. In *Philosophical Foundations of the Law of Torts*, edited by John Oberdiek. Oxford: Oxford University Press.
- Pettit, Philip. 2012. *On the People's Terms: A Republican Theory and Model of Democracy*. Cambridge: Cambridge University Press.
- . 2015. *The Robust Demands of the Good: Ethics with Attachment, Virtue, and Respect*. Oxford: Oxford University Press.
- Placani, Adriana. 2017. 'When the Risk of Harm Harms'. *Law and Philosophy* 36 (1): 77–100.
- Pritchard, Duncan. 2005. *Epistemic Luck*. Oxford: Clarendon Press.
- Pummer, Theron. 2016. 'Whether and Where to Give'. *Philosophy & Public Affairs* 44 (1): 77–95.
- . 2019. 'All or Nothing, but If Not All, Next Best or Nothing'. *Journal of Philosophy* 116 (5): 278–291.
- Quinn, Warren S. 1989a. 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing'. *The Philosophical Review* 98 (3): 287–312.
- . 1989b. 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing'. *Philosophical Review* 98 (3): 287–312.

- Quong, Jonathan. 2009. 'Killing in Self-Defense'. *Ethics* 119 (3): 507–37.
- . 2015. 'Rights Against Harm'. *Aristotelian Society Supplementary Volume* 89 (1): 249–266.
- . 2016. 'Agent-Relative Prerogatives to Do Harm'. *Criminal Law and Philosophy* 10 (4): 815–29.
- Rabenberg, Michael. 2014. 'Harm'. *Journal of Ethics and Social Philosophy* 8 (3): 1–33.
- Rainbolt, George W. 2006. *The Concept of Rights*. Dordrecht, Nether.: Springer.
- Rawls, John. 1999. *A Theory of Justice*. Rev. ed. Oxford: Oxford University Press.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.
- . 1994. 'Rights and Individual Well-Being'. In *Ethics in the Public Domain: Essays in the Morality of Law and Politics*, 44–59. Oxford: Oxford University Press;
- Rice, Christopher M. 2013. 'Defending the Objective List Theory of Well-Being'. *Ratio* 26 (2): 196–211.
- Ripstein, Arthur. 2006. 'Beyond the Harm Principle'. *Philosophy & Public Affairs*, no. 3: 215.
- . 2009. *Force and Freedom: Kant's Legal and Political Philosophy*. Harvard University Press.
- Scanlon, T. M. 1984. 'Rights, Goals and Fairness'. In *Theories of Rights*, edited by Jeremy Waldron, 137–52. Oxford: Oxford University Press.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, Mass: Harvard University Press.
- . 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass. ; London: Belknap.
- Scheffler, Samuel. 1994. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Oxford University Press.
- Shafer-Landau, Russ. 1995. 'Specifying Absolute Rights'. *Arizona Law Review* 37: 209.
- Shiffrin, Seana Valentine. 1999. 'Wrongful Life, Procreative Responsibility, and the Significance of Harm'. *Legal Theory* 5 (2): 117–48.
- . 2012. 'Harm and Its Moral Significance'. *Legal Theory* 18 (3): 357–98.
- Shrader-Frechette, Kristin. 2002. *Environmental Justice: Creating Equality, Reclaiming Democracy*. Oxford, New York: Oxford University Press.
- Sidgwick, Henry. 1981. *The Methods of Ethics*. 7th ed. Indianapolis, IN: Hackett.
- Simmonds, N. E. 2000. 'Rights at the Cutting Edge'. In *A Debate Over Rights: Philosophical Enquiries*, edited by Matthew H. Kramer, N. E. Simmonds, and Hillel Steiner. Oxford: University Press.
- Sinclair, Thomas. 2018. 'Are We Conditionally Obligated to Be Effective Altruists?' *Philosophy & Public Affairs* 46 (1): 36–59.
- Skorupski, John. 2010. *The Domain of Reasons*. Oxford: Oxford University Press.
- Smith, Holly M. 2011. 'The "Prospective View" of Obligation'. *Journal of Ethics and Social Philosophy* 5 (1): 1–8.
- . 2018. *Making Morality Work*. Oxford, New York: Oxford University Press.
- Smith, Martin. 2016. *Between Probability and Certainty: What Justifies Belief*. Oxford: Oxford University Press.
- Snedegar, Justin. 2016. 'Reasons, Oughts, and Requirements'. In *Oxford Studies in Metaethics: 11*, edited by Russ Shafer-Landau, 155–81. Oxford: Oxford University Press.
- Sosa, Ernest. 1999. 'How to Defeat Opposition to Moore'. *Noûs* 33: 141–53.

- Southwood, Nicholas. 2015. 'Democracy as a Modally Demanding Value'. *Noûs* 49 (3): 504–21.
- Sreenivasan, Gopal. 2005. 'A Hybrid Theory of Claim-Rights'. *Oxford Journal of Legal Studies* 25 (2): 257–74.
- . 2010. 'Duties and Their Direction'. *Ethics* 120 (3): 465–94.
- Steiner, Hillel. 1994. *An Essay on Rights*. Oxford: Blackwell.
- . 2000. 'Working Rights'. In *A Debate Over Rights: Philosophical Enquiries*, edited by Matthew H. Kramer, N. E. Simmonds, and Hillel Steiner. Oxford: Oxford University Press.
- . 2007. 'Moral Rights'. Edited by David Copp. *The Oxford Handbook of Ethical Theory*. <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195325911.001.0001/oxfordhb-9780195325911-e-17>.
- . 2013. 'Directed Duties and Inalienable Rights'. *Ethics* 123 (2): 230–44.
- Stewart, Hamish. 2010. 'The Limits of the Harm Principle'. *Criminal Law and Philosophy* 4 (1): 17–35. <https://doi.org/10.1007/s11572-009-9082-9>.
- Sumner, L. W. 1987. *The Moral Foundation of Rights*. Oxford: Clarendon Press.
- . 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon.
- Tadros, Victor. 2011a. 'Harm, Sovereignty, and Prohibition'. *Legal Theory* 17 (1): 35–65.
- . 2011b. *The Ends of Harm*. Oxford: Oxford University Press.
- . 2015. 'Wrongful Intentions without Closeness'. *Philosophy & Public Affairs* 43 (1): 52–74.
- . 2016a. *Wrongs and Crimes*. Oxford: Oxford University Press.
- . 2016b. 'Causation, Culpability, and Liability'. In *The Ethics of Self-Defence*, edited by Christain Coons and Michael Weber. Oxford University Press.
- Tasioulas, John. 2010. 'Taking Rights out of Human Rights'. *Ethics* 120 (4): 647–78.
- . 2015. 'On the Foundations of Human Rights'. In *Philosophical Foundations of Human Rights*, edited by Rowan Cruft, Massimo Renzo, and S. Matthew Liao, 45–70. Oxford: Oxford University Press.
- Thompson, Michael. 2004. 'What Is It to Wrong Someone? A Puzzle about Justice'. In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. J. Wallace, Phillip Pettit, Samuel Scheffler, and Michael Smith, 333–84. Oxford: Oxford University Press.
- Thomson, Judith Jarvis. 1986a. 'Self-Defence and Rights'. In *Rights, Restitution, and Risk: Essays in Moral Theory*, edited by William Parent, 33–48. Cambridge, Mass: Harvard University Press.
- . 1986b. 'Some Ruminations on Rights'. In *Rights, Restitution, and Risk: Essays in Moral Theory*, edited by William Parent, 49–65. Cambridge, Mass: Harvard University Press.
- . 1990. *The Realm of Rights*. Cambridge, Mass: Harvard University Press.
- . 1991. 'Self-Defense'. *Philosophy & Public Affairs* 20 (4): 283–310.
- . 2011. 'More On The Metaphysics of Harm'. *Philosophy and Phenomenological Research* 82 (2): 436–58.
- Timmerman, Travis. 2016. 'Reconsidering Categorical Desire Views'. In *Immortality and the Philosophy of Death*, edited by Michael Cholbi. London: Rowman & Littlefield.
- Van Duffel, Siegfried. 2012. 'The Nature of Rights Debate Rests on a Mistake'. *Pacific Philosophical Quarterly* 93 (1): 104–123.
- Velleman, J. D. 2008. 'Persons in Prospect'. *Philosophy & Public Affairs* 36 (3): 221–44.
- Waldron, Jeremy. 1985. *Theories of Rights*. Oxford: Oxford University Press.

- . 1988. *The Right to Private Property*. Oxford: Clarendon Press.
- Wallace, R. Jay. 2019. *The Moral Nexus*. Princeton, NJ: Princeton University Press.
- Wellman, Carl. 1985. *A Theory of Rights: Persons under Laws, Institutions, and Morals*. Totowa, NJ: Rowman & Allanheld.
- . 1999. *The Proliferation of Rights: Moral Progress or Empty Rhetoric?* Westview Press.
- Wellman, Christopher. 1995. 'On Conflicts between Rights'. *Law and Philosophy* 14 (3): 271–95.
- Wenar, Leif. 2005. 'The Nature of Rights'. *Philosophy & Public Affairs* 33 (3): 223–52.
- . 2013. 'The Nature of Claim-Rights'. *Ethics* 123 (2): 202–29.
- . 2015. 'Rights'. In *The Stanford Encyclopaedia of Philosophy*, edited by Edward N. Zalta, Fall 2015. Stanford, CA: Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2015/entries/rights/>.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.
- . 2009. 'Replies to Critics'. In *Williamson on Knowledge*, edited by Patrick Greenough and Duncan Pritchard. Oxford: Oxford University Press.
- Woollard, Fiona. 2012. 'Have We Solved the Non-Identity Problem?' *Ethical Theory and Moral Practice* 15 (5): 677–90.
- . 2015. *Doing and Allowing Harm*. Oxford: Oxford University Press.
- Zimmerman, Michael J. 1997. 'A Plea for Accusers'. *American Philosophical Quarterly* 34 (2): 229–43.
- . 2008. *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- . 2014. *Ignorance and Moral Obligation*. Oxford: Oxford University Press.