

THE AUTHORITY OF DEONTIC CONSTRAINTS

by

Andrew P. Ross

A thesis submitted to the Department of Philosophy
In conformity with the requirements for
the degree of Doctor of Philosophy.

Queen's University
Kingston, Ontario, Canada
August, 2013

Copyright ©Andrew P. Ross, 2013

Abstract

Non-consequentialists agree that Luke may not kill Lorelai in order to prevent Kirk from killing Richard and Emily. According to this view, Luke faces a deontic constraint: he is forbidden from killing Lorelai, even though doing so will bring about fewer killings overall. The justification of constraints, in my view, faces two challenges. First, constraints must meet the Irrationality Challenge: it needs to be demonstrated that there is nothing inconsistent about the claim that Luke should allow more killings to come about. And, secondly, a successful explanation of constraints must meet the Authority Challenge: we need to know why Luke's reason not to kill Lorelai is normatively categorical.

This dissertation takes up different aspects of Authority Challenge. The first introductory chapter aims to motivate the question of authority as a pressing challenge to non-consequentialism. I argue that the violation of constraints is not just motivated by the thought that they are rationally inconsistent, but by the claim that their intuitive importance cannot be explained.

Chapters two and three take up the connection between the authority of constraints and their interpersonal character. In chapter two, I argue that Stephen Darwall's account of the second-person standpoint cannot yield an account of constraints that satisfies the Authority Challenge and that T.M. Scanlon's contractualism offers us a better way of accounting for the interpersonal significance of constraints. Chapter three argues that Frances Kamm's inviolability approach cannot be reconciled with the intuitive distinction between acting wrongly and wronging someone. The arguments of this chapter are meant to demonstrate that in order for wronging to carry any normative significance, it must play a foundational role in our account of permissibility.

The fourth chapter argues that Moderate deontologists—those who posit a threshold on the killing of the innocent—cannot make sense of the intuitive authority of deontic constraints. The failure of Moderate deontology, I argue, reveals the overlooked appeal of Absolutism. The fifth chapter argues that the authority of restrictions extends to a prohibition on killing non-responsible threats. I argue that a prohibition on killing non-responsible threats accords with the demands of fairness.

Acknowledgements

I am very grateful to have worked under the supervision of Rahul Kumar. Always generous with his time and ready with incisive comments, he made this thesis possible. Kerah Gordon-Solmon had many conversations with me about falling persons, exploding trolleys, and other moral puzzles. With her uncanny ability to construct and dissect thought experiments, Kerah helped me to iron out my arguments. I would also like to thank my examining committee, Will Kymlicka, David Sussman, and Christopher Essert, for a dynamic discussion. It was an honor to discuss my work with them.

At Queen's, I had the pleasure of working with many wonderful philosophers. I would particularly like to thank David Bakhurst, Stephen Leighton, Christine Sypnowich, Deborah Knight, Nolan Ritcey, Eamon Quinn, Jordan MacKenzie, and Octavian Busuioc. I would also like to thank Marilyn Lavoie and Judy Vanhooser for their assistance.

Finally, I would like to thank my family for their encouragement. To Stephanie, I owe more than I can say.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Chapter 1 Introduction.....	1
Chapter 2 Deontic Constraints and the Second-Person Standpoint.....	14
Chapter 3 Wronging, Wrongness, and Inviolability.....	54
Chapter 4 The Authority of Moderate Deontology.....	76
Chapter 5 Threats, Fairness, and the Moral Relevance of Causal Placement.....	105
Bibliography.....	125

Chapter 1

Introduction

Consider the following case:

Constraint: Kirk intends to kill Richard and Emily. Luke can prevent Kirk from doing so by killing Lorelai. Neither Lorelai, Richard nor Emily has done anything that would make them responsible for the situation at hand.

Non-consequentialist theories of moral reasoning are united by the claim that although killing Lorelai would bring about fewer killings overall, Luke is *forbidden* from doing so. Ordinary moral reasoning includes deontic constraints: restrictions that prohibit us from minimizing a certain type of evil by performing an act of that kind.

Consequentialists have challenged deontologists by arguing that constraints are *irrational*: if constraints are based upon the idea that the killing of the innocent is bad, then it is difficult to explain why one should not pursue a course of action that leads to the minimization of such killings (Nozick 1974, Nagel 1986, Scheffler 1982). What is strange about restrictions is that they seem to be reasons for us not to perform killings, but not reasons for us to prevent killings. As Robert Nozick express the question, “How can a concern for the non-violation of C lead to refusal to violate C even when this would prevent other more extensive violations of C?” (30). This is the so-called paradox of deontology. For ease of exposition, I will refer to this as the Irrationality Challenge.¹

¹ There are different ways to formulate the paradox of deontology. For our purposes, skepticism about restrictions seems to be motivated by the thought that if killing is bad, something that minimizes killings

A related, though slightly different, challenge to deontic constraints asks after their intuitive *authority* rather than their rationality. Intuitively, deontic constraints seem to furnish particularly stringent All-Things-Considered-Oughts insofar as they seem to be decisive in nature. To put the point another way, constraints do not provide ordinary reasons, but reasons that seem to be categorical: Luke is *prohibited* from killing Lorelai. In this sense, Luke does not simply have an optional prerogative to refrain from killing Lorelai; rather, he must not do so. The question of authority asks whether or not the alleged normative stringency of constraints can be accounted for. Call this the Authority Challenge.

Non-consequentialists have, in my view, left the Authority Challenge relatively unexplored. That is, non-consequentialists usually attempt to explain the rationality or consistency of constraints without undertaking the question of authority. For the purposes of this introduction, I aim to motivate the question of authority and, following this, to preview the following chapters. Each chapter is meant to be self-standing, but they are all connected insofar as they take up different aspects of the Authority Challenge.

G.E.M. Anscombe first issued a version of the Authority Challenge over fifty years ago in her seminal essay, “Modern Moral Philosophy” (1981). Anscombe argues that moral philosophers rely upon theological understandings of normative authority, despite the fact that they have abandoned the theological assumptions that render such concepts coherent. Anscombe argues that the intelligibility of a moral must—the inescapability of moral demands—depends upon a theological framework that we have abandoned. Moral concepts such as “morally ought” or “morally obligated” are supposed to have the normative force of a moral law (4). The problem however, is that, treating these concepts in a legalistic fashion becomes unintelligible if we no

overall should not be forbidden. Kasper Lippert-Rasmussen (1999) discusses different ways of formulating the paradox.

longer believe in a divine lawmaker. Although the authority of deontic constraints holds intuitive appeal, Anscombe's argument underscores the pressing need to give support to moral instinct.

Anscombe's challenge is meant to apply to morality as a whole, but the question of authority seems to be particularly pertinent to the issue of deontic restrictions. Specifically, non-consequentialists do not typically claim that Luke simply has a reason to refrain from killing Lorelai; rather, they claim that he is forbidden from doing so. Treating constraints as prohibitions implies that the considerations that count in favor of killing Lorelai lack the normative force that they typically have. If Luke is prohibited from killing Lorelai, then he ought not to be "tempted by fear or hope of consequences" (9). Notably, this makes sense within a Hebrew-Christian framework: the consequences have no force because we are ordered by God to do the best that we can within the rules he has given us (8). Within a divine command framework, the idea of a prohibition gains its normative force from the existence of divine lawmaker: we are bound to obey the restriction because we understand ourselves as being obligated to God. But, of course, most non-consequentialists do not think of themselves in this way. As such, the question arises as to whether or not we are entitled to think of constraints as normatively binding prohibitions. The Authority Challenge highlights a different form of irrationality than the one identified by consequentialists: how can secular moral philosophers explain the distinctive normative stringency of deontic constraints?

If we cannot explain the intuitive authority of deontic constraints, it is not clear why we should continue to believe in their existence. To demonstrate the gravity of the worry, consider the following example borrowed from Alasdair MacIntyre. Captain Cook, in the journal of his third voyage, records the discovery by English seamen of the Polynesian word *taboo* (111). The English seamen found that although the Polynesians had liberal sexual practices, there remained a

stringent prohibition on men and women eating together. When the seamen inquired as to why mixed dining was forbidden, they were told that such a practice was taboo. Now, taboo did not simply mean that the practice was disallowed; rather, the use of the word taboo was supposed to imply that it was prohibited for a *particularly significant* reason. Yet, although the Polynesians used the word taboo in this way, the seamen were unable to draw an intelligible justification from the Polynesians as to why the practice was forbidden (111).

MacIntyre argues that the natives themselves did not really understand the word that they were using (111). Instead, their failure to articulate the normative authority of the taboo indicates that the prohibition was simply a remnant from a previous ethical framework that made the proscription intelligible. The Polynesians, in other words, had lost the social background that made intelligible the taboo status of men and women eating together; their continued use of the word was simply a hangover from a collapsed social framework. Moreover, the inability of the Polynesians to make the authority of the taboo intelligible indicated that the concept itself was endangered; in fact, only forty years later the taboo disappeared without major consequence (111). The extinction was no doubt eased by the fact that once the taboo was extracted from its original context it appeared to be *entirely arbitrary*. In MacIntyre's words, once moral concepts "have been deprived of any status that can secure their authority and, if they do not acquire some new status quickly, both their interpretation and their justification become debatable" (112). The question at hand is whether or not non-consequentialists are treating constraints as *taboo*. If the normative stringency that we attach to constraints depends upon a cultural framework that we no longer have access to, then the worry is that we are under the spell of an illusory mesmeric force that will, overtime, disappear.

To a certain extent, the Irrationality Challenge and the Authority Challenge bleed together: we should expect our rationale for constraints to tell us something about their intuitive importance. That said, many responses to the Irrationality Challenge leave the question of authority untouched. Philippa Foot, for example, argues that we can dissolve the paradox of deontology by seeing a blank where the consequentialist sees a better state of affairs (1985, 204). That is, on Foot's line of argument, we should simply deny that the ranking of better and worse states of affairs has any meaning when we are considering questions of justice.²

In one respect, responses such as Foot's are quite insightful: the teleological conception of rationality, advanced by consequentialists, does not seem to apply to all values. The paradox of deontology seems to assume that respect for constraints must take the form of a goal; i.e. it assumes that we care about constraints because we have the goal of making sure that there are fewer murders in the world. Yet, as T.M. Scanlon argues, "not all the considerations that figure in determining the eligibility of an action have to take the form of "goals"... (1998, 85). To claim that one values friendship, for example, does not seem to be equivalent to the claim that one is committed to having as many friends as one can or to bringing about more friendships in the world (88-89). With respect to values such as friendship, it is not obvious that we must adopt a maximizing conception of rationality. And, this opens up room for thwarting the charge of irrationality: if constraints provide non-teleological reasons, then there is nothing irrational about not killing Lorelai.

Yet, the problem with strategies such as Foot's and Scanlon's is that they seem incomplete in a significant way. If we simply see a blank where the consequentialist sees a better state of affairs, then it is not at all obvious why we should think of constraint type cases as

² John Taurek (1997) argues, similarly, that there is no sense in which one state of affairs is worse than another.

haunting in any way whatsoever. As Samuel Scheffler argues, if Foot is correct, then it is not clear that we should *even understand* the question posed by the consequentialist (1985, 414). If we saw a blank, where the consequentialist sees a better state of affairs, then there would not be anything troubling about constraint type cases. The problem, obviously, is that the consequentialist's question *does* make sense: more deaths simply seems worse than fewer deaths. Although we might deny that the reasons provided by constraints take a teleological form, it does not seem open to us to insist that we cannot see the *prima facie* appeal of violating constraints.

The violation of constraints need not be motivated by the claim that constraints are irrational, but by the claim that they do not provide reasons that are strong enough to prevent us from acting on considerations that we recognize as holding normative force. Derek Parfit, for example, does not argue that constraints are irrational or unimportant, but that there is no explanation as to why they are *so* important that they defeat other normatively compelling considerations (2011, 365). If Luke kills Lorelai in order to save Richard and Emily, then we should on Parfit's view, regard this "in a sober way" as good news (393). Parfit's challenge to constraints is not to their consistency or rationality, but to their intuitive authority.

Dissolving the irrational appearance of constraints, then, is only half of the battle. Constraints are perplexing, not just because they seem to be inconsistent, but because they require us to act against particularly weighty considerations. If there is a straightforward sense in which we can prefer fewer deaths to more deaths, then we need to explain, not only why constraints are rational, but also why they are so stringent that they defeat other normatively compelling reasons.

Deontic Constraints and the Second-Person Standpoint

A familiar strategy for attempting to vindicate the normative stringency of moral obligations is to connect them to formal features of practical reason or agency. Kant's moral theory is the most obvious example of this type of explanation. On a Kantian view, anyone who regards themselves as a rational agent is committed to recognizing the authority of the Categorical Imperative. Accounts such as Kant's can be considered formal accounts in the sense that they attempt to connect constraints to considerations that are independent of the desirability of any other end that an agent might have. The upshot of a formal strategy would be to reveal constraints—from the perspective of the deliberating agent—as normatively inescapable. Formal strategies aim to meet the Authority Challenge by revealing constraints to be something that we cannot conceptually refuse.

Stephen Darwall, in *The Second-Person Standpoint*, attempts to give a formal account of constraints by connecting them to the concept of second-personal address. As Darwall presents it, to address an agent from the second-person standpoint is to make a claim on her will in such a way that she is normatively bound to comply with your command. When Lorelai, for example, orders Luke not to kill her, her command is distinctively second-personal insofar as it presupposes her authority to hold Luke accountable. In coming to recognize the constraint as second-personal, Luke must recognize himself as accountable to Lorelai.

The second-person standpoint appears to offer a promising way of meeting the Authority Challenge. Darwall's project aims to answer Anscombe's challenge by replacing a divine lawmaker with secular legislators. What makes constraints authoritative, on Darwall's view, is not the fact that God mandates them, but that you and I mandate them.

The aim of this chapter is two-fold. First, I argue that the second-person standpoint cannot yield a plausible account of restrictions. Although the second-person standpoint is widely

regarded as a non-consequentialist account of moral reasoning, I argue that it fails to answer the Irrationality Challenge. Specifically, I argue that constraints are threatened, not just by appeals to states of affairs, but also by competing agent-relative reasons. Secondly, I present T.M. Scanlon's contractualism as an interpersonal alternative to Darwall's second-person standpoint. On the contractualist account of moral reasoning, I argue, we can understand constraints as something that no one, suitably motivated by mutual recognition, could reasonably reject. Restrictions, on the contractualist account that I present, can be understood as considerations that enable us to relate to one another as independent self-governors. Understood in this way, constraints are not inescapable—they are something that we can conceptually refuse—but they are duties that carry a distinct sense of relational importance.

Formal accounts of deontic constraints, such as Darwall's, are appealing, in large part because they attempt to vindicate deontic constraints in such a way that the idea of their violation can be nothing more than a confused, amoral fantasy. An additional aim of this chapter is to demonstrate that the appeal of formal accounts is deceptive. The problem, I argue, with formal strategies such as Darwall's is that they fail to give us a compelling explanation of why constraints actually matter to us. In contrast, a substantive account of constraints, such as the one that I propose, can better uncover the special significance that constraints hold for us.

Wronging, Wrongness, and Inviolability

In my third chapter, I examine Frances Kamm's inviolability approach to deontic constraints. Kamm attempts to dissolve the paradox of deontology by connecting restrictions to the value of inviolability. If we are forbidden from killing one person to save many, then Richard, Emily, and Lorelai each possesses an especially high level of inviolability. In contrast, if it were permissible to kill Lorelai, then although we might have a better chance at survival, our level of inviolability,

as individual persons, would be lower.

The primary appeal of the inviolability approach is that it offers a response to the Irrationality Challenge. If inviolability is not a value that we maximize—i.e. it is not a value that we bring about through our actions—then there is nothing paradoxical about the thought that Luke is forbidden from killing Lorelai. The concern not to kill Lorelai, in other words, does not stem from a consideration to minimize rights violations, but to respect the moral status of each individual person. Luke cannot be moved to kill Lorelai out of a concern to protect the inviolable status of Emily and Richard because doing so would be *self-defeating*: if it were permissible to kill Lorelai then everyone would have a lower level of inviolability.

In Kamm's most recent presentation of the inviolability approach, she expands upon the theme of inviolability in order to account for the distinction between acting wrongly and wronging someone in particular. If Luke, for example, burns down a tree, we might claim that he has acted wrongly. In contrast, if Luke kills Lorelai then we will think that he has not simply acted wrongly, but that he has *wronged* Lorelai in particular. The distinction between wronging and wrongness—what I call the Wronging Distinction—is, arguably, one way of explaining the intuitive stringency of constraints.³ By incorporating the Wronging Distinction into her account of constraints, Kamm attempts to add a dimension of interpersonal authority to the inviolability approach.

The aim of this chapter is to explore whether or not Kamm can, in fact, account for the Wronging Distinction. I argue that she cannot and that her attempt to integrate the Wronging Distinction threatens to render the concept of inviolability explanatorily redundant. In addition to

³ Scanlon (1998), Hurley (2009), Darwall (2006), and Wallace (forthcoming).

this claim, I argue that my diagnosis of Kamm’s approach has implications for the way that we think about significance of the Wronging Distinction. In particular, I argue that if we wish to incorporate the Wronging Distinction as a significant moral feature, our account of moral reasoning must be foundationally interpersonal. Accounts of wrongdoing cannot be meaningfully extracted from non-interpersonal values or standards of permissibility. Unless wrongdoing plays a foundational role in accounting for the permissibility of an action, it is normatively speaking, an idle fifth wheel.

The Authority of Moderate Deontology

Typically, non-consequentialism is understood as being aligned with common sense morality.⁴ It seems to be part of common sense, for example, that Luke ought not to kill Lorelai even if doing so will bring about a greater good. However, although common sense morality includes deontic prohibitions, it also posits that restrictions admit of a threshold: there are certain scenarios where the consequences of not killing an innocent person are so horrible, that we are required to do so. In my fourth chapter, I argue that taking the Authority Challenge seriously pulls non-consequentialism out of alignment with common sense: if we hope to account for the authority of restrictions, then we ought to forego the positing of a threshold. Meeting the Authority Challenge requires us to be Absolutists: constraint against killing the innocent must never be violated *no matter what the consequences*.

Most deontologists describe themselves as Moderate deontologists. Although it is wrong, they claim, for Luke to kill Lorelai in order to save Richard and Emily, it would *not* be wrong to kill Lorelai in order to save a great many people. Additionally, Moderate Deontologists claim that if we accept the existence of a threshold, this does not alter the authoritative character of sub-

⁴ Shelly Kagan, for example, claims, “common sense morality is deontological rather than consequentialist” (1998, 73).

threshold constraints. Generally, Moderate deontologists make the following authority-related claims. First, sub-threshold constraints remain particularly stringent, All-Things-Considered Oughts. And, secondly, Moderates claim that catastrophe cases should be understood as exceptional moral tragedies.

I argue that the Moderate does not have the grounds for making either of these claims. Specifically, I argue that in order to be intelligible, Moderate deontology must be committed to the claim that constraints ought to be weighed against the value of aggregate outcomes. Accepting this latter claim, I argue, comes at a great cost. If constraints are simply considerations to be weighed, then there can be situations where we may permissibly alter the constraint's normative force. Additionally, understanding constraints as weighty considerations casts considerable doubt on the claim that catastrophe cases are, by the Moderate's own standards, exceptional moral tragedies. In fact, as I argue, Moderates have reason to see threshold situations, not as exceptional moral tragedies, but as fortunate situations.

Threats, Fairness, and Burden Distribution

The final chapter of my project argues that the authority of restrictions extends to a prohibition on killing non-responsible threats. Non-responsible threats are agents who pose—through no fault of their own—a lethal, unjustified threat to another agent. For example, suppose that someone has been pushed off of a cliff and will land on me unless I vaporize them with my ray gun.⁵ This falling person is a non-responsible threat: they threaten me in virtue of their place in the causal architecture, but they are not in anyway responsible for the threat that they pose.

⁵ Robert Nozick is the original author of this example (34).

Many non-consequentialists believe that it is permissible to kill non-responsible threats (Thomson 1991; Tadros 2011; Quong 2009). For these non-consequentialists, constraints admit of an agent-relative exception: when it is *my* life that will be lost, it seems too demanding to forbid me from exercising lethal defensive force. I follow Jeff McMahan (2009) and Michael Otsuka (1994) in arguing that the prevailing view is mistaken. Though, as I will explain, my own analysis of the problem departs sharply from theirs.

In the first half of this chapter, I consider a recent defense of the claim that it is permissible to kill non-responsible threats. Building on the work of Warren Quinn (1991), Jonathan Quong (2009) argues that the difference between killing threats and bystanders corresponds to a distinction between killing someone eliminatively and killing them opportunistically. Killing a threat, or so the argument goes, involves a less pernicious form of agency than killing a bystander. According to Quong, the distinction between opportunistic and eliminative agency reveals that there are two significantly different ways of killing someone as a means to save your own life.

Most non-consequentialists, including McMahan and Otsuka, accept that the distinction between opportunistic and eliminative agency carries some moral significance, but I argue that the distinction's importance is illusory. Here, the question of authority is pertinent in two respects. First, non-consequentialists, in my view, tend to place a great weight on certain moral distinctions without explaining their intuitive normative stringency. We do not, I argue, have any reason to believe that opportunistic killings are, in fact, morally worse than eliminative killings. Secondly, even if we accept that the distinction carries some weight, it is not obviously operational in threat-type cases.

In the second-half of this chapter, I argue that threat-type cases have been miscast as matters of responsibility and that we should, instead, understand such scenarios as posing questions of fair burden distribution. McMahan and Otsuka deny that causal placement carries any moral relevance, but I argue that this is a mistake: both non-responsible threats and threatened parties are required, in virtue of their causal placement, to mitigate the costs of the situation by taking on especially heavy burdens. Generally speaking, when two agents face a threat they are required, if they are able, to split the costs of the situation between them. In line with this view, I argue that a prohibition on killing non-responsible threats can be understood as being in accordance with the demands of fairness.

Chapter 2

Deontic Constraints and the Second-Person Standpoint

Suppose that Ted can save Bob and Carol by killing Alice. Common sense morality dictates that Ted is forbidden from killing Alice despite the fact that doing so will save Bob and Carol. According to common sense, Ted faces a deontic constraint: a restriction that prohibits us from minimizing a certain type of evil by performing an act of that kind. Consequentialists have challenged common sense by arguing that the intuition is *irrational*: if constraints are based upon the idea that burdens are bad, then it is difficult to explain why one should not pursue a course of action that leads to the minimization of burdens (Nozick 30, Nagel 1986 177, Scheffler 80). Call this the Irrationality Challenge.

A related, though slightly different, challenge to deontic constraints asks after their intuitive *authority* rather than their rationality. Deontic constraints seem to furnish particularly stringent All-Things-Considered-Oughts: non-optional reasons that defeat concerns that we would, other things being equal, take very seriously. Deontic constraints seem to be *categorical* in nature: Ted seems to have a decisive reason not to kill Alice and not simply an optional prerogative to refrain from doing so. The question of authority asks whether or not the alleged normative authority of constraints can be accounted for. Call this the Authority Challenge.

In *The Second-Person Standpoint* (2006), Stephen Darwall argues that the normative significance of morality can only be understood by entering into the second-person standpoint. Darwall argues that entering into the second-person standpoint requires agents to think in terms of an authoritative, interpersonal circle; when we address one another, person-to-person, we presuppose a relationship of accountability and authority. These conceptual presuppositions,

according to Darwall, account for the authoritative nature of moral duties. In addition to this claim, Darwall also argues that the second-person standpoint should be thought of as the foundation for non-consequentialist moral reasoning. In particular, Darwall argues that the second-person standpoint ought to be understood as the foundation for T.M. Scanlon's contractualism.

Commentators have found the second-person standpoint appealing, in part, because Darwall presents it as paradigmatically non-consequentialist. Darwall's focus on the authoritative nature of moral duties seems to capture what is distinctive about deontological positions. Additionally, Darwall's emphasis upon the interpersonal nature of second-personal reasoning seems to provide the necessary resources for deontic constraints. Insofar as the second-person standpoint operates person-to-person, it seems to provide a distinct alternative to the impersonal reasoning of consequentialism. Paul Hurley, for instance, argues that

Alice has an independent claim on Ted, in such a case, not to be harmed by him, a claim that is independent of the third person, impersonal badness of that harm resulting. It is a claim that Alice has on any other "you" any "second person", not to be harmed by that person. (186)

Mark Lebar also appeals to the second-person standpoint as fertile ground for restrictions:

Your nature gives me an (agent-relative) reason not to act on you in that way, and you are entitled to hold me accountable to such reasons...Second personal reasons in this way offer just the sorts of resources needed to make sense of deontic constraints (649).

Despite its non-consequentialist appearances, however, it remains to be shown that the second-person standpoint is, in fact, non-consequentialist. In this respect, deontic constraints serve as *the*

benchmark for testing non-consequentialist credentials: if the second-person standpoint operates in a deontological fashion, then it certainly ought to produce restrictions.

I argue that the second-person standpoint fails to yield an account of restrictions that meets both the Authority Challenge and the Irrationality Challenge. The structure of this chapter is as follows. Firstly, I argue that deontic constraints are not just threatened by appeals to states of affairs, but that they are also challenged by the idea that certain individuals may have agent-relative reasons that seem to count against them. In order to be properly non-consequentialist, the second-person standpoint must provide a way of delineating between reasonable and unreasonable demands. Properly meeting the Irrationality Challenge, I argue, requires that Darwall provide a way of adjudicating between competing second-personal reasons. The problem, for Darwall, is that neither the conceptual structure nor the agent-relativity of second-personal reasons can accomplish this latter task. Without a method of arbitration, I argue, there is nothing inherently non-consequentialist about the second-person standpoint.

The above criticisms may seem to suggest that deontic constraints cannot be extracted from an interpersonal account of morality. As such, in the second half of this chapter, I present Scanlonian contractualism as an interpersonal alternative to the second-person standpoint. Darwall argues that the second-person standpoint ought to be viewed as the foundation to Scanlonian contractualism, but I argue that contractualism offers a more plausible theory of interpersonal moral reasoning that evades the criticism outlined above.

Scanlon's contractualism, I argue, serves as a useful foil to Darwall's second-person standpoint. On Darwall's view, considerations of substantive value are reasons of the wrong kind to ground the moral claims that we have on one another. Contractualism, in contrast, takes the substantive appeal of mutual recognition as its foundation. I argue that taking this aspect of the

contractualist picture seriously enables us to develop a compelling account of constraints.

Restrictions, according to the contractualist, are duties that no one could reasonably reject insofar as they structure the very terms of interpersonal recognition. More specifically, I argue that deontic constraints can be thought of, on the contractualist picture, as prerequisites for relating to one another as individual persons.

Finally, I return to the Authority Challenge. On Darwall's view, restrictions are categorical because we are, in entering the second-person standpoint, conceptually committed to recognizing our own accountability. I argue, however, that accountability only captures a small fragment of the intuitive authority attached to restrictions. Although we cannot, on the contractualist picture, think of constraints as inescapable nature, I argue that this is not the misfortune that it initially appears to be. The significance of deontic constraints, for those of us in the moral community, stems from the importance of relating to one another as persons, rather than a formal concern about practical reason. These considerations, I argue, offers us a different way of thinking about the authority of deontic constraints.

Darwall's Second-Person Standpoint

Before continuing, it will be helpful to give an overview of Darwall's second-person standpoint and the account of constraints it is meant to produce. Darwall uses two examples to explain second-personal reasoning. The first example is of a platoon sergeant ordering her troops to fall in (12). What is important, for Darwall, is the conceptual structure of the sergeant's command. In issuing the command, the sergeant does not simply intend that the soldiers will fall in line; rather, the sergeant intends that that the soldiers will fall in line because *she* has issued the command. If the soldiers fall in for some other reason—say, because they want to move out of the sunlight—then the sergeant's command will have failed as a command. What is important about

the sergeant's command is that it carries a built in request for recognition: the soldiers must fall in line and the reason they must do so is because the *sergeant* has made the order.

In order to highlight the conceptual structure of the sergeant's command, compare the military example to a case where I point out that you have reason to eat a healthier diet. In pointing this out, there is no sense in which I presuppose the authority to instruct you about your diet. If my advice is good advice, it is because I stand in a certain relation to facts and evidence and *not* because of the relation that I stand in to you (12). Unlike the sergeant's command, there is no sense in which you should change your diet because *I* have told you to do so. If you should change your diet, then you should do so for the sake of your health or some other agent-neutral reason. My dietary instructions are epistemic, rather than authoritative, in nature: I am trying to get you to see a reason that is there independently of my pointing it out to you. Second-personal reasons, in contrast, have a different conceptual structure. In issuing the command, the sergeant presupposes both her authority to demand this of her troops and their ability to comply. Unlike my nutritional advice, the sergeant's command is not merely epistemic—she is not trying to get her troops to see a reason that is there anyway—rather, the sergeant's command *creates* a reason for the troops to fall in.

Darwall argues that the sergeant gives the troops a second-personal reason—a reason that can only be grasped by presupposing the sergeant's authority and the troop's accountability (12).⁶ From the viewpoint of the soldiers, recognizing the reason as second-personal requires them to assume that a relationship of authority and accountability obtains between them and their superior officer. Again, the significance of this, according to Darwall, is that commands have a different conceptual structure from agent-neutral reasons. Grasping second-personal reasons as

⁶ As will become evident, the military example is of limited use as an analogy for moral reasoning. For present purposes, it is useful insofar as it highlights the conceptual structure of a demand.

second-personal requires us to think in terms of a second-personal circle: in order to recognize the reason as second-personal, one must presuppose the authority of their interlocutor to make the claim and, in so doing, presuppose their own accountability for compliance (13). To put the point another way, authority and accountability can be understood as the “normative felicity conditions” of second-personal address: addresser and addressee must assume that such a relationship obtains between them in order to make sense of the command (74-76). The upshot of this is that the soldiers, in recognizing the sergeant’s command as second personal, have already stepped into a conceptual circle of authority and accountability.

Darwall borrows his second example from David Hume’s *Enquiry Concerning the Principles of Morals* (1985). Suppose that Alice is afflicted with gout and that Ted accidentally steps on her inflamed toes. Alice protests by asking Ted to remove his foot from her gout-ridden appendage. There are two ways to think of Alice’s protest. We might think of Alice’s protest as drawing attention to an agent-neutral reason: Alice is in pain and everyone has reason to reduce the amount of pain in the world. Anyone who is in a position to remove the offending foot should do so. Understood in this way, Alice’s protest is closer to my dietary advice: she is simply pointing to an agent-neutral reason.

In contrast, on the understanding that Darwall favors, Alice’s protest functions in the same way as the sergeant’s command insofar as it has a built in, second-personal, bid for recognition: Ted ought to remove his foot because *Alice* has the authority to demand this of him. From Alice’s perspective, her protest would fail if Ted removed his foot because he got bored or because it would decrease the aggregate amount of suffering in the world. In order for Alice’s protest to succeed *qua* second-personal reason, Ted must remove his foot because *she* has ordered him to do so.

Darwall claims that the conceptual circle of second-personal reasons imbues moral reasons with their distinctive normative pull. Recognizing his reason to remove his foot as second-personal requires Ted to presuppose his own accountability and Alice's authority. Accordingly, in coming to grasp his reason as second-personal, Ted's practical deliberations come to be framed in terms of a "moral must." Alice's protest, much like the sergeant's command, presupposes the normative felicity conditions of second-personal address. Grasping the reason as second-personal requires Ted to presuppose Alice's authority and his own accountability. Thus, at a conceptual level, Ted is normatively bound to act on the reason insofar as he recognizes himself as accountable *to* Alice.

The difference between agent-neutral reasons and second-personal reasons on Darwall's view lies in their relation to accountability. In coming to see something as desirable, I may see it as an eligible option, but I do not see it as normatively inescapable. That is, I do not see myself, in virtue of having such a reason, as accountable to anyone. In contrast, second-personal reasons are supposedly claim-exacting in a way that makes them normatively binding. In grasping the reason as second-personal, Ted recognizes himself as accountable *to* Alice; i.e., he recognizes that Alice may hold him responsible for his failure to comply with the reason. Darwall is not exactly clear about what counts as holding someone responsible. Often, he seems to mean that in holding someone responsible we would be entitled to express reactive attitudes or bestow some sort of penalty or sanction on them (68-69). Alice, for example, could feel resentment towards Ted and the Sergeant could punish her troops in some other substantive sense.

It is important to note that Darwall understands second-personal reasons as "deeply agent-relative" (5-6). Darwall seems to mean two things by this. First, when Alice addresses Ted, she creates a reason that exists solely for him. Unlike agent-neutral reasons—which exist for

everyone—second-personal reasons are agent-relative reasons that are indexed to particular people. In addressing Ted, Alice gives a reason directly to him rather than the world at large.

Secondly, Darwall also understands second-personal reasons as *irreducible* to the relationship that obtains between addresser and addressee. That is, Darwall posits a sharp divide between second-personal reasons and considerations of substantive value. Appeals to something such as say the badness of pain or the goodness of standing in a certain relationship counts, for Darwall, as “reasons of the wrong kind” (66). Since such reasons fail to be conceptually linked to notions of accountability, Darwall argues that they cannot ground our moral duties. Second-personal reasons, in other words, are “second personal all the way down” insofar as they stem *directly* from relations of authority and accountability (59).

Darwall argues that second-personal reasoning provides an answer to both the Irrationality and Authority Challenge. To explain, let us change the gout example into a podiatric restrictions case where a villain will stomp on Bob and Carol’s feet unless Ted stomps on Alice’s foot. Darwall argues that second-personal reasoning avoids the Irrationality Challenge because of the way that second-personal reasons function. Agent-neutral reasons are concerned with states of affairs, and in light of this, an agent with agent-neutral reasons is not prevented from comparing one state of affairs (fewer foot stomplings) with another (more foot stomplings) (7). It is this sort of comparison, Darwall argues, which generates the paradox of deontology: if Ted’s reason not to step on Alice’s foot is to produce a better state of affairs, then once our villain enters the scene, he seems permitted to secure the best overall state of affairs (fewer foot stomplings).

Comparatively, if Ted’s reason is second-personal—if it stems from Alice’s authority—then he does not have a reason to minimize foot stompling because Alice’s demand does not refer to the value of states of affairs but to her own authority (7). If comparing one state of affairs with

another generates the paradox of deontology, then second-personal reasoning seems to avoid the paradox of deontology by avoiding appeals to states of affairs. Put simply, Darwall argues that second-personal reasoning operates person-to-person—*not* person-to-state-of-affairs—and, in light of this, answers the Irrationality Challenge by circumventing the paradox of deontology (7 n.13).

Darwall answer to the Authority Challenge should be obvious. Recognizing constraints as second-personal requires both the addressee and addresser to presuppose that a relationship of authority and command obtains between them. And, the conceptual circle of second-personal reasoning structures the restriction in categorical terms: entering the second-person standpoint requires Ted to recognize Alice’s authority and, thereby, his own accountability. Constraints, as understood from the second-person standpoint, are “claims and demands that we cannot avoid assuming we have the authority to address to each other as one free and rational person among others” (38).

Reasonable and Unreasonable Reasons

For explanatory purposes, it is important to flag a potential problem plaguing Darwall’s account. The military example and the gouty toe example seem to be crucially different in the following respect. Pointedly, the sergeant’s address is supposed to create a new reason that did not previously exist; but, if address plays a generative role, then this would suggest that Ted does not have a second-personal reason to refrain from stepping on Alice’s toes *until* she addresses him (26). Yet, as R. Jay Wallace (2007) and Gary Watson (2007) have pointed out, this seems absurd: Ted has reason not to step on Alice’s foot whether she protests or not. Problematically, Darwall’s emphasis upon second-personal *address* seems to render moral reasons implausibly

hostage to actual speech acts. If Alice's address actually creates a reason that was not there before hand then Darwall seems to be committed to a form of moral voluntarism.

Often, it seems that Darwall uses the idea of address as a metaphor for illustrating the conceptual circle of authority and accountability. Yet, simultaneously, Darwall presents address not as a metaphor but as something crucial to the concept of a second-personal reason: "They [second-personal reasons] simply wouldn't exist but for their role in second-personal address" (8). Such equivocations complicate the task of Darwall exegesis. Nevertheless, for the sake of argument, and on a charitable reading, it would seem that address ought to be understood in hypothetical or possible terms. In his response to Wallace and Watson, Darwall claims that:

...the moral community as I understand it is not any actual community composed of actual human beings. It is like Kant's idea of a "realm of ends," a regulative ideal that we employ to make sense of our ethical thought and practice. (64)

Address, whether it is understood as hypothetical or actual, seems to be important to Darwall because it draws interlocutors into a conceptual circle of authority and accountability. The suggestion seems to be that it is only when we hypothetically take up the roles of addresser and addressee that we are fully in the conceptual circle of the second-person standpoint. On this understanding, when Ted is reasoning about whether or not to take his foot of Alice's he is imagining how a representative from the ideal moral community would address him. Whether or not this response successfully avoids moral voluntarism is a question that I shall leave to the side.⁷

⁷ Darwall's suggestion seems to be that our duties are those that a hypothetical community would ideally agree to. Problematically, however, this would seem to make Ted's duty one that is owed to a regulative ideal rather than to Alice herself. As Wallace points out, once the move is made to a regulative ideal, the concept of second-personal address ceases to be genuinely interpersonal. Moreover, note that this shifts the

For present purposes, I wish to highlight the problem of distinguishing between reasonable and unreasonable reasons. When we are deliberating about what to do, certain considerations strike us as reasonable or unreasonable. When I am, for example, designing a syllabus for a first year philosophy course it might seem reasonable for me to teach Descartes or Kant, but it would seem unreasonable to expect first year students to read and understand Hegel. When thinking about which philosophers to teach, I do not need to consider every possible option—it would never occur to me to teach Hegel in a first year course—instead, my appraisal of the task is framed in a certain way given the task at hand.

How do we know that this consideration is unreasonable? If it seems unreasonable of me to expect first-year students to read *The Phenomenology of Spirit*, then this is because it conflicts with a certain aim that I ought to have in teaching a first philosophy course. More specifically, since I am aiming to design a course for people who have never studied philosophy it makes sense for me to select more accessible works of philosophy.

Reasonable and unreasonable reasons are also present in our deliberations with other people. In the giving and taking of requests and demands, certain considerations strike us as reasonable or unreasonable. For example, consider the giving and taking of reasons that happens between a supervisor and a graduate student. It certainly seems reasonable for a graduate student to expect her supervisor to read and comment on a draft of her thesis. In contrast, it seems

explanatory weight off of the concept of address onto the idea of a hypothetical agreement. If an action is made impermissible by its violation of an abstract agreement, then address seems to reveal reasons rather than generate them. Put another way, if address does not render an action impermissible then it seems to play an epistemic role.

unreasonable for the student to ask her supervisor to read her eighty-fifth draft or to proofread and format her work for her.

Notice that considering the agent-relative reasons for and against the request is unhelpful. From the supervisor's perspective the request is too demanding insofar as it would take up too much of her time and she has, accordingly, a fairly strong agent-relative reason for objecting to the request. But the student also has a very strong agent-relative reason for making the request: having someone to proofread her work will free up more time for her to work on her dissertation. This seems to leave us with a case of gridlock where each agent has an agent-relative reason.

Again, in this case, what makes certain requests reasonable and others unreasonable has something to do with an overarching aim or concern. In this case, what seems to make the request for feedback reasonable is that the aim of a student-supervisor relationship is to foster intellectual growth and to help students improve their work. The unreasonable requests are, in contrast, unreasonable because they seem to be at odds with what we think a healthy or valuable student-supervisor relationship ought to be like. If the goal of a student-supervisor relationship is to produce an independent scholar, then, obviously, relying too heavily on one's supervisor is unreasonable precisely because it conflicts with this aim. Placing limits on the amount and type of help that a student can ask for has a relational impact: both student and supervisor are better able to fulfill the aim of their relationship by minimizing co-dependency. Students and teachers can better fulfill the aims of their relationship if there are limits on how much help students can request from their supervisors.

What is important to note about the above examples is that the question of reasonableness gains traction by having an overarching aim or concern in mind. Given the aim of a first year philosophy course, there are certain parameters to the sort of workload that I can expect students

to handle. Similarly, given the purpose of student-supervisor relationships it is reasonable to assume that there are certain limits on how much help we can ask for. Notably, what counts as reasonable in these cases is not simply a matter of intuition; rather, it is something that we can elucidate by thinking about what our given aim ought to be. What makes a reason reasonable is a substantive—rather than a purely conceptual—issue.

Darwall seems to define reasonable reasons as agent-relative reasons that can be addressed person-to-person. In doing this, as previously explained, he means to exclude states of affairs as relevant moral considerations. Additionally, agent-neutral reasons seem to be unreasonable reasons insofar as they are reasons of the wrong kind. Whether or not this provides a sufficient framework for elucidating the reasonableness of constraints is the question I shall turn to next. For present purposes it will help to note that the question of reasonableness is in tension with Darwall's claim that second-personal reasons are second-personal "all the way down" (59). As previously explained, Darwall views second-personal reasons as irreducible to authority relations. Furthermore, the fact that a certain set of teacher-student relationship standards would seem more desirable would be, on Darwall's view, a "reason of the wrong kind" to ground second-personal claims.

The problem with this line of argument is that it leaves us no way of adjudicating between reasonable and unreasonable reasons. At a conceptual level, reasonable and unreasonable reasons are the same. Consider, for instance, the student's unreasonable request that her supervisor correct her spelling, grammar, and formatting. Pointedly, the request bears the *same conceptual structure* as the reasonable demand for critical feedback. If the student requests that the supervisor proofread and format her paper, then she presupposes the same normative felicity conditions that structure a reasonable request. This indicates that if we are going to

consider the question of reasonableness, we cannot simply appeal to authority and accountability. In asking whether or not the request is reasonable, we are asking a substantive question about how students and teachers ought to relate to one another—the fact that we must assume a relationship of authority and accountability is unhelpful. We can assume that teachers and students are accountable to one another because they are in a relationship of some sort, but if we are thinking about the particular ways that they are accountable to one another then we are thinking in substantive rather than conceptual terms. More specifically, the assumption that the purpose of student-supervisor relationships is to produce independent scholars is *not* a conceptual assumption; it is, rather, a substantive claim. Assumptions of authority and accountability do not, by themselves, give us a grip on the distinction between reason and unreasonable reasons.

It might be objected that this is simply not what Darwall is trying to do; that the domain of the reasonable lies beyond the scope of his project. However, it is far from obvious that this response is open Darwall. Without some account of reasonable reasons there is nothing inherently moral about the second-person standpoint. If I order you to dig a hole in my backyard I assume, at a conceptual level, that I hold the authority to make such a request and that you are accountable to me for doing so. And, in recognizing my request as second-personal, you must also step into the conceptual circle of the second-person standpoint. Of course, whether or not I have actually have the authority to order you to do this will depend upon further features of the situation, but we can note that the bare fact of that the request exhibits this sort of conceptual structure does not mark it as moral. More problematically, if we assume that I can give you a reason to dig a hole in my backyard simply by my mere say-so, then there is no limit to the reasons that can be pressed from within the second person standpoint. Without a grip on reasonable reasons, the second-person standpoint gives us a categorical, but vacant sense of authority.

Thus, it is not open to Darwall—if he hopes to illuminate the nature of moral reasoning—to leave the gap between reasonable reasons and the second-person standpoint unabridged.

Without a way to demarcate the morally relevant, it is not clear that there is anything *moral* about the second-person standpoint, let alone non-consequentialist. As such, for the sake of argument I will assume that Darwall is committed to having some way of accounting for reasonable reasons.

Competing Second-Personal Reasons

In this section, I aim to demonstrate that answering the Irrationality Challenge requires us to have a way of demarcating between reasonable and unreasonable reasons. Second-personal reasons may not make reference to states of affairs but it is not clear that the agent-relativity of second-personal reasons can accomplish this task. Deontic constraints are threatened, not just by appeals to states of affairs, but also by competing agent-relative demands. For example:

Forced Aid: Suppose that the villain is threatening to kill Ted and that Ted can spare himself death by paralyzing Alice.

Deontic constraints are not just threatened by numbers: they are also challenged by the idea that respecting them can come at great cost to the would-be violator. In *Forced Aid*, for example, Ted may be able to avoid death by paralyzing Alice. Yet intuitively, Ted still faces a deontic constraint since it is implausible that he may bestow *that* great a burden on Alice. But the question arises as to how we can reconcile this intuition with Ted's perspective: what prevents Ted from addressing Alice second-personally and demanding that she bear the lesser cost?

Cases such as *Forced Aid* are tricky. In certain scenarios, it is entirely plausible that one may burden others in order to avoid harm to oneself. It seems permissible, for example, to lightly step on someone else's toes in order to avoid being hit by a car. That said, as *Forced Aid*

illustrates, duties of beneficence have an intuitive limit: although it may be permissible for us to bestow light burdens on other people (in order to avoid greater burdens to ourselves) we cannot always bestow a burden simply because it is the least harmful burden. The right not to be paralyzed, the right not to be blinded, and the right not to have a limb amputated are, intuitively, maximally stringent constraints: they are not constraints that can be overridden to save one life let alone many lives.⁸

Reflecting on maximally stringent constraints reveals that what counts as a reasonable cost is not simply the cost that we, personally, would prefer to bear. I might prefer to suffer paralysis rather than death, but this does not mean that I may permissibly paralyze you in order to avoid death. If we simply compare death to blindness, amputation, or paralysis it may be the case that we would, personally, prefer to suffer any one of these burdens rather than the burden of death. Yet, even if this is the case—even if we would personally prefer to be blinded rather than killed—it seems impermissible to impose such burdens on other people.

While it seems counter-intuitive that Ted's reason to avoid death would trump Alice's reason to avoid paralysis, it is not clear that Darwall has the resources to tell us why this is so. If we simply weigh the burden of paralysis against the burden of death, then the second-personal scales will indicate that Ted should be permitted to paralyze Alice and this, from a non-consequentialist perspective, is the wrong answer. Furthermore although Ted's demand may be unreasonable, it bears the same conceptual structure as Alice's legitimate demand insofar as both demands presuppose a relationship of authority and accountability. Additionally, while Alice has an agent-relative reason that Ted not paralyze her, Ted also has an agent-relative reason that he be allowed to do so. Markedly, the agent-relativity of second-personal reasons proves unhelpful

⁸ I borrow the term maximally stringent from Judith Jarvis Thomson (1990, 168).

since each agent brings their own reason to the second-person standpoint.

Darwall argues that the idea of a reasonable cost is tied up with the second-person standpoint. Not stepping on someone's foot, for example, does not place an unreasonable demand upon anyone since it is at no cost to the agent not to step on other people's feet. But, to tweak the example, suppose that stepping on your foot is the only way that I can avoid an excessive burden from being placed on myself. In such a case, Darwall argues, "it seems arguable that we (and you) cannot reasonably demand that I not step on your foot, and so it is neither wrong nor a violation of your right for me to do so" (2010, 155). Reasonable costs, on this line of thought, can be understood as operating on a continuum. If we hold the burden to the victim fixed (a stepped-on foot) and alter the burden to the agent who would act, then we end up with the following scale: at one end, there is no cost to the agent and his action would therefore be wrong; at the other end, the cost to the agent is very great and the action is therefore permissible (155). Darwall's suggestion is simple and intuitive: if Ted can avoid stepping on Alice's foot at no cost to himself, then it is clearly wrong for him to do so, but if Ted can only avoid stepping on Alice's foot at great cost to himself, then Ted can demand that Alice absorb the cost of him treading on her foot. Reasonableness, on Darwall's understanding, is a matter of comparing and balancing burdens.

Darwall's response is inadequate in two respects. First, if reasonableness is a function of weighing and balancing reasons then it is not clear that the second-person standpoint is doing any work. Weighing and balancing is a feature of ordinary practical reason: there is no need to enter the second-person standpoint if permissibility follows from the weighing of ordinary reasons. Second, as *Forced Aid* demonstrates, assessing reasonable costs is not always tantamount to weighing immediate burdens. Ted may suffer the more severe burden, yet intuitively, this does not make it reasonable for him to paralyze Alice.

Aggregation also presents a challenge to the second-person standpoint. Darwall assumes that second-personal reasons do not combine in order to tip the balance of reasons, but it is not obvious why this is so. Alice's address might occlude the ranking states of affairs, but it does not necessarily prevent Ted from appealing to an aggregate of second-personal demands. For example:

Interpersonal Aggregation: Bob and Carol will be burdened unless Ted burdens Alice. Bob and Carol both demand that Ted save them. Ted decides that he must preserve as many second-personal relationships as he possibly can, so he decides to kill Alice.

Interpersonal Aggregation remains a live possibility for Darwall in two respects. First, there is nothing about the concept of a second-personal reason that rules out the move of moral clustering. Ted may have a second-personal relationship with Alice, but he also has one with Bob and Carol. This means that whatever he decides to do, Ted's action must be justifiable to all of them. If their demands conflict with one another, then Ted could satisfy as many second-personal demands as possible.

It may be objected that Darwall obstructs the aggregation of second-personal reasons by defining them as reasons that are addressed person-to-person. In order for interpersonal aggregation to be a live possibility, Bob and Carol would need to address Ted as a collective. But, if second-personal reasons are addressed person-to-person, and not person-to-collective, then interpersonal aggregation cannot get off the ground. Bob and Carol may address Ted- but they do so as individuals and *not* as members of a group.

There are two responses that can be made to this objection. First, it is not obvious why the moral community could not demand that Alice sacrifice herself for the community. We can,

plausibly, be accountable to groups. If I break the rules of my country club, then I do not violate the expectations of any one person but of the club as a whole. And my fellow club members can accordingly hold me responsible. Furthermore, in demanding that Alice sacrifice herself for the community, the request does not have to appeal to an ideal state of affairs. Instead, the moral community could say to Alice: “We all have the agent-relative reason of wanting to minimize our chances of being harmed. As such, we should all agree that in situations such as this, the one person ought to sacrifice themselves”. Notably, this demand does not appeal to the value of aggregate outcomes qua aggregate outcomes, but to an agent-relative reason that all of the agents have.

Secondly, we can also note that the possibility of *Interpersonal Aggregation* does not have to depend upon the idea that Bob and Carol address Ted as a group. Instead, the possibility could arise because there is nothing about the nature of second-personal reasons that prevents Ted from treating them as reasons that can be added together. As long as Ted thinks that he has conflicting obligations, it is open to him to try to treat these reasons in an aggregative fashion; i.e. to try to meet as many obligations as possible. Once the reasons presented by Bob and Carol are on the table, there is nothing that prevents them from tipping the balance of reasons in favor of burdening Alice. Generally speaking, when we cannot perfectly comply with all of the reason that apply to us then we should come as close to complete compliance as possible; this is part of the logic of practical reason generally (Raz 2003, 296). If I cannot keep all of my promises, for example, I should keep as many of them as possible. In a similar fashion, if Ted cannot meet all of his second-personal demands then he ought to meet as many of them as possible.

An obvious objection can be raised against *Interpersonal Aggregation*. Second-personal reasons are claims that we have the authority to demand of one another. But clearly, Bob and

Carol do not have the authority to demand that Ted burden Alice: their demands are unreasonable. Since Bob and Carol do not have the requisite authority, it will be argued that they do not issue genuine second-personal reasons.

In response, it can be noted that Bob and Carol's demand certainly is unreasonable. The present point, however, is that Darwall's understanding of second-personal reasons gives us no obvious way of differentiating between reasonable and unreasonable demands. More pointedly, the conceptual structure of a legitimate demand is no different from the conceptual structure of an illegitimate demand. All claims have a built in bid for recognition: they demand the addressee's compliance and they presuppose the authority of the addresser. When Bob and Carol demand that Ted save them they assume that they have the authority to make this demand and that Ted is accountable for complying with their demand. At a conceptual level their demand is no different from Alice's legitimate demand. Agent-relativity, moreover, does not add any helpful content. Bob and Carol—just like Alice—have agent-relative reasons for demanding that Ted save them. Strikingly, neither the agent-relativity nor the conceptual structure of second-personal reasons can tell us anything about their acceptability.

It is not open to Darwall—if he wishes to secure a plausible account of non-consequentialist reasoning—to simply deny that the presence of multiple demands can make a difference. To illustrate:

Lifeboat: Alice is stranded on a rock in the middle of the ocean. Bob and Carol are similarly stranded. Ted can save either Bob and Carol or Alice, but not both.

In order to make sense of *Lifeboat*, Darwall cannot simply deny that numbers can make a difference. Darwall, presumably, will want to conclude that we ought to save the larger number in

Lifeboat though not in *Interpersonal Aggregation*. However, the conceptual structure of second-personal demands does not tell us why we may aggregate in one case but not the other. The accommodation of standard aggregation cases such as *lifeboat* requires that numbers, at least in certain cases, be able to make a difference.

To be clear, the problem facing Darwall is not that Bob and Carol should not count. Rather, the problem is that the agent-relative reasons that Bob and Carol have should not enter the equation. *Interpersonal Aggregation* arises from the fact that there appears to be a conflict between the agent-relative reasons of Bob, Carol, and Alice. Once we have conflict between competing second-personal reasons, constraints begin to take on a paradoxical air: if Ted cares about complying with second-personal demands, then it is not obvious why he should not conform to as many of them as he possibly can. In order to avoid this it needs to be shown that Ted does not have such an obligation to Bob and Carol; i.e. that the reasons which they address to Ted do not count as genuine second-personal reasons. Darwall, in other words, needs to circumscribe the reasons that can be pressed from within the second-person standpoint in order to avoid aggregative outcomes.

Interpersonal Aggregation and *Forced Aid* illustrate the need to meet two requirements:

1. The constraint needs to be justified, not only to the potential victim, but also to everyone affected.
2. The reasons that can be pressed from within the second-person standpoint need to be circumscribed as reasonable or relevant for the purposes of moral argument.

In order to be a moral viewpoint—let alone non-consequentialist—the second-person standpoint will need to meet the above criteria. Securing the above criteria, however, pushes the ideas of

deeply agent-relativity and second-personal address—whether it is actual or hypothetical—out of the picture entirely.

Consider the first requirement. Numerous agents can be benefited by the violation of deontic restrictions. In order to avoid *Interpersonal Aggregation* and *Forced Aid*, the constraint needs to be justified, not just to the victim, but also to each person involved. We need to know why Ted - in respecting the constraint - also respects the second-personal authority of Bob and Carol and, in the latter case, why Ted cannot demand that Alice suffer the lesser cost. After all, Ted has a second-personal relationship with Bob and Carol and, as such, it needs to be explained why they are not wronged by Ted respecting the constraint.

Developing an interpersonal justification of deontic constraints requires us to reconcile the constraint with multiple and possibly conflicting points of view. This means that when we are deliberating about what is morally relevant in a given situation, we have to think about what people are *generally* owed; what anyone—whether it is Bob, Carol or Alice—can normally be expected to demand from us. What Bob, Carol, and Alice have the authority to demand must be common to each of them. Interpersonal authority—if it is to be justified from each person's point of view—must make use of reasons that are available to everyone. Ted may owe it to Alice in particular not to burden her, but the restriction must be justifiable to *everyone* involved in order to be a genuinely interpersonal justification. Notice that this changes the structure of Ted's deliberations. In thinking about what other people are entitled to demand of us we are not necessarily thinking of actual people, but of nonspecific points of view. As Darwall presents it, the phenomenology of second-personal reasons is directed to the event of address: when the sergeant addresses her troops, the individual soldier focuses upon the actual command because it is the speech act itself, which creates the new reason. Ted's phenomenology, however, must be

different if it is directed towards generic reasons rather than particular commands. If Ted's attention is focused on considerations that could justify his action to everyone—Bob, Carol and Alice—then Ted's phenomenology is not really directed towards Alice, but to all of those involved. Rather than think about what a particular agent will demand of him, justifiability to each person requires Ted to think about what *anyone* could demand of him.

There is, to be sure, something appealing about second-personal address. Address directs us towards the potential victim of a violation and, in this manner, forces us to consider the situation from a perspective that is not our own. Hypothetical address—understood as a way of putting oneself in another person's place—can be morally useful in that it makes vivid the perspective of another person. However, its usefulness is limited when it comes to fulfilling the second desideratum, since the idea of address does not, by itself, give us enough content to come to a conclusion about what can be morally justified and what cannot. Ted can think about what particular people might have reason to want, or what they might be inclined to demand of him, but this does not provide any real traction for separating reasonable from unreasonable demands.

Notice that the first requirement—justifiability to each person—puts pressure on Darwall's claim to deep agent-relativity. If the reason that Ted acts upon is justifiable to each person, then it cannot be a reason that exists solely between Ted and Alice. Instead, if the reason is justifiable to each person, this will be because it is what *anybody* in such circumstance would have the authority to demand of anyone else. Under the current circumstances the reason is indexed to Alice, but it could just as easily be indexed to another agent. Consequently, it is difficult to suppose that constraints are deeply agent-relative in Darwall's sense; rather, the fact that they could be used to justify the action to *anyone* seems to make them interpersonal, yet agent-neutral.

Rather than think of interpersonal reasons as agent-relative, it is more helpful to think of them as *generic*: they are reasons that anybody would accept for the purposes of moral deliberation. Significantly, this gives us a much more plausible account of how Ted deliberates about what to do. That is, Ted knows what to do by thinking about what anyone would have reason to want in a situation such as this. Ted can recognize the constraint as one that would obtain in a situation with different individuals.

Darwall should want to admit this latter point to be important. A plausible account of moral reasoning should be able to make sense of the claim that Ted knows what to do without knowing all of the situation's particularities and without knowing what particular individuals will happen to say to him. It should be obvious to Ted, in both scenarios that he is forbidden from burdening Alice and it should be obvious without his being told.

Darwall's shift from address to a regulative ideal may seem to accomplish this latter task. In deliberating about whether or not to burden Alice, Ted does not need to think about what Alice would say to him. Rather, Ted must think of what a representative of the kingdom of ends would say to him. But note, however, that thinking about what the representative from the ideal community would demand can only be helpful if we already have a way of demarcating the reasonable from the unreasonable. But the shift to a regulative ideal can only obfuscate this task: why would ideal persons want to have restrictions on what may done for the sake of the greater good? Why would it be unreasonable, within an ideal community, to demand that others bear the greater cost in a situation such as *Forced Aid*? Thinking of the second-person standpoint as a regulative ideal is unhelpful because there is no content to this regulative ideal.

To sharpen the point, note that the moral motivation that we attribute to the regulative ideal

alters the outcome of hypothetical address. For instance, if the citizens of Darwall's kingdom are of a Hobbesian motivation, it is not obvious that they will agree to restrictions or maximally stringent rights. Rather, whether or not they will do so depends upon the probabilities of the empirical world. A Hobbesian regulative ideal will have different accounts of what counts as a reasonable demand and they will, as a result, form agreements that are not distinctly deontological.

To summarize, the two criteria outlined above put pressure on the second-person standpoint in the following respects. First, Darwall's focus on address obscures the general nature of deontic constraints. Constraints need to be justified to each person affected by the duty. Ted's deliberations are not directed towards the possibility of address but towards considerations that seem relevant from each person's point of view. Secondly, delineating reasonable from unreasonable demands requires us to appeal to resources outside of the second-person standpoint. Neither agent-relativity nor the bare conceptual structure of second-personal reasons can perform this latter task. Without a grasp on what makes certain demands legitimate, problem cases such as *Forced Aid* and *Interpersonal Aggregation* remain open possibilities.

Contractualism as Interpersonal Alternative

Thus far, I have argued that the concept of second-personal address does not have enough content to give us a grip on the concept of reasonable reasons. Non-consequentialist constraints require more content than the second-person standpoint can furnish. The idea that something is "reasonable" or "unreasonable" is intuitive but a simple appeal to intuitions will not justify constraints to anyone who does not already share the intuition. The deficiency in Darwall's account gives us a starting point: we need to give content to the idea of reasonable reasons so that

certain considerations are excluded from moral reasoning. The idea of the reasonable needs to be framed in a particular way before it can do any work.

In this section, I want to argue that the idea of reasonable reasons can be given shape by thinking of interpersonal relations in substantive terms. When we are deliberating about what to do we know that certain reasons are reasonable or unreasonable because of the sort of relationship that we stand in to other people. Returning to the example of the teacher and the student proves helpful. Intuitively, it is reasonable for the student to ask for feedback on her first draft. And, it is equally intuitive that it is *unreasonable* for the student to ask her supervisor to proofread her fiftieth draft. I have argued that we know this by thinking about the aim of student-teacher relationships in substantive terms. When we are thinking about whether or not certain demands are reasonable between teachers and students we are supposing that teachers and students have some sort of common aim that unites their deliberation.

Contractualism, as I will argue, offers us a way of thinking about ordinary interpersonal relationships in a manner that is similar to the supervisor-student relationship. The contractualist begins with the idea that the value of rationality lies in the capacity for rational self-governance, understood as the ability to act in accordance with the reasons that count against or in favor of a particular action (Scanlon 1998, 105-106). The ability to assess reasons and, in turn, the ability to select amongst the reasons that there are enables us to plan and live our lives in an active sense: we can choose how we want our lives to go and we can, accordingly, live life intentionally. Valuing rational self-governance in other people, according to the contractualist, requires us to treat other rational beings “in ways that they could, by proper exercise of this capacity, recognize as justifiable” (169). Put another way, we value rational self-governance in others when we live according to principles that no one, suitably motivated, could reasonably reject.

But why value rational self-governance in this way? The contractualist answer to this question is to note that living in a way that is justifiable to others makes it possible to stand in a relationship of mutual recognition and respect. Living according to principles that no one can reasonably reject means that we can live with one another knowing that we respect our distinctive value as persons (154). One way to understand mutual recognition is to see it as form of social harmony. It is, as John Stuart Mill might explain, the “desire to be in unity with our fellow rational creatures” (303).

Significantly, mutual recognition is not something that we are conceptually or normatively bound to seek; it is, rather, a substantive consideration that we value for its own sake. In this respect, Scanlonian contractualism serves as a useful foil to Darwall’s second-person standpoint. Darwall takes considerations such as mutual recognition to be reasons of the wrong kind—they cannot vindicate the intuitive importance of morality—but for the contractualist it is the substantive appeal of standing in a relationship of mutual recognition that serves as the impetus to seek moral agreement. Unlike the second-person standpoint, Scanlon views contractualist deliberation as being shaped by a shared evaluative commitment: On Scanlon’s understanding the value mutual recognition serves as the focal point of moral reasoning and, as I will argue, gives shape to the concept of reasonable reasons.⁹

Just as thinking about the aim of a student-teacher relationship can clarify reasonable reasons, so too can mutual recognition elucidate reasonableness within ordinary morality. Contractualism aims to uncover principles that no one, suitably motivated, can reasonably reject. The key idea here is that “suitable motivation”—understood as the desire to stand in a

⁹ Darwall objects that this is a “reason of the wrong kind” to ground moral norms. This is an objection that I will return to later on.

relationship of mutual recognitions—helps us determine whether or not a proposed principle is reasonable.

Mutual recognition gives content to the idea of the reasonable in two ways. First, the commitment to respecting one another's capacity for rational self-governance brings in certain obvious reasons. Being able to properly live a rational, self-governed life requires us to have access to enough food and water in order to survive and to be free from danger. What makes requests for assistance reasonable is that they have a direct bearing on one's life and they are considerations that anyone would have reason to want. Anyone who is drowning, for example, has reason to want assistance and it is the generality of this consideration that makes it a consideration relevant to the question of a principle's reasonableness.¹⁰

Whether or not a request for assistance is in fact reasonable depends upon the other features of the situation. It may make a difference, for example, whether or not we can save the drowning person without severely injuring ourselves. The present point, however, is simply to note that thinking about the value of rational self-governance gives us a way of framing certain reasons as being, at least, *prima facie* reasonable. The idea here is that anyone committed to standing in a relationship of mutual recognition—anyone committed to respecting the capacity for rational self-governance in oneself and in others—has reason to regard such considerations as reasonable for the purposes of moral deliberation. As Rahul Kumar expresses it, the contractualist formulates the question of reasonableness as: “Can this kind of consideration be considered to be important for being able to live a rationally self-governed, meaningful life?” (17).

¹⁰ There are, obviously, more controversial goods and duties that contractualism should want to accommodate. I do not have room to explore this issue here. For a more full explanation of contractualism's account of reasonableness, see Rahul Kumar's “Reasonable Reasons in Contractualist Moral Argument” (2003).

Mutual recognition gives content to the idea of the reasonable in a second way. The considerations that are relevant are reasonable, not just because they impact our ability to exercise rational self-governance, but because they alter way in which we *relate* to one another as persons. As previously mentioned, the aim of moral deliberation, on the contractualist view, is to be able to recognize one another as self-governors. That is, when we are considering a proposed principle, we are not only considering the potential impact on the individual's life, but also the way in which accepting the principle might change our view of one another as persons. Reasonableness is tied, in part, to how a given principle will alter the character of our social relations. Understood in this way, contractualism also frames the question of reasonableness as: "In what way does this principle alter or structure our ability to *relate* to one another as rational, independent self-governors?"

An example can bring the above claim into sharper focus. Intuitively, it seems reasonable for my neighbor to ask me to watch his children in an emergency and it seems unreasonable for him to demand that I donate a kidney to him. If we try to assess the reasonableness of each request by comparing agent-relative costs then we might think that because the cost of giving up a kidney is so high, I cannot be expected to do so. But, note, however, that the agent-relative cost for my kidney strapped neighbor is even higher. My neighbor, if he is dying from kidney failure, has a very strong agent-relative reason for wanting a principle that requires me to give up my kidney. The question is: why would a principle of aid require that I watch my neighbors' children, but not require that I give up my kidney?

If we frame the question of reasonableness within the aim of relating to one another as separate persons, then we can give support to our intuitions. On the one hand, we are neighbors and this is what makes it reasonable for him to ask me to watch his children in an emergency—it

is reasonable for anyone to want to be able to depend upon others for help in times of crisis, especially when such help comes at a such low cost to oneself. Yet, simultaneously, the limit of the principle may also be shaped by the relational question: a neighbor-to-neighbor relationship governed by such a demanding principle would not be one that would enable us to relate to one another as separate individual persons in charge of their own lives. The idea that someone else would have a claim to my kidney would greatly diminish my ability to see my body as *belonging to me* and such a principle would, in turn, make us unable to relate to one another as separate persons.

To elaborate on the above, consider how I view my body under the principle that requires me to give up a kidney. Under such a principle, I would have a much diminished sense of authority over my body, because it would not belong to me, but to anyone who would be in a position similar to my neighbor. Such a principle would greatly affect our ability to relate to one another as *self-governing* agents. In contrast, a principle that limits kidney demands grants us a particular sense of self-ownership: I have authority over my body in the sense that that nobody else has a claim to it in the way that they would have a claim on a public resource.

It might sound as though the consideration being appealed to here is agent-relative. Under a limited principle of aid, I enjoy the benefit of seeing my body as mine. But, this is not quite right. What makes the consideration reasonable—what makes it a consideration that my neighbor also ought to accept—is that we both have reason to want to relate to one another as self-governing agents. And, in order to see ourselves as self-governors—as persons in charge of our own lives—it is not unreasonable that we should maintain authority over our own bodies.

We can connect the above point to deontic constraints. By rooting the idea of the

reasonable in mutual recognition rather than agent-relativity or well-being, contractualism can understand the value of deontic constraints in distinctly interpersonal terms. Principles that would permit us to randomly interfere with one another would seriously impede our ability to view one another as independent, self-governing persons. Deontic constraints, on the contractualist understanding, have an expressive significance that shapes interpersonal relations: they are prohibitions that enable us to relate to one another as individuals in control of our own bodies.¹¹

To be clear, the value of self-ownership should not be understood as an agent-relative value or as something that we just want for ourselves. The value of self-ownership does not necessarily lie in its ability to protect us from being harmed; i.e. agreeing to restrictions may not, in fact, improve our chances of survival. Instead, the value of self-ownership, as I am presenting it here, ought not to be understood in welfarist terms but as having an expressive dimension that alters the way in which we relate to one another. What seems to be important about self-ownership is that it carries a distinctive relational value. In agreeing to deontic restrictions, we set the terms of moral conduct so that we relate to one another as individual *persons* who direct their own lives, rather than as people who are subservient to the interests of society at large.¹² To put the point another way, having a certain degree of authority over one's self can be understood as a prerequisite for being able to relate to one another as self-governing agents.

Consider how our self-image—our understanding of what we are—changes under different moral principles. In agreeing to restrictions we recognize Alice as an individual independent

¹¹ Here, I am largely following Kumar (1999).

¹² Jeffrey Brand-Ballard argues that contractualism cannot accommodate expressive considerations. According to Ballard, contractualism as a patient-focused theory and only agent-centered theories of restrictions can accommodate the expressive value of constraints (2004, 290). However, as I understand it, contractualism is neither exclusively agent nor patient focused. As John Oberdiek (2007, 122) points out, contractualism is an interpersonal moral theory and it is, in this way, both agent and patient focused.

being. In contrast, if we view Alice as someone who may be permissibly killed, we do not view her as an individual person with her own ends but as someone who belongs to the human community, someone “to be dealt with according to [the community’s] best overall interests” (Quinn 170). Under such a principle, we would not relate to one another as self-governing agents, but as subjects of the human community.

If we think of *Forced Aid* in terms of balancing welfare claims, it is difficult to justify our intuition that it is impermissible for Ted to blind or paralyze Alice. However, thinking of the relationship between Ted and Alice in normative terms—as one that ought to be shaped by mutual recognition—can move us past this stalemate. A relationship of mutual recognition is one in which both parties aim to relate to one another as *individual persons*. Alice, or anyone else, can reject Ted’s demand on the grounds that it would impede this aim. This latter consideration can be justified to Ted on the grounds that he too, insofar as he is committed to respecting Alice’s value as a person, has reason to want self-ownership over his person.

A similar analysis can be made with respect to *Aggregation*. If Alice is required to die for the sake of the group, there is very little sense in which she can maintain a view of herself as an independent person. Alice would be, in Quinn’s terms, a subject of the human community. Of course, it may be objected that Bob and Carol have reason to object to restrictions in *Aggregation*. After all, although we gain a sense of self-ownership in agreeing to constraints, we may also decrease our chances of survival.

Why would it be unreasonable to trade the value of self-ownership for a better chance at not being harmed? From a Hobbesian perspective, we might think that it makes sense to give up the value of self-ownership for an improved survival rate. But, it is worth thinking about what we

would be doing in making this kind of bargain. In particular, we would be trading the value of viewing one another as self-governors for an increase in survival odds. In making this trade, we might ensure that we have a better chance at living a longer life, but we would be living such lives, not as persons who own their bodies or as persons in charge of their own lives.¹³ Although we might gain more years of life, we would lose our distinctive status as self-governing agents.¹⁴ Valuing one another as independent self-governors seems to be deeply incompatible with the sort of bargain that asks us trade self-ownership for an improved chance of survival.

It is worth noting that viewing persons in the way that I have proposed—as self-governors—dovetails with the moral psychology of deontic constraints. If Ted, for example, can push Alice off of a bridge in order to save Bob and Carol, it seems plausible that, outside of a philosophy classroom, this consideration would not even occur to him. From Alice’s perspective, while Ted’s compliance with the constraint could be deliberate—he could weigh the benefits and burdens of compliance—there would be something unnerving about Ted’s moral psychology if this were the case. Not only should Ted refrain from killing Alice, he should not consider the fact that doing so would bring about a better outcome. There is, in other words, a preemptive phenomenology to deontic constraints that seems to preclude Ted’s deliberating in this way.¹⁵

¹³ Anderson notes that it seems suspicious for us to compare such benefits, because they are incommensurable values (1993, 74-79).

¹⁴ Joseph Raz argues that a refusal to compare certain values—such as money and friendship—lends our social relations a dimension of meaning that they would otherwise lack. Refusing to compare the value of one’s friends with money carries a certain social significance insofar as it marks one’s friends as holding a different type of value rather than simply a weighty value (1986, 353). What I am proposing is somewhat similar: in refusing to treat people’s lives in cost benefit terms, our social relations take on a dimension of significance that they would otherwise lack. People are seen not as mere use-values but as having a distinctively higher value. In Elizabeth Anderson’s words we “value people [through deontic restrictions] in higher ways than those in which we value lower goods, such as mere use-values” (1993, 77).

¹⁵ To be fair, not all cases involving deontic constraints will involve the sort of preemptive phenomenology outlined above. In cases involving large numbers, agents may weigh the possible benefits and burdens of their actions. Nevertheless, the force of the consideration remains. In ordinary cases where it is possible to

In more technical terms, constraints seem to function as protected reasons. Ted has a reason not to burden Alice and a reason not to treat other considerations, such as aggregate welfare, as eligible reasons.¹⁶ The argument advanced here offers a plausible way to explain the preemptive phenomenology of constraints. Ted should not weigh the disvalue of killing Alice against the lives of Bob and Carol because to do so would be inconsistent with viewing her as a self-governor. Thinking about Alice in this way would, in other words, count as instance of misrecognition: to treat Alice's life as something to be weighed would be to view Alice as something other than a person in charge of her own life.

Answering The Authority Challenge

These arguments may seem disappointing insofar as they leave the inescapability of deontic constraints elusive. The substantive values of self-ownership and of relating to one another as independent self-governors, though intuitively familiar, is not a formal component of practical reason. Although its substantive value may seem appealing it is not a value that one cannot be rationally refused in the sense that to do so would be a mark of deep irrationality or incoherence. It is, rather, one that cannot be reasonably refused given the shared aim of living on terms of mutual recognition. Consequently, although self-ownership might illuminate the appeal of deontic constraints, it does not produce an unavoidable moral must. Thus, it may be objected that the Authority Challenge remains unanswered.

In his critique of Scanlonian contractualism, for example, Darwall argues:

violate a deontic constraint the moral agent's deliberative process will automatically exclude certain considerations.

¹⁶ For more on protected reasons see McDowell (1979) Raz (1986), Enoch (2011), and Scanlon (1998).

Considerations of desirability...are simply reasons of the wrong kind to ground (second-personal) reasons of moral obligation and rights. If, consequently, the reasons we have to stand in relations of mutual recognition with others are only that it is desirable that we do so (either impartially or from our own point of view), and then no matter how weighty these reasons may be, they cannot ground the distinctive (second-personal reasons) of moral right. The most they can ground, again, is the desire to be such as to recognize the validity of moral claims; they can give us no reason to accept or recognize moral claims in their own terms. (317)

The objection, in other words, is that the desire to see certain claims as normatively stringent does not render them normatively stringent. The fact that constraints would change the character of the social world in an appealing way does not mean that they are, in fact, authoritative. Thus, Darwall will object that contractualism fails to account for the special normativity of restrictions.

In line with this argument, Darwall may argue that second-personal address re-enters Ted's deliberations at the level of authority; i.e. it is Alice's implicit address that grants Ted's duty the special normative stringency that it has. The fact that Ted would be accountable *to* Alice—the fact that she could say “you can't do that to me”—may be what accounts for the intuitive authority of deontic constraints. Connecting constraints to second-personal address would give Alice a conceptual trump card: when she addresses Ted, her word is final. Unlike the relational value of self-ownership, second-personal constraints would be something that we cannot conceptually refuse.

To be clear, Darwall is surely correct to argue that Alice has the standing to demand compliance or compensation from Ted. It would be strange to argue that Alice cannot demand his compliance nor demand compensation if Ted has wronged her. However, it is questionable as to

whether or not holding someone accountable—demanding compliance or compensation—fully accounts for the authority of constraints. Nor is it obvious that we should think of normative inescapability as being as important as Darwall presents it to be.

To support these claims, consider what it means to take Darwall's wrong kind of reason objection seriously. The objection argues that although it is appealing to relate to one another on terms that respect our value as self-governors, there is a sense of 'should' that remains to be elucidated. If this is the case, then Darwall's objection can be rephrased as follows:

“Yes, this action satisfies mutual recognition—it facilitates our living together on terms that affirm the value of our being able to live self-directed, meaningful lives—but why should I?”

The above question is not incoherent, but it is somewhat disconcerting. It is unnerving, in part, because it suggests that without an additional reason we ought to give up on morality.

Additionally, it suggests that an agent might be moved to obey constraints without caring about their substantive value. Darwall's agent, supplied with an inescapable should, would not necessarily care that constraints reflect your distinctive value as person—mutual recognition would be irrelevant. In fact, such an agent might even wish that things could be otherwise, regretting the fact that they are conceptually committed to accepting reasonable reasons. Second-personal authority, in other words, might be regarded as a formal, yet unwelcome, metaethical truth. This is worrying: why should we view the inescapability of the second-personal circle as a benefit rather than some sort of moral snare?

It may be objected that this cannot be possible: if we are conceptually committed to recognizing the authority of other persons, and hence the authority of deontic constraints, then there must be a sense in which we are required to view this fact in a positive light.

This is not true. There is nothing inherent in the idea of an inescapable truth that requires us to regard it in a positive light. As Derek Parfit explains: “We can regret truths even when it is logically impossible that these truths be false” (1984, 175). The Pythagoreans, for example, wanted the square root of two to be a rational number. Even though it is impossible for things to be otherwise, the Pythagoreans deeply regretted the fact that the square root of two cannot be a rational number (175). It may also, for example, be an inescapable truth of personal identity that we could not have been born earlier than we actually were. It might be impossible for us to have been born in the 18th Century. However, although this fact of identity may be unavoidable, this does not mean that we cannot regret not having lived during the 18th Century. In short, there is nothing about inescapable truths that would prevent us from regretting their validity.

We can, I think, regret relationships of authority and accountability much like the Pythagoreans regretted that the square root of two cannot be a rational number. Suppose, for example, that I am attempting to park my car in an area where the city of Toronto says that I cannot park. I am accountable to the parking authority of Toronto: they have authority to demand that I not park my car where they say so and they can hold me accountable, through some form of penalty or sanction, if I do so. In grasping that this would be a parking violation, I am recognizing the relationship of authority and accountability that obtains between the city of Toronto and me. However, although I may recognize such a relationship, nothing prevents me from regretting the fact that I cannot, in fact, park my car here. Nor does anything prevent me from regretting the fact that the city of Toronto has this authority over me. I can, moreover, decide against parking simply because I would be held accountable for doing so; I need not care about the reasons that there are for the no parking rule (such as the fact that it is reserved for disabled people).

Ted could, on Darwall’s view, reason about killing Alice in the same way that I reason

about violating the parking restriction. Suppose, for instance, that Ted deliberates about whether or not to push Alice off of the bridge, but decides, ultimately, that he cannot do so because the conceptual nature of Alice's address prevents him from doing so. Ted decides that although he would like to save Bob and Carol by killing Alice, he does not want to be held accountable to Alice. In this case, even though Alice has not been harmed in any way, and even though Ted has complied with her second-personal authority, there still seems to be something amiss with Ted's view of Alice. In particular, what seems to be askew, from Alice's perspective, is that Ted does not see anything valuable about relating to her as an individual person who has a certain degree of sovereignty over herself. Ted might view Alice as holding the conceptual trump card, but he does not necessarily recognize her as the sort of person she takes herself to be.

If Ted complies with the constraint simply because he would be held accountable—and not because he values the relationship between him and Alice—then he treats Alice like a parking authority: she is someone to be obeyed, but she is not necessarily someone who he should value as a separate person. This reveals that what is important about constraints is not simply that other people can hold us accountable for complying with them; rather what seems to be important about constraints is that they tie in with a substantive way of valuing other people. Accordingly, I do not think that we should deem considerations of substantive value as reasons of the wrong kind. Instead, I would submit that acting purely for reasons of accountability seems to be acting for a reason of the wrong kind.

The above point undermines the allure of attempting to reveal interpersonal morality as being a formal, unavoidable component of practical reason. Even if we can show Ted that he would be violating a conceptual requirement of second-personal address, this would not be the same thing as showing Ted that Alice is a person who matters. Making constraints conceptually

inescapable does not explain why we find them to be important, or why they imbue our relations with a level of meaning they would otherwise lack.

In conclusion, this chapter has argued that developing an interpersonal account of constraints requires us to think of interpersonal morality as being somewhat different from Darwall's second-person standpoint. I have argued that properly responding to the Irrationality Challenge necessitates that we have a way of delineating between reasonable and unreasonable demands. Neither the agent-relativity nor the conceptual structure of second-personal reasons can accomplish this latter task. Additionally, I have argued that accommodating the problem of reasonable reasons reveals that Darwall's shift from second-personal address to a regulative ideal fails to accomplish anything of significance.

In the second half of this chapter, I have presented T.M. Scanlon's contractualism as an alternative account of interpersonal moral reasoning. In particular, I have argued contractualism offers us a way of framing the question of reasonable reasons by appealing to the substantive value of mutual recognition. Reasonable reasons, on the contractualist view, are those that have bearing on our ability to relate to one another as distinct individual persons. This is not a conceptual presupposition but a substantive way of valuing human life. I have argued that deontic restrictions can be understood as duties that no one could reasonably reject insofar as they too share the ideal of relating to one another as independent, self-governing persons. Constraints, on the contractualist view, structure the very way in which we relate to one another as persons.

Darwall argues that considerations of substantive value are reasons of the wrong kind and that, instead, focusing on accountability can only capture the authority of restrictions. This, I have argued, is a mistake: acting merely for reasons of accountability is a reason of wrong kind. Divorced from considerations of substantive value, the second-person standpoint gives us an

impoverished account of morality's authority: even if constraints could be shown to be conceptually unavoidable, this does not necessarily explain their importance to us.

Chapter 3

Wronging, Wrongness, and Inviolability

Non-consequentialists often attempt to capture a familiar, if slightly elusive, sense of moral wrongness. In particular, many non-consequentialists give a central role to the idea that there is a distinction to be made between acting wrongly and wronging someone. To explain, consider the difference between my duty not to trample sunflowers and my duty not to trample you. In the case of sunflowers, I might act wrongly in trampling them without good reason, but it does not seem that I can wrong them by doing so. In juxtaposition, if I trample *you* it does not seem that I have merely acted wrongly, but that I have *wronged you* in particular because I owe it to you not to do so. Call this the Wronging Distinction.

Arguably, the Wronging Distinction is important because it offers one possible way of explaining the intuitive authority of moral reasons. Deontic constraints, for example, seem to carry an unusual normative stringency insofar as they seem to be non-optional. Ted, for instance, seems to have a categorical reason not to kill Alice—even if doing so will save Bob and Carol—and not simply an elective consideration to refrain from doing so. The absolute priority of Ted’s deontic constraint might be explained by the fact that he owes it to Alice in particular not to kill her.¹⁷

In *Intricate Ethics*, Frances Kamm elaborates on a theme from her earlier work—that it is

¹⁷ Recently, non-consequentialists such as Stephen Darwall, R. Jay Wallace, and Paul Hurley have attempted to explain the authority of morality by emphasizing its interpersonal nature. Although Darwall, Wallace, and Hurley all differ in how they understand the nature of interpersonal morality, they all share the thought that it is the relational structure of morality that explains its normative stringency. See Stephen Darwall (2006), R. Jay Wallace, (forthcoming), and Paul Hurley, (2009).

the status of persons as inviolable that justifies why it is often impermissible to do what will best promote the good—in order to incorporate the Wronging Distinction. By reconciling inviolability with the Wronging Distinction, Kamm attempts to add moral gravity to her account of deontic constraints by imbuing them with a dimension of interpersonal authority.

I argue that the inviolability approach cannot make sense of the capacity to be wronged. The structure of this chapter is as follows. To begin, I give a brief explanation of the inviolability approach, highlighting the way in which Kamm’s earlier version overlooked the Wronging Distinction. In turn, I examine the way in which Kamm has modified the inviolability account in order to include the capacity to be wronged. Following this, I argue that the inviolability approach faces the following dilemma: either Kamm’s explanation of wronging renders the concept of inviolability superfluous or the capacity to be wronged is a mysterious, metaphysical accident. Finally, I argue that my diagnosis of Kamm’s inviolability approach has general implications for the way that we think about interpersonal authority. To anticipate, I argue that wronging cannot be wrung from wrongness: interpersonal authority cannot be extracted from non-interpersonal explanations of permissibility. In order for the Wronging Distinction to carry normative weight, it must play a formative role in accounting for standards of permissibility.

Deontic Constraints and Inviolability

The inviolability approach is most easily explained by examining its proposed solution to the paradox of deontology. In its most basic form, inviolable status is a conjunction of propositions detailing the ways in which it is impermissible to treat a given entity (28).¹⁸ The number of constraints associated with a given entity corresponds to its level of inviolability; i.e.,

¹⁸As I will note later on, in *Intricate Ethics* Kamm draws a distinction between moral status and inviolability. For present purposes, I will focus solely on inviolability.

the more constraints that are connected to an individual thing—the more ways in which it is impermissible to treat it—the higher its level of inviolability. For instance, if more restrictions are attached to human beings than there are attached to animals then this means that human beings have a higher level of inviolability than animals.

Human beings, according to Kamm, have a very high level of inviolability because it is impermissible to do a great many things to humans. Humans are not, of course, maximally inviolable; it is, for example, consistent with our inviolable status to flip the switch in Judith Jarvis Thomson's Trolley Case (25).¹⁹ But, human beings are unique, according to Kamm, in that our level of inviolability is high enough that it forbids violating one person's right not to be killed even when doing so will minimize rights violations overall. That is, human inviolability is especially high insofar as it forbids the performance of minimizing violations—acts that perform a rights violation in order to minimize rights violations overall. Ted, for example, is forbidden from killing Alice even if doing so will prevent a villain from killing Bob and Carol, because human inviolability is especially high.

In order to explain how this responds to the paradox of deontology, consider what it means for Ted to be forbidden from performing a minimizing violation. If it is *impermissible* to kill Alice, then humans possess a very high level of inviolability. Conversely, if it is *permissible* to kill Alice, then everyone will have a lower level of inviolability. Notably, in the latter situation Bob and Carol will live, but they will not possess the same high level of inviolability that they would in the former situation (28). Even if Bob and Carol are killed, they die with a higher level of inviolability than they would have if it were permissible to kill Alice. In this sense, the

¹⁹ As Michael Otsuka (1997) notes, this feature of the inviolability approach implies that there could be beings with an even higher level of inviolability than human beings.

impermissibility of killing Alice signals that human beings have an especially high level of inviolability.

Kamm argues that the inviolability approach avoids the paradox of deontology because inviolability cannot be maximized. Inviolability is a value that we recognize and respect in our actions, but it is not a value that we can bring about by minimizing harms; rather, it is *the fact of impermissibility* that secures the high moral status of Bob, Carol and Alice. This means that Ted cannot be moved to kill Alice out of a concern to protect the inviolable status of Bob and Carol because doing so would be *self-defeating*: if it were permissible to kill Alice then everyone would have a lower level of inviolability (28). Someone like Ted can acknowledge the high level of human inviolability that there is by respecting the constraint on killing Alice.

The primary appeal of the inviolability approach is that it explains deontic constraints as adequately consistent. Minimizing violations cannot promote inviolability, because if they were permissible then humans would have a lower level of inviolability. A prohibition on minimizing violations reflects the moral fact that human beings possess an exceptionally high level of inviolability. This leads Kamm to claim that the inviolability approach clears the air of paradox: a concern with human inviolability cannot lead us to perform actions that would—if they were permissible—lower human inviolability.

Elsewhere in the literature, it has been questioned as to whether or not the inviolability approach successfully answers the paradox of deontology.²⁰ Setting this question aside, I would like to draw attention to the fact that the concept of inviolability does not, by itself, provide a

²⁰ For an argument that Kamm fails to solve the paradox of deontology, see Kasper Lippert-Rasmussen (2009).

ready explanation of the Wronging Distinction. Pointedly, if inviolability is understood as a list of restrictions detailing how we may treat an entity, then animals and plants also possess inviolability insofar as they too cannot be treated in certain ways. Yet, although we may be morally required to refrain from the wanton destruction of plants and animals, it does not seem that we owe it to them not to do so. As such, there is nothing inherent in the concept of inviolability that requires us to think of Ted's duty as being *owed to Alice*. As a result, the early inviolability approach leaves the Wronging Distinction unexplained. In the absence of such an explanation, the duty not to kill Alice appears to be similar to the duty not to destroy sunflowers: it may be a moral duty that we have, but it seems to be a duty that is owed to no one in particular.

Kamm's earliest presentation of the inviolability approach actually emphasized the non-interpersonal nature of inviolability. In particular, Kamm argued that the inviolability approach narrowed the distinction between consequentialism and non-consequentialism insofar as both are concerned with good states of affairs. In Kamm's words:

If we conceive of deontological constraints as expressions of the status of each person and this status of inviolability is itself a good, then to some degree we narrow the distinction between consequentialism and deontology. For consequentialism is concerned with bringing about what some theory of value determines to be the best state of affairs. If the high status of persons is a good whose existence is reflected in constraints, then, in this sense, constraints are connected with a good and a good state of affairs. (1999, 388-389)

To explain Kamm's claim of reconciliation, consider the connection she draws between deontic constraints and the appeal to states of affairs. A higher level of inviolability, according to Kamm, ensures a better state of affairs because "[T]he world is, in a sense, a better place, as it has more important creatures in it"(386). As such, in respecting deontic constraints we are not producing a

lesser state of affairs; rather, we are recognizing the value of inviolability and thus, acknowledging the better state of affairs that we live in. This narrows the gap between consequentialism and non-consequentialism because the comparison between one state of affairs and another forms the crux of Kamm's justification for deontic constraints: Ted should not kill Alice because a world in which doing so would be permissible would have less inviolability and, consequently, would be a worse world.

To be clear, the inviolability approach does not ask us to maximize inviolability. Unlike welfare, the value of inviolability does not require promotion; even if Ted kills Alice this does not change the moral fact of their high inviolable status. Ted's duty is not to bring the best state of affairs about and, in this respect, it remains distinctly non-consequentialist. Despite this difference however, the early inviolability approach remains similar to a consequentialist approach insofar as it reconciles deontic constraints with the idea that permissibility is indexed to the best state of affairs. Ted's duty is not to bring the best state of affairs about, but it is secured by the comparison between one state of affairs and another.

Notably, narrowing the gap between consequentialism and non-consequentialism in this way makes the duty to respect inviolability non-interpersonal. If the motivation to respect constraints is a concern to act in accordance with the value of a state of affairs, then there is no sense in which the duty to respect Alice's rights are owed to Alice in particular. Instead, the motivation to respect Alice's rights is tantamount to a concern to recognize the level of inviolability in the world and to act in accordance with the best state of affairs that we happen to live in. As such, the duty to respect inviolable constraints is much like the duty to promote well-being: it may be a moral duty, but nothing about it invites the thought that it is owed to anyone in particular. By connecting the value of inviolability with a good state of affairs, the early

inviolability approach eclipses the Wronging Distinction and renders the duty to respect inviolability a non-interpersonal duty.

Interpersonal Inviolability?

In *Intricate Ethics*, Kamm responds to the above objection by attempting to meld the inviolability approach with the Wronging Distinction. In particular, Kamm moves away from her previous emphasis on valuable states of affairs by arguing that:

...ideas of respect for persons and the dignity of the person are connected to the idea that one *owes it to a* person to behave in certain ways... This is in contrast to it being simply wrong to treat someone in certain ways because, for example, one owes it to God not to or because it would not maximize utility to do so, and one has a general duty as a rational being—but not owed to anyone—to maximize utility. (230)

Put another way, Kamm's qualifies the inviolability approach so that that the duties associated with human beings are "directed duties": duties that are owed to the particular person that is acted upon (230). On this understanding, "[W]e wrong the person if we violate the constraint, as we owe it to him not to do it"(231).²¹ Accordingly, we are not moved to respect duties that comprise inviolable status because of a valuable state of affairs, but because such duties are owed to particular persons.

Of course, though the Wronging Distinction is intuitive, its incorporation into the inviolability account calls for explanation. To say that the duties constituting the inviolable status of human beings take a directed form simply asserts what is at issue. Moreover, as previously mentioned, if we define inviolability simply as a list of restrictions, then entities such as tigers

²¹ These quotes are from Kamm's chapter on rights. Inviolability and rights, on Kamm's view, are connected insofar as the rights we have make up, or express, our inviolable status (253).

and sunflowers also possess inviolability insofar as we have duties not to treat them in certain ways. Yet, as Kamm herself emphasizes, the constraints associated with such creatures do *not* take a directed form (230). Again, there is no necessary connection between *generic* inviolability—inviolability understood as a conjunction of propositions concerning impermissibility—and the capacity to be wronged.

How can the inviolability account integrate the Wronging Distinction? It might seem that Kamm can explain the Wronging Distinction by emphasizing that inviolability comes in different levels: something has a higher level of violability if there are more ways in which it is impermissible to treat it. This might explain why the inviolability of persons differs from the inviolability of tigers and sunflowers. Human beings possess a higher level of inviolability because there are more ways in which it is impermissible to treat us. As previously explained, Kamm argues that the level of inviolability attached to human beings is high enough to issue deontic constraints: it is forbidden to kill one innocent person even if doing so prevents three other innocent people from being killed (28). In contrast, the level of inviolability associated with tigers is not high enough to generate deontic constraints. Arguably, it seems morally permissible—though perhaps not morally obligatory—to kill one tiger in order to prevent three others from being killed.

This difference seems to open up one way of explaining the Wronging Distinction. The presence of directed duties might be thought to correlate with a particular *level* of inviolability. One might, for example, argue that although inviolability *simpliciter* may not generate directed duties, a sufficiently high level of inviolability might. In order to explain the Wronging Distinction all that Kamm needs to do is to identify the degree of inviolability that creates the capacity to be wronged.

This line of response is not available to Kamm: a high level of inviolability does not necessarily entail the presence of a directed duty. For example, while our high level of inviolability generates deontic constraints, so too does the high inviolability attached to entities that carry symbolic value. As Kamm herself notes, “flags or holy entities, should not be destroyed for the sake of saving others of their type” (256). Yet, even if there are deontic constraints associated with flags or holy entities, these constraints do not take a directed form. Instead, these duties seem to be directed towards God, or to no one in particular, rather than to the flags or religious icons themselves. In such cases, what we have are beings that possess a high level of inviolability—high enough to issue deontic constraints—but without the presence of a directed duty. Thus, the high inviolability of human beings cannot explain our capacity to be wronged.

A more promising explanation might be mounted by noting the distinction that Kamm develops between moral status and inviolability. Moral status is similar to, though not synonymous with, inviolable status: like inviolability, something’s moral status details the ways in which it is permissible to treat it (227). However, Kamm does not treat moral status and inviolable status as equivalent concepts: inviolability is something that makes up our moral status, but it is not the only feature that does so (227). For present purposes, what is important to note about Kamm’s explanation of moral status is that she views moral status as changing the nature of an entity’s inviolability. In particular, moral status can alter what it means to respect duties of inviolable status.

To support this claim, Kamm argues that the moral status of an entity supervenes upon its non-moral properties. Drawing a connection between non-moral properties and moral status gives

rise to the idea that there are different *types* or *genres* of inviolability.²² To explain, consider the following examples. A tree or a work of art “may count in its own right in the sense that it gives us reason to constrain our behavior toward it (for example, not to destroy it) just because that would preserve this entity...In that sense it counts morally” (228). In contrast, the moral status of an animal supervenes upon its consciousness or sentience, which means that it can *benefit* from the duties attached to it. This means that “when I save a bird, I can do it for its sake, because it will get something out of continuing to existence” (229). One can refrain from destroying trees, artworks, and animals, but while one can act *for the sake of* the animal, one cannot act for the sake of an artwork or a tree.

In light of the above, Kamm argues that the non-moral properties that make up something’s moral status can change what it means to respect its inviolability. Since the moral status of a tiger is based upon its capacity for pain, respect for a tiger’s inviolability can be in the interests of a tiger, and this entails that we can respect a tiger’s inviolability *for the sake of* the tiger itself (228). By contrast, respect for sunflower inviolability cannot be for the sake of a sunflower itself since plants do not have the non-moral properties that would give them interests (229). In this fashion, tigers and other animals can be understood to have a different type or genre of inviolability than plants, in virtue of the non-moral fact that they have interests.

Citing T.M. Scanlon, Kamm picks out “judgment-sensitive attitudes” as the non-moral property upon which human moral status supervenes (232). Creatures with judgment-sensitive attitudes can act in accordance with reasons and “are entities to whom we can owe certain

²² Here I am following Kasper Lippert-Rasmussen’s interpretation of Kamm’s connection between moral status and the value of inviolability: that differences in moral status give rise to different types or genres of inviolability (2009, 167-170). Kamm herself is not explicit about the way in which the two concepts are meant to interact. Nonetheless, I take it that Lippert-Rasmussen’s interpretation is the most plausible understanding of the connection between moral status and inviolability.

treatment”(232). In light of this fact about human moral status, it is open for Kamm to claim that human beings have a different type of inviolability than plants and animals. Since our moral status supervenes upon judgment-sensitive attitudes, the duties that make up our inviolable status take a directed form. In this fashion, the supervenience relation makes the Wronging Distinction intelligible: when we disregard our duties towards human beings we not only act wrongly, we wrong *them* in particular because human inviolability, insofar as it supervenes upon judgment-sensitive attitudes, takes an interpersonal direction.

Implicit in Kamm’s claim is the thought that individuals can be wronged because they can recognize that they instantiate the value of inviolability. In this sense, when Ted kills Alice he acts wrongly because he disregards the level of her inviolability *and* he wrongs her, in particular, because she can recognize that he violates a duty attached to her. Judgment sensitive attitudes, understood in this way, change the nature of human inviolability insofar as they make us capable of recognizing the duties that comprise our inviolable status. If inviolability “already resides *in* persons and we act in the light of it”, then these persons can recognize the duties attached to them and, in this way, are capable of being wronged (29).

Although this explanation seems intuitive, it does not fully explain the Wronging Distinction. To see this, consider how a simple version of utilitarianism could offer a similar explanation of wronging. If Ted should not kill Alice because doing so would lead to a decrease in utility and if Alice can recognize that it is her utility that is at stake, then there is a straightforward sense in which Ted’s duty and Alice’s ability to recognize this duty overlap. Establishing this overlap, however, does not explain why the duty should be thought of as being owed to Alice in particular. Although Alice can recognize the she instantiates the duty, the duty is not really owed to her in particular because she is simply a particular manifestation of the general,

non-interpersonal value of utility.

In a similar vein, the duty to respect Alice's inviolability seems to be parasitic upon the general duty to respect inviolability wherever it is encountered. Alice may be able to recognize that her presence in a situation brings certain duties in tow, but this does not entail that these duties are owed to her in particular. Additionally, the claim that inviolability resides *in* persons is of no real help in establishing the Wronging Distinction. The fact that people instantiate certain values does not explain why the duties associated with such values take an interpersonal direction. The value of utility, for instance, resides in persons—in the sense that we are a source of utility—but it also resides in creatures that cannot, intuitively, be wronged. Similarly, inviolability—understood simply as a list of restrictions—also resides, in lower levels, in plants and animals.

The above objection may only demonstrate that Kamm's explanation is incomplete. Nothing argued thus far has demonstrated that Kamm cannot work out, in more detail, the connection between inviolability, judgment-sensitive attitudes, and wronging. However, as I shall go on to argue, working out the supervenience connection in more detail can only result in a pyrrhic victory for Kamm. Explaining inviolability as a value that supervenes upon non-moral properties creates the following dilemma: either the concept of inviolability is redundant or the capacity to be wronged is a metaphysical accident.

Let me explain. Consider, first, the redundancy horn of the dilemma. Put simply, incorporating non-moral properties and, the values that come with them, into the inviolability approach renders the concept of inviolability explanatorily unnecessary. To see this, consider what the supervenience move means for something like a work of art. If we think of the

inviolable status of a painting as supervening upon its non-moral properties—such as the fact that it is beautiful—it is not clear what is gained by saying that it has a particular type of moral status. The fact that the painting is beautiful, for instance, seems to give us reason to admire it, to hang it on our wall where people can see it, and not to destroy it. The non-moral property of aesthetic beauty, in other words, seems to be reason providing in a way that makes references to its particular status explanatorily unnecessary. Another way to put this objection would be to say that the value of its inviolable status “passes the buck” to the value of the artwork’s non-moral properties. To say that the painting is also inviolable is to put on an explanatory fifth wheel.²³

A similar analysis can be made with respect to tigers. If the moral status of tigers supervenes upon their sentience, then again, the concept of inviolability seems to pass the buck from inviolability to the property of sentience and other evaluative concepts. To see that my action causes the tiger pain –and to see this pain as bad thing—is to see that I have a consideration that counts against my action; i.e., the tiger’s pain gives me a reason not to harm it. Again, what is not clear is what is to be gained by saying that that the tiger also has a certain level of inviolability. More precisely, to say that the tiger has a certain moral status would not seem to give me an additional reason not to harm the tiger.

To be clear, I do not wish to claim that we can make judgments about what we have reason to do without relying on evaluative concepts whatsoever. To claim, that we should not torture tigers for fun, relies on the evaluative claim that gratuitous suffering is a bad thing. Nevertheless, the point to be made is that we can see the fact of a tiger’s pain as reason providing, without

²³ The term “buck passing” is borrowed from Scanlon’s so-called buck passing account of goodness (1998, 96-100). For present purposes, I wish to adopt the basic structure of the argument without necessarily endorsing Scanlon’s analysis of goodness. Note, also, that in using this term I do not mean to claim that all values pass the buck to non-natural properties; rather, I am simply claiming that inviolability passes the buck to non-natural properties and other evaluative concepts.

making the additional claim that we need to understand tigers as holding some sort of unique moral status. Instead, constraints on harming tigers can be understood as being rooted in the fact that tigers are sentient. Similarly, the aesthetic properties of a painting give us reason not to wantonly destroy it.

To speak of tiger inviolability or artwork inviolability would seem, at best, to be a form of moral shorthand—a way of gesturing to the fact that we have more specific reasons to treat such entities in particular ways. For entities such as tigers and paintings, it is easy to see the way in which Kamm’s appeal to supervenience makes the concept of inviolability redundant.

In a similar fashion, the appeal to judgment-sensitive attitudes threatens to make the concept of human inviolability superfluous. Notably, for someone like Scanlon, the capacity to be wronged is not simply attached to creatures with judgment-sensitive attitudes; rather, wronging comes about because of the way that judgment-sensitive attitudes function (169). To illustrate, consider what it means for someone to govern themselves in accordance with reasons. At a basic level, rational self-governance involves the ability to plan and revise one’s actions: to see certain considerations as counting against or in favor of a particular action and to plan our lives in accordance with these assessments (106). Simultaneously, at a secondary level, we can respect the value of rational self-governance in other human beings by living according to principles that that they could not reasonably reject insofar as they too were motivated by this ideal (154). At this secondary level, we are able to ask other persons to justify their behavior to us: to explain why their conduct conforms to standards of conduct that no one, suitably motivated, could reasonably reject.

It is important to note that the contractualist understanding of justifiability makes no

reference to states of affairs. For an agent motivated to live on terms that no one could reasonably reject, the relevant reasons for the purposes of moral argument will be those that bear on an agent's ability to live a self-governed, meaningful life. Notably, by limiting the relevant moral reasons to those that have an impact upon one's ability to live on terms of self-governance, contractualism pushes states of affairs out of the picture. Appeals to states of affairs are impersonal because an overall state of affairs is better or worse for no one in particular. The idea that a state of affairs is "best overall" or "worst overall" has no bearing from the perspective of the individual. Of course, a good state of affairs can be good for me if my position in it is good one, *but* it is not good for me qua good state of affairs: the fact that something is best *overall* does not have relevance from the individual's point of view. This latter feature makes contractualism foundationally interpersonal in the sense that there are no appeals to impersonal values.

The importance of this last point is worth emphasizing. On the contractualist view, it is the ability to justify our behavior to others that gives rise to both wronging *and* standards of permissibility: others wrong us, and thereby act impermissibly, when they treat us in ways that can be reasonably rejected. It is a distinctive mark of contractualism that wronging underwrites standards of permissibility; the possibility of justifying one's conduct to another forms the basis for standards of permissible conduct.

Suppose, for example, that Ted is hitting golf balls off of his roof and into Alice's backyard. Fortunately, although Alice is in her garden, none of the golf balls strike her. On the contractualist picture, Ted wrongs Alice because he acts in a way that cannot be justified to her. It seems uncontroversial that, for the purposes of living a self-governed life, Alice has reason to want to be free from the sort of random harm that Ted exposes her to. Alice, in virtue of her value as a self-governing agent, can reasonably expect that Ted take her safety into account. What is

distinctive about contractualism is that it is the fact that Alice can reasonably expect this of Ted, that makes his action impermissible and an instance of wronging. In violating Alice's legitimate expectation—that she holds in virtue of being a rational self-governor—Ted wrongs Alice because he cannot justify his behavior to her.

The importance of the above example is two-fold. First, Ted has reason not to treat Alice in the way that he does because she could reasonably reject the permissibility of his conduct. This makes the explanatory power of inviolability questionable: we have reason to treat other people in certain ways, not because they hold a certain level of inviolability, but because they are rational self-governors. Once Ted has such a reason, it is not obvious what would be gained by adding that such behavior is incompatible with Alice's moral status. Explaining the capacity to be wronged as a function of judgment-sensitive attitudes shifts the explanatory weight off of the concept of inviolability and on to the property of rationality. Secondly, the contractualist framework offers an explanation as to why the duty is owed to Alice in particular: Ted wrongs Alice because he violates her reasonable expectations.

Again, appeals to inviolable status seem to operate as a sort of moral shorthand. At best, when we claim that we cannot be treated in specific ways because we are creatures with a certain value, we are really claiming that there are specific reasons having to do with us that count against certain actions. At worst, assertions of inviolable status become redundant when the reasons that disfavor an action have already been made explicit. Thus, Kamm's appeal to judgment-sensitive attitudes renders the concept of inviolability explanatorily superfluous.

It may be objected that human inviolability need not be a redundant concept. Although we may not need the concept of inviolability for tigers, art, and plants, that does not mean that we do

not need it in the case of human beings. Basic duties, such as “don’t torture tigers” can be provided by properties such as sentience, but more complex duties such as “do not kill one person in order to minimize overall killings” do not follow straightforwardly from an appeal to rationality. Since the constraints associated with human beings function in a more complex way, it may be the case that concept of inviolability is needed to make sense of the way in which they function. We may need the concept of inviolability, in other words, to make sense of a prohibition on minimizing violations. Hence, in the case of human beings, it may be objected that the notion of inviolability has not yet been revealed to be redundant.

I will not defend the claim that contractualism can produce deontic constraints here.²⁴ For present purposes however, we can note that even if appeals to inviolability are not superfluous—even if the concept does real theoretical work—the inviolability approach still faces the second horn of the dilemma. In particular, Kamm still faces the charge that the Wronging Distinction, on the inviolability approach, is a metaphysical quirk rather than an important moral feature.

To elaborate, note that a non-redundant account of inviolability leaves the connection between wronging and judgment-sensitive attitudes obscure. If the capacity to be wronged does not arise from the way that judgment sensitive attitudes function—or what they make us capable of doing—then it could be the case that the capacity to be wronged could attach itself to the non-moral properties of tigers and sunflowers instead. Without such clarification, moreover, the connection between the Wronging Distinction and rationality seems to be a serendipitous accident: rationality does not generate the capacity to be wronged, it *just happens* to be attached to creatures with judgment-sensitive attitudes. Put bluntly, it is not enough to assert that the capacity to be wronged is connected to judgment sensitive attitudes: Kamm needs to give an

²⁴ See pages 38-47.

explanation as to how judgment-sensitive attitudes transform acts of wrongness into acts of wronging. Without a more robust connection between rationality and the capacity to be wronged, Kamm's supervenience explanation hardly seems to be an explanation at all.

It may be noted that although Kamm follows Scanlon in connecting the capacity to be wronged to the possession of judgment-sensitive attitudes, she does not follow Scanlon in connecting the Wronging Distinction with the ability to demand justification. In fact, Kamm explicitly denies that the capacity to be wronged is tantamount to the capacity for reasonable rejection: "It is tempting to think that those we could wrong in the course of a permissible act would also have a higher moral status than those not so wronged" (231). In other words, the fact that we could be *permissibly* wronged would indicate that we have a very high moral status. Thus, Kamm argues that a "wrong act is not a necessary condition of wronging" and that, consequently, wronging and reasonable rejection come apart (468).²⁵

The above objection raises a plethora of questions regarding the relationship between wronging and permissibility that extend beyond the scope of this chapter. But, for present purposes, we might note that *even if* there is such a thing as permissible wronging it does not remove the above dilemma. Any attempt to avoid the first horn of the dilemma—any attempt to avoid the redundancy objection—simply draws our attention to the second horn. That is, if Kamm avoids the first horn of the dilemma by raising the possibility of permissible wronging, then the inviolability approach still faces the task of explaining the connection between wronging and

²⁵ Kamm expands upon the original Trolley Case to make this point. If we flip the switch, it does not seem impermissible for the one person stranded on the track to resist our doing so (say, by redirecting the trolley back towards the five). Kamm argues that her permissible resistance indicates that we have permissibly wronged her (2007, 31). I do not share the intuition that you may permissibly redirect the trolley. This requires further argument, but the present point is that Kamm's diagnosis is not necessarily the most obvious. For more on the infringement/violation distinction see Oberdiek (2003).

judgment-sensitive attitudes. Even if Kamm finds Scanlon's account of wrongness lacking, we can still expect her to provide an alternative explanation of the relationship between judgment-sensitive attitudes and the capacity to be wronged. Without such an explanation, moreover, the capacity to be wronged appears to be a mysterious, metaphysical accident. Thus, avoiding the redundancy objection by making a claim about permissible wronging only serves to make the connection between wronging and judgment-sensitive attitudes unintelligible.

Leaving the connection between wronging and rationality opaque diminishes the extent to which we can explain the significance of the Wronging Distinction. Kamm suggests that wronging adds a dimension of moral gravity to human inviolability: "it is tempting to think that these entities [those who can be wronged] have a *higher* moral status than other entities which also have moral status" (230). But, if the connection between wronging and judgment-sensitive attitudes remains obscure, then the importance of wronging seems to be equally unclear. Without a substantive connection between wronging and rationality, it is difficult to see why the Wronging Distinction should be considered to be an important feature of our moral status instead of an arbitrary quirk about human beings. Moreover, for those who are not already convinced that rationality carries an additional sense of moral importance, the connection that Kamm draws between wronging, rationality, and their respective importance will simply beg the question. Thus, insisting on a blurry connection between wronging and rationality prevents Kamm from explaining the significance of the Wronging Distinction.

In summation, Kamm's appeal to a supervenience relation between moral status and judgment-sensitive attitudes gives rise to the following dilemma. On the first horn, the appeal of judgment-sensitive attitudes threatens to make inviolability a redundant concept. If the connection between rationality and wronging is robust—if wronging arises from reasonable rejection—then

rationality can yield a satisfactory account of wronging and standards of permissibility that no one can reasonably reject. This threatens to make the concept of inviolability explanatorily superfluous: if reasonable rejection can generate standards of permissibility, there is simply no need to appeal to the concept of inviolability.

Kamm may resist the first horn by arguing that inviolability remains explanatorily vital in cases where reasonable rejection cannot provide a satisfactory explanation of our moral duties. However, as I have argued, the success of this strategy brings up the second horn of the dilemma. Specifically, the connection between rationality and the capacity to be wronged, on a non-redundant account of inviolability, appears to be a metaphysical accident—wronging *just happens* to be attached to creatures with judgment sensitive attitudes. Embracing this horn comes at a significant cost: without a more robust explanation of the connection between wronging and judgment-sensitive attitudes, the moral significance of the Wronging Distinction seems to be a mere ad-hoc appeal to intuition. If wronging has nothing to do with what makes human beings distinct creatures, then the importance of the Wronging Distinction seems unexplainable.

Interpersonal Authority

The dilemma faced by Kamm's approach to deontic constraints has wider implications for the way that we think about the authority of the Wronging Distinction. Specifically, any moral theory that begins with a non-interpersonal account of wrongness cannot incorporate the Wronging Distinction as a *significant* moral feature. That is, non-consequentialists who advocate non-interpersonal accounts of moral permissibility cannot expect the Wronging distinction to add

normative weight to their accounts of wrongness.²⁶

The above point can be made sharper by considering the following example. Suppose, for example, that Ted is about to kill Alice. As a religious observer, Alice subscribes to divine command theory of permissibility. According to Alice, God's will renders Ted's action wrong. *Prima facie*, Alice and Ted disagree about whether or not Alice is entitled to be in control of her life. However, from Alice's perspective, the root of their disagreement is not about her, but about God, and whether or not He forbids murder. When Alice claims, "You can't kill me!", she is referring not to her own authority, but to God's. Alice's protests, on her view, simply serve to *remind* Ted of God's authority; her protests and demands carry no authority in themselves.

From Alice's perspective, what could the Wronging Distinction add? If there is any importance in Ted recognizing her as a being to whom things are owed, it is, for her view, parasitic upon her cardinal concern that Ted recognize the importance of obeying God's will. Alice may be able to recognize that she is a creature with a certain level of value or moral status—she may view herself as a child of God for example—but her view of herself, as a morally important subject is tempered by her subscription to divine command. Alice herself cannot carry any real authority because she does not play a fundamental role in accounting for the permissibility of the constraint. As such, the claim that Ted has wronged her in particular cannot carry any normative weight.

The above example illustrates that a redundant moral feature cannot play a substantive role in accounting for a duty's normative weight. If instances of wronging simply signal the violation

²⁶ The argument of this section bears clear similarities to Stephen Darwall's claim that morality must be second-personal "all the way down" (59). I obviously agree with the broad meaning of this statement, but as argued in Chapter 2, Darwall's understanding of this claim is suspicious.

of a non-interpersonal value, then wronging itself seems to be nothing more than a reminder of this other value's authority. That is, if standards of permissibility are secured by a non-interpersonal value such as inviolability or welfare, then wronging, at most, refers to the authority of these other values. For any moral theory that is based upon a non-interpersonal value, Alice's demands will only call attention to the authority of something else—whether it is inviolability, well-being, or God. On non-interpersonal views, the capacity to be wronged can, at most, direct our attention towards to the authority of another value, but it cannot, by itself, carry a meaningful sense of authority in its own right. If wronging simply serves to remind us of the authority of non-interpersonal values, then its value seems to be purely instrumental and the idea that wronging adds a dimension of moral gravity falls by the wayside.

In conclusion, I have argued that Kamm does not succeed in her attempt to reconcile the inviolability approach with the Wronging Distinction. My argument has proceeded in two stages. First, I have argued that there is no necessary correlation between a high degree of inviolability and the capacity to be wronged. Second, I have argued that Kamm's supervenience explanation threatens the explanatory viability of the inviolability approach. In particular, I have argued that Kamm's appeal to judgment-sensitive attitudes either renders the capacity to be wronged into an inexplicable feature of morality, or it reveals the concept of inviolability to be redundant. Finally, I have argued that interpersonal authority cannot be extracted from non-interpersonal standards of permissibility. In order for the Wronging Distinction to carry normative sway, the capacity to be wronged must play a formative role in the duties that we have.

Chapter 4

The Authority of Moderate Deontology

All deontologists agree that there are constraints on what may be done in pursuit of the greater good, but not all deontologists agree about whether or not these constraints may sometimes be overridden by considerations of the greater good. On the one hand, Absolutists maintain that constraints may *never* be overridden, no matter what the consequences. On the other, Moderate deontologists claim that when the consequences will be catastrophic, constraints may be violated permissibly. For example, both Moderates and Absolutists agree that, in

Low Stakes, There is a poisonous gas threatening 10 people. If Alice the innocent bystander is killed, her body will release an antitoxin that will dissolve the gas before it reaches the 10. John McClane is nearby and can save the 10 people by killing Alice.

McClane is forbidden from Killing Alice. But, in

High Stakes, There is a poisonous gas threatening 1,000 people. Killing Alice the innocent bystander will release an antitoxin that will dissolve the gas. John McClane is nearby and can save the 1,000 by killing Alice.

Moderates and Absolutists disagree about the existence of a threshold. The former argue that the constraint against killing Alice has reached a limit: the consequences of not killing Alice are so bad that McClane may permissibly kill Alice. Absolutists, in contrast, maintain that it does not matter how many people McClane can save—such considerations can never justify killing Alice.

27

²⁷ Moderate deontologists include Charles Fried (1978), Warren Quinn (1993), John Rawls (1995), Samantha Brennan (1995), Michael Walzer (1997), Michael Moore (2008), Martha Nussbaum (2001), Thomas Nagel (1979), Frances Kamm (2007), and Victor Tadros (2011). Absolutists include G.E.M.

Absolutism—with its *whatever the consequences* zeal—seems to stretch our deontological intuitions beyond plausibility. If Absolutism forbids McClane from saving an entire city, then many would argue that this is a view for die-hard fanatics only. Charges of fanaticism are compounded by the idea that Absolutism is a remnant of religious thinking. Bernard Williams, for example, argues that one can only put their faith in Absolutism if they view it as having “supernatural warrant” (90). Absolutists, or so the suggestion goes, are simply in the fog of a religious hangover. Although most people reject consequentialism, they do not accept deontology *whatever the consequences*.²⁸

Threshold issues notwithstanding, Absolutists and Moderates are united in their view that deontic constraints, when they apply, supply *authoritative* reasons. In *Low Stakes*, for example, both Moderates and Absolutists maintain that the constraint against killing Alice supplies McClane with a particularly stringent All-Things-Considered ought. Constraints exercise a distinctive authority insofar as they supply decisive, non-optional, reasons for action. Despite their obvious differences, the intuitive authority of constraints provides an ecumenical tie between Absolutists and Moderates. The difference being, of course, that on the Moderate view, only sub-threshold constraints remain decisive.²⁹

Moderates maintain that even if there is a threshold this does nothing to change the

Anscombe (1981) and P.T. Geach (1969). Judith Jarvis Thomson (1990) endorses a threshold conception of rights, yet maintains that the right not to be killed is absolute. Others, such as Robert Nozick (1974), are agnostic on the question.

²⁸ There are other questions to be asked about Absolutism. For example, may we permissibly kill Alice if she has created the threat? Does Absolutism extend to property rights, or the right not to be kicked in the shins? These are important questions, but for present purposes I will be focusing on Absolutism as it pertains to the intentional killing of innocent people. It is conceptually possible to be an Absolutist about the right not to be killed, but not about lying, shin kicking, etc. See Judith Jarvis Thomson (1991).

²⁹ W.D. Ross may be an exception to the Authority Claim. In characterizing constraints as *prima facie* duties that cannot be determined in advance, Ross does not present restrictions in distinctly deontological terms. As Saul Smilansky expresses the point, “The price of such a meta-deontological position is that it lacks, anywhere (e.g. even before the threshold) precisely the sort of absolute constraint that has been the attraction of deontology” (2003, 74).

normative force of low-level restrictions. Thomas Nagel, for example, argues that the admission of a threshold does not make constraints “disappear, or change their basic character”; sub-threshold, constraints remain starkly deontological (62-63). Moderates, in other words, maintain that the reality of a threshold does not adulterate the normative force of low-level constraints. For ease of exposition, I will refer to this as the Authority Claim.

Although the authoritative character of deontic restrictions is largely intuitive, the substance of the Authority Claim can, for present purposes, be broken down into two statements. First, the Authority Claim states that the basic normative character of low-level constraints remain categorical in character; they are particularly stringent All-Things-Considered-Oughts. Secondly, the Moderate tends to characterize situations such as *High Stakes* as exceptional moral tragedies. From the Moderate perspective, threshold situations are, in comparison to *Low Stakes*, the more regrettable circumstances.

I argue that Moderate deontologists cannot make sense of the Authority Claim. I argue, *pace* Nagel, that the existence of a threshold changes the basic character of low-level constraints and thereby threatens to make them disappear. The structure of this chapter is as follows. To begin, I outline a form of Absolutism and its explanation of the Authority Claim. Following this I outline two different forms of Moderate Deontology—what I will refer to as the Weighing Model and the Specificationist Model—and their explanations of threshold deontology. Ultimately, I argue that the Weighing Model is the most plausible characterization of Moderate deontology.

Adherence to the Weighing Model, I argue, unravels the Moderate’s commitment to the Authority claim. First, I argue that there are certain cases where Moderate agents may alter the normative force of sub-threshold constraints. Secondly, I argue that the Moderate cannot claim that threshold situations are more regrettable than sub-threshold situations. Moreover, I argue that

from the perspective of an agent attempting to live up to her normative commitments, the Moderate ought to regard *High Stakes* as preferable to *Low Stakes*. In light of these arguments, I argue that if we wish to maintain the intuitive authority of constraints then we must be Absolutists. Authoritative constraints, I argue, must be all or nothing.

Absolute Authority

Absolutists and Moderates agree that the constraint that McClane faces in *Low Stakes* supplies him with an All-Things-Considered ought, but they disagree about how this All-Things-Considered ought comes about. Absolutists understand constraints as supplying agents with protected reasons. Protected reasons are combinations of reasons that give the agent a reason to x *and* a reason not to consider other reasons that count in favor of x-ing or not x-ing. Suppose, for example, that I am going to play a game of tennis and that I need to decide whether or not I am going to play to win, or play just to have fun. In the end, I decide that I will probably have more fun if I play to win. Noticeably, this decision changes my deliberative process. Once I begin to play, certain considerations will not strike me as relevant; spiking the ball, for instance, might hurt my opponent's feelings, but if I am playing to win, I have a reason not to weigh this consideration. That I might hurt my opponent's feelings is still a reason not to spike the ball, but playing to win gives me a reason not to think about this. Protected reasons, in other words, mute the normative force of competing considerations.³⁰

Constraints, for the Absolutist, function in a similar manner to the tennis example. McClane, for example, has both a reason not to burden Alice and a reason that excludes considerations that count in favor of burdening Alice. That is, McClane has a reason not to burden Alice and reason not to consider the fact that doing so would minimize the amount of

³⁰ For more on protected reasons see Raz (1986) and Enoch (2011). The tennis example is borrowed from Scanlon (1998, 51).

burdens overall. This changes the way that McClane arrives at the All-Things-Considered judgment that he ought not to burden Alice. Specifically, McClane *does not weigh* burdening Alice against not burdening Alice; rather, McClane's reason not to burden Alice forms a direct route to the All-Things-Considered judgment by circumventing competing considerations.

The above can be helpfully illustrated with a metaphor borrowed from Robert Nozick. As Nozick explains it, constraints draw "a line (or hyper plane)...in moral space around the individual" (57). On the moral space conception, constraints function as a barrier in the sense that individuals are shielded from consequentialist considerations. Much like the tennis example, the normative force of competing considerations is muted: the shield that protects Alice from being burdened gives McClane both a reason not to burden Alice and a reason not to weigh the competing consequences. On this understanding, if McClane weighs the option of burdening Alice against the option of not burdening, then something has gone wrong with his deliberative process. The barrier that surrounds Alice is supposed to occlude both burdening *and* the weighing of options. As Gopal Sreenivasan explains the point: "...a deontological claim-right is meant precisely to block or prevent this kind of trade off, rather than to depend upon its outcome" (476).

The Weighing Model

There is no received or standard definition of Moderate deontology and, as such, it is difficult to give a latitudinarian outline of the view. Although many deontologists purport to be Moderates, explanations of a threshold are gestured at rather than expounded. As Anthony Ellis observes, "Writers sympathetic to [moderate] deontology often remark that an account will have to be given, but do not go on to give the account" (859). The following explanation of Moderate deontology is meant to offer the basic conceptual structure of the position.

In its simplest form, Moderate deontology attempts to combine a concern for constraints

with a concern for aggregative outcomes. In particular, Moderates attempt to manage these dual concerns by positing a point of severity at which an action that is normally wrong will be right given its consequences. If the threshold is tied to the *severity* of the consequences, it is plausible to interpret the Moderate as weighing the consequences of not killing against the badness of killing.

On this view, constraints supply ordinary reasons. Unlike protected reasons, these reasons are simply weighed against other reasons, and All-Things-Considered Oughts are yielded by what we have, on-balance, most reason to do. When I wake up in the morning, for example, I may need to choose between going to the gym and going back to sleep. There are multiple considerations that count in favor of either option and by weighing them against one another I can determine what I have most reason to do. In a similar vein, constraints can be understood as considerations that are weighed against states of affairs; what we should do, All-Things-Considered, will depend upon the relative weight of the constraint and the weight of competing considerations.³¹ For Moderate deontologists, constraints are simply considerations with a certain amount of weight. They carry enough force that McClane should not kill in *Low Stakes*, but not enough force that they outweigh killing in *High Stakes*. A threshold exists, on this view, simply because constraints are not so important that they outweigh all possible consequences.

When we are weighing reasons, the route to the All-Things-Considered ought is not a direct one. My reason to go to the gym, for example, does not, if I have other competing considerations, supply me with an All-Things-Considered ought all by itself. What I have most reason to do only emerges as a result of the weighing process. This puts the Moderate position in stark contrast to the Absolutist: on this understanding, there would be nothing amiss if McClane

³¹ It may seem that the Moderate is weighing incommensurable values against one another. I will return to this question later on. For more on the Moderate and incommensurability, see Ellis (1992).

weighed the burdening of Alice against the burdening of 10 other people. There would, of course, be something askew if McClane failed to weigh the considerations properly—if he failed to take not killing seriously enough—but there would be nothing wrong with the weighing process itself.

To emphasize, the primary difference between Absolutist and Moderate understandings of constraints is the way in which the reasons connect to an All-Things-Considered judgment. Moderate constraints supply ordinary reasons that simply tip the balance of reasons in cases such as *Low Stakes*, but not in *High Stakes*. Absolutist constraints supply reasons that circumvent the weighing process entirely. For Moderates, constraints are connected to All-Things-Considered judgments simply by having the right weight, but for the Absolutist, constraints occlude these sorts of moral trade-offs entirely. As Nagel explains, from the Absolutist perspective “something has gone wrong when certain measures are admitted into consideration in the first place” (58).

The Moderate’s explanation of the Authority Claim should be obvious: reasons that outweigh other reasons *just are* important reasons, and on-balance reasons *just are* what we have, All-Things-Considered, most reason to do. Furthermore, on the Moderate line of thought, the more it takes to outweigh a constraint, the more important the constraint must be. The stringency of constraints, on this view, can be measured by the amount of good it takes to violate them. As Frances Kamm explains it, “The more good we must sacrifice rather than transgress the right, the more stringent is the right by at least one measure” (249).

The Specificationist Model

Certain Moderate deontologists may be willing to endorse the weighing process. Others may be resistant to my characterization of Moderate Deontology. Instead, they may wish to argue that constraints are not outweighed, but that they *fail to obtain* in threshold situations. Frances Kamm, for example, makes a distinction between a right being overridden and its failure to apply:

“Some moral philosophers support threshold deontology, where this means that when the overall bad consequences of not transgressing someone’s right is great enough, the right is overridden. (An alternative way of understanding the matter is that the right when properly specified does not apply in certain circumstances)” (2011 49).

Kamm does not fully explain the distinction at hand, but she may mean that the content of rights are defined, in advance, by a list of qualifications that enumerate when a constraint does and does not apply.³² On this view, rights are limited in the way that they are specified. For example, the right not to be killed may be delimited in that it does not include a prohibition on killing in cases of self-defense; i.e. the constraint against killing includes an exemption in situations where the victim presents a lethal threat to an innocent agent. The constraint against killing, understood in this way, is not outweighed by competing factors; rather, it does not apply given the particularities of the situation. In this fashion, it may be argued that the restriction on killing the innocent includes a threshold specification. More precisely, the right not to be killed might be specified so that it does not extend to the right-not-to-be-killed-at-the-expense-of-1,000-other-people.

Although the Weighing Model and the Specificationist Model are conceptually different in how they explain defeated constraints, it is not clear that the Specificationist Model can offer a viable explanation of a threshold. To explain, note that on the Specificationist Model, specifications cannot be made without a rationale that explains or justifies the exception. A self-defense clause, for example, might be included on the grounds that victim has forfeited their right not to be killed because they are, in some way, responsible for being a threat.

The question facing the Specificationist Model, then, is whether or not a rationale for the

³² For more on the specificationism of rights see Shafer-Landau (1995) and Oberdiek (2004; 2008).

threshold-specification can be formulated without implicitly relying on the Weighing Model. Pointedly, it is difficult to see how a rationale for a threshold-specification could fail to rely on the *severe weight* of the situation's consequences. If the existence of the threshold *is* predicated upon the severity of the consequences, then it seems that the Specificationist Model just collapses into the Weighing Model: Alice's right is limited because it can be outweighed. If the severity of the consequences has an affect on the constraint against killing Alice, it seems to have one only because it outweighs the constraint.

Proponents of the Specificationist Model may deny that their view relies on the weight of the consequences, but in lieu of a more plausible rationale, a threshold-specification borders on incoherence. If the threshold has nothing to do with the weight or the severity of the consequences, then it is unclear why the specification exists or why it should be thought important. Without a link to the magnitude of the consequences, the idea of a threshold-specification appears conceptually possible but wholly arbitrary. Although there may be nothing formally inconsistent about such a view, it abandons Moderate Deontology to an implausible moral arbitrariness.

It is doubtful, then, that the Specificationist Model can explain the existence of a threshold without becoming another version of the Weighing Model. That said, there is one crucial difference between the Weighing Model and the Specificationist Model that I will return to later on. Specifically, the Specificationist Model trades on the idea that Alice has a right not to be killed *impermissibly* rather than a right not to be killed *per se*. This means that when exceptions apply there is no right being infringed or overridden. Accordingly, on the Specificationist view Alice *does not have the right* not to be killed in *High Stakes* since the circumstances mark the limit of that right. In contrast, on the Weighing Model, Alice continues to

have a right beyond the threshold although it is permissible for McClane to infringe it. This difference—whether Alice retains a right past the threshold or not—bears significance for how we ought to understand the implications of a threshold. I will return to this point later on.³³

The failure of the Specificationist model is instructive. If the appeal of a threshold is thought to stem from the claim that the consequences of not killing would be especially severe, then the Moderate seems committed to the idea that they are weighing constraints against states of affairs. In contrast, if the threshold has nothing to do with the weight of the consequences, then it is difficult to discern why it would exist or why we should find it compelling. I submit, then, that the Weighing Model offers the most plausible way of understanding Moderate deontology.

Tipping the Balance

Moderates make the following claims:

- (1) Deontic constraints, in situations such as *Low Stakes*, hold authority over aggregative considerations by outweighing them.
- (2) Constraints are limited: they may be outweighed, in situations such as *High Stakes*, by a significant increment of harm.

In addition to (1) and (2), Moderates often claim:

- (3) Aggregative considerations only become relevant after a certain threshold is reached.³⁴

Note that the Moderate's answer to the Authority Claim indicates that (3) is misleading. If constraints outweigh consequences in *Low Stakes*, then the consequences may be outweighed but

³³ At issue here is the question of “moral residue” or “moral traces”. Feinberg (1980) and Thomson (1990) argue against the specificationist approach on the grounds that it cannot explain the “moral residue” that lingers after a right has been defeated. Instead, Thomson argues that we should think of defeated rights as being permissibly “infringed” rather than impermissibly violated. For a specificationist response to the infringement/violation distinction see John Oberdiek (2004).

³⁴ Nagel (1979) presents the position this way. Williams (1973) and Walzer (2000) also characterize the idea of a threshold in this way.

they are not irrelevant. Outweighed reasons still hold normative force; they still count in favor of a particular action. My reason to get more sleep might be outweighed by my reason to go to the gym, but it does not simply disappear; I still have a reason to sleep in. Similarly, Moderate constraints, even when they connect to an All-Things-Considered ought, are still being balanced against prospective increments of good. So, instead of (3), Moderates should be understood to claim:

(3*) There are two types of reasons that we ought to be committed to—aggregative and deontological—and we should conform to whichever type of reason holds the most weight in a given situation.

The idea of a threshold suggests that there is a point at which aggregative considerations *become* relevant. Often the threshold line is presented as a sort of moral borderline between the realms of deontology and consequentialism—once you have crossed the threshold, you are in a different realm of value: a place where deontological standards fail to apply and consequentialist standards take over. However, if constraints are simply outweighed by states of affairs, then *the image of a borderline is misleading*: states of affairs do not become relevant, they are always relevant and it just so happens that they are often outweighed. In this sense, McClane's reason to kill Alice does not simply appear above the threshold; rather it is simply of a decisive *weight* above the threshold. The above point has the implication that low-level constraints are acts whose stringency comes in degrees. As the Moderate moves from *Low Stakes* to *High Stakes*, the strength of McClane's reason diminishes.

Moderates are best understood as a particular type of value pluralist who attempts to reconcile consequentialist and deontological moral standards. Moderates are unlike full-blooded consequentialists in thinking that reasons to promote the greater good must be balanced against

deontological reasons, but they do, nonetheless, accept a qualified version of the consequentialist's commitment to securing better aggregate outcomes. In this respect, Moderates are best understood as having two practical commitments: a commitment to promoting the aggregative good and a commitment to respecting rights.³⁵

To clarify, Absolutists need not have any objection to the betterment of the world nor need they deny that there is more than one value in the world. Absolutists do not cease to be deontologists, for example, when they donate money to charity. Nevertheless, the Absolutist's concern for the greater good remains sequestered from their concern for restrictions. To emphasize, for the Absolutist, constraints are always exclusionary, but for the Moderate constraints are one type of reason to be weighed against the prospective value of states-of-affairs. The difference, in other words, is that the Moderate's commitments interact, whereas the Absolutist's commitment to constraints retains a lexical priority. As Saul Smilansky explains the point, Absolutists are "pluralists only when it does not matter" (73).

The weighing process unravels the Moderate's explanation of the Authority Claim. To explain, consider the nature of on-balance reasons. On-balance reasons can be decisive reasons, but they are only decisive insofar as we have reason not to tip the balance of reasons one way or the other. For example, on-balance, I might have more reason to go to the gym than to sleep in, and, in this case, my reason seems to yield an All-Things-Considered-Ought. However, if I rearrange my schedule so that I have time to go to the gym in the evenings, then my reason to go to the gym in the morning is no longer decisive. I can, in other words, alter the balance of reasons in favor of sleeping in. More importantly, I have no independent reason *not* to tip the scales in

³⁵ Saul Smilansky also points out that Moderate deontologists are better thought of as value pluralists: "The proper description of [the Moderate] is that she is a pluralist, plausibly combining concern for consequences with a concern for deontological constraints" (2003, 72).

this way. In fact, from my perspective, it would be *better* if I tipped the balance since it would allow me to conform to as many of my reasons as possible. On-balance reasons can be sufficient, but they are not authoritatively decisive if we are permitted to rework the balance.

Moderate constraints are on-balance reasons and are, as such, susceptible to being reworked. To explain, consider the following non-threshold example,

Gas Distribution, A poisonous gas is travelling towards a village of 1000 people. It is travelling so that when it reaches the town, it will only kill 800 of the villagers. However, McClane can turn on a nearby fan so that the gas will dissipate over the entire population. By dispersing the gas in this way, each person in the city will live a life of chronic pain. Their lives will still be worth living, but their well-being will be greatly reduced.

It is vital that we set the issue of a threshold aside and consider how the non-consequentialist—whether they are Absolutist or Moderate—ought to understand *Gas Distribution*. The case presents a question of fair burden distribution. McClane's choices are to let 800 people die, or to alter the situation so that each person lives a life filled with chronic pain. If we assume that no one in the village is responsible for the gas leak, then it seems fair that, even though more people will be harmed, McClane should impose the lesser burden. When we compare death to a life that is painful (but still worth living) it only seems fair that the burden be distributed more evenly.

Let us assume that both Moderates and Absolutists are committed to resolving such cases using a non-aggregative method. *Gas Distribution* is an aggregation-type case that non-consequentialists ought to be able to make sense of. Most deontologists will accept that there are cases where we can impose a lesser burden on others in order to save lives. This sort of burden

distribution can be justified, according to some non-consequentialists, by pairwise comparison.³⁶

On the pairwise method, individuals are compared one-to-one to determine who has the biggest concern or potential complaint against a proposed action. Each individual on the one side is compared with each individual on the other, in order to determine how badly off each person would be if not aided. Pairwise comparison is most often appealed to in cases involving large numbers. For non-consequentialists, it is way of accommodating the intuition that one should save the larger group without directly appealing to numbers. Additionally, pairwise comparison also supports the claim that there are situations where we can impose a burden on someone else in order to spare another agent a much more severe burden. For instance,

Disease-Inoculation, 800 villagers have contracted a disease that they are certain to die from. The only cure for the disease is a therapeutic vaccine. The vaccine will only work if the entire village is inoculated. Although the vaccine will remove the certainty of death from the 800 of the population, it will give everyone inoculated a mild, non-lethal flu.

Although the vaccine in *Disease-Inoculation* will impose a burden on those who would not otherwise be affected, pairwise comparison can justify the imposition. If we compare certain death to a mild flu, it is better to opt for the latter. Hence, we can, in this way, justify the redistribution of the burden.

Just as pairwise comparison can give us the intuitive upshot in *Disease-Inoculation*, so too can it give us the intuitive answer in *Gas Distribution*. If we compare death to years of chronic pain, pairwise comparison recommends redistributing the burden so that each person suffers the lesser burden. Chronic pain is, of course, worse than a non-lethal flu, but it is still

³⁶ Non-consequentialists who endorse pairwise comparison include Frances Kamm (2007), T.M. Scanlon (1998), Stephen Darwall (2010), and Rahul Kumar (2001).

much better than death. Thus, both the Moderate and the Absolutist can agree that McClane ought to flip the switch in *Gas Distribution*.

That McClane ought to turn the fan on in *Gas Distribution* seems, to me, to be an intuitive conclusion that both Moderates and Absolutists ought to accept. Since no one in the village is responsible for the gas leak, it only seems fair that the entire city share the burden in this way. But, matters can be made more difficult for the Moderate by slightly altering the case:

Rising Stakes, 800 of 1000 villagers face death from the poisonous gas. McClane can turn on a nearby fan, so that the gas will dissipate over the entire village. By dispersing the gas in this way, each person in the village will develop a severe chronic pain condition. Their lives will be worth living, but their well-being will be dramatically decreased. Unlike *Gas Distribution*, however, Alice is nearby and killing her will dissolve the gas completely so that no one is harmed.

How is the Moderate supposed to understand *Rising Stakes*? As I have argued, McClane has reason to more evenly distribute the burden at hand by trading death against chronic pain. This, as I have argued, is a reason that both the Moderate and the Absolutist ought to be able to make sense of. However, although it is fair to flip the switch, it also creates a terrible state of affairs insofar as everyone will lead a life of chronic pain. Their lives, as mentioned, will still be worth living, but only marginally. It seems, to me, that the Moderate should consider this to be awful enough to bring the threshold about.

If flipping the switch will bring about the threshold level, then McClane's action will also render the killing of Alice permissible. Since the situation falls within the Moderate's catastrophe clause, McClane can permissibly kill Alice. The balance of reasons has been tipped: by changing the weight of the consequences, McClane makes the killing of Alice permissible. Since, the

purported consequences of not killing Alice have now reached catastrophic proportions, McClane ought, according to the Moderate Deontologist, to abandon pairwise comparison and aggregate the harm faced by the many.³⁷

Rising Stakes pulls the first aspect of the Authority Claim—the claim that the normative character of sub-threshold constraints remains starkly deontological—into doubt. If it is permissible for McClane to bring about the threshold, then it is permissible for him to alter the normative force of the constraint against killing Alice. If constraints are supposed to be decisive All-Things-Considered-Oughts, then we ought not to be able to alter their normative force. Yet, if constraints are simply reasons to be weighed against other reasons, then it is possible to alter the balance of reasons and diminish the force of the constraint against killing Alice. Sub-threshold constraints cannot be stringent considerations if we are permitted to alter their normative force.

It may be objected that it is clearly impermissible, on deontological grounds, to render the killing of Alice permissible. Arguably, there must be some sort of deontological prohibition on bringing about a threshold situation. It might be argued, for instance, that Alice's killing ought to alter the permissibility of adopting the pairwise method outlined above: if we know that redistributing the burden at hand will bring about the killing of Alice, then McClane ought not to do so.

In response, note that this is a case where McClane has reason—independently of threshold considerations—to fairly redistribute the burden at hand. The aggregate outcome of the

³⁷ *Rising Stakes* may be especially problematic for Frances Kamm given her adherence to the Doctrine of Triple Effect (DTE) (2007, 92-93). The DTE posits a distinction between acting in-order to and acting because-of. In the *Loop* trolley problem for example, Kamm argues that an agent could flip the switch without intending to hit the one; rather the agent could flip the switch *because* it would hit the one. This seems to imply, that McClane could permissibly flip the switch without intending to risk the city; instead, according to the DTE, he would flip the switch *because* it risked the city and *because* it would bring about a permissible means of saving the 10. Put another way, the DTE would make it permissible to exploit a deontic threshold.

redistribution is certainly awful, but considerations of fairness seem to recommend that the burden be distributed in this way. The Absolutist also has reason to flip the switch in *Rising Stakes*—it is the fair way to distribute the prospective burden—but for the Absolutist flipping the switch will not bring about the threshold.

Keeping the above point in mind, it can be noted that if the Moderate forbids McClane from flipping the switch, then they prevent him from performing what would otherwise be a permissible action. This latter move would have the counter-intuitive implication that Moderate deontology can be *more demanding* than Absolutism insofar as McClane would be forbidden from acting on considerations of fairness. This implication, moreover, does not seem to be one that the Moderate can accept insofar as the primary appeal of the position is its *moderate* take on restrictions. Forbidding McClane from flipping the switch would transform Moderate deontology into the sort of fanatical position to which it is supposed to provide an alternative.

Rising Stakes can be understood, then, to present the following dilemma to Moderate deontologists: either sub-threshold constraints lose their intuitive authority or their authority is maintained at the cost of Moderate deontology's common sense appeal. If McClane is forbidden from flipping the switch because it would bring about a threshold, then he is forbidden from fairly distributing the burden at hand. And this, in turn, would diminish Moderate deontology's centrist flavor. We can think of this as the paradox of Moderate deontology.

It may be objected that *Rising Stakes* is only problematic for non-consequentialists who accept pairwise comparison. Additionally, it may also be argued that the threshold is only reached if the entire village faces certain death rather than a lifetime of chronic pain.

These objections are not open to the Moderate. I have used pairwise comparison simply to illustrate the way in which non-consequentialists might make sense of burden distribution

cases using non-aggregative reasoning. Since Moderate deontology is supposed to accommodate our common sense moral intuitions, it needs to find some way of making sense of *Gas Distribution*; whether the Moderate prefers some other non-aggregative method over pairwise comparison is irrelevant. The point of *Rising Stakes* is to illustrate that once the Moderate distributes the burden, they will alter the normative force of the constraint against killing Alice. Furthermore, stipulating that we can only aggregate certain death, rather than years of chronic pain, strikes me as unacceptably ad-hoc. A Moderate deontology that does not weigh many years of severe pain is hardly a *moderate* position. Again, adding such modifications to the Moderate position diminishes its common sense appeal.

Moral Residue and Living up to One's Commitments

The Moderate may be willing to admit that there will be cases where we might alter the force of constraints, but they may object that this does nothing to damage the Authority Claim in cases such as *Low Stakes* where McClane cannot alter the balance of reasons. In sub-threshold situations, the on-balance reason not to burden Alice remains decisive because the scales are not in danger of being tipped. *Rising Stakes*, much like *High Stakes*, is simply an unfortunate situation where aggregative considerations hold sway. This brings us to the second dimension of the Authority Claim: that threshold situations are, in comparison to *Low Stakes*, the more regrettable circumstances.

Even if constraints remain decisive in situations such as *Low Stakes*, the existence of cases like *Rising Stakes* and *High Stakes* imply that lower-level constraints are an unfortunate fact about morality. If we found ourselves, like McClane, bound by a deontic constraint, then the possibility of threshold circumstances would give us reason to think that it would be *better* if more people were being threatened. Constraints might continue to apply, but they would seem to

be an unfortunate imposition: even if we would have reason to respect the constraint, we would also have reason to hope that we could override it by increasing the numbers. This would reveal low stakes constraints not to be arbitrary or irrational, but to be *unfortunate*. Constraints, on the Moderate view, end up looking more like moral obstacles than significant duties.

To explain, consider the Moderate's substantive commitments. As previously argued, the Moderate is best understood as someone who attempts to live up to two commitments: a commitment to promoting aggregate welfare and a commitment to respecting rights. From the perspective of an agent trying their best to comply with all of their commitments, *Rising Stakes* should be regarded as good news since it puts McClane in a position to satisfy both of his goals: he can save as many people as possible and he respects Alice's right not to be killed.

The latter claim may sound strange. If McClane kills Alice, then how does he respect her right not to be killed? If we understand the importance of rights to lie in their weight, then McClane has given Alice's right all of the respect that it can be accorded. In weighing aggregative concerns against rights, McClane seems dedicated to the idea that the properly respecting rights is a matter of taking adequate account of their weight. The fact that the killing is permissible in *High Stakes* means that McClane can say to Alice "I am sorry that I killed you, but you need to realize that I respected your right not to be killed because I took full account of its normative weight. It is not as if I treated your right as being unimportant. You have no legitimate complaint against me."

Now obviously *Rising Stakes* is not unadulterated good news since it will involve the killing of Alice, but it is better news than a case such as a *Low Stakes*. In *Low Stakes*, McClane respects Alice's right not to be killed—he gives it enough weight—but he fails to promote the greater good. In contrast, *Rising Stakes* seems preferable since McClane permissibly promotes the

greater good while respecting Alice's right. McClane, in both *Low Stakes* and *High Stakes* respects Alice's right not to be killed, but it is only in *High Stakes* that he also fulfills his consequentialist commitment. Complying as best as one can with one's commitments, moreover, is just part of the nature of practical reason; we ought, insofar as we are able, to comply with as many of the commitments that we have. Normatively speaking, *High Stakes* ought to be viewed as preferable to *Low Stakes*.

It would be entirely plausible for McClane—given his commitment to aggregate welfare—to feel disappointed if he found himself in *Low Stakes* rather than *High Stakes* or *Rising Stakes*. Suppose for example that McClane thought that he was in *High Stakes*, but after recounting found that he was in *Low Stakes*. From McClane's perspective, this should seem *unlucky* insofar as he would not be able to live up to his own standards. The existence of a threshold would not necessarily make the low-level constraint against killing Alice disappear but it would fundamentally alter its basic normative character.

Here is another way to put the point. Moderate deontologists often present cases such as *High Stakes* as exceptional moral tragedies, where it is unfortunate but necessary, to act on consequentialist reasons. What I am arguing, is that given the Moderate's general commitment to promoting the greater good, it is difficult to see how the label "unfortunate" is used in anything but an ad-hoc way. Given that the Moderate deontologist always has a reason to kill Alice in *Low Stakes*—an outweighed reason is still a reason—it would be equally consistent for them to view *Low Stakes* as the truly tragic scenario—not tragic in the sense that people will die, but tragic in the sense that it is not permissible to minimize harm. The Moderate's commitment to aggregative outcomes invites the thought that deontic constraints are simply considerations that make it more difficult to be a full-blooded consequentialist. Moderates may view constraints as having a certain

amount of authority—constraints may outweigh aggregate outcomes in most situations—but, given the full set of normative commitments to which they subscribe, they also have reason to hope that this authority will be overturned.

The Moderate may object to the above line of argument by relying on a distinction, drawn by Judith Jarvis Thomson, between the infringement and the violation of a right. On this view, rights always apply even when it is permissible to violate them; rights that are outweighed are *infringed* rather than violated (84; 70). The infringing/violating distinction is often thought to capture the idea that in cases of permissible rights infringements, something of moral value has been lost, leaving behind a form of “moral residue”. The intuition underlying the moral residue claim is that acting permissibly may still leave something to be desired. What is left to be desired, in cases such as *High Stakes*, may very well be the infringement of a right.

Building on the infringing/violating distinction, the Moderate may argue that although McClane may permissibly kill Alice, he nevertheless *infringes* her right not to be killed. In this sense, crossing the threshold may mean that the circumstances have rendered killing permissible, but it does not change the fact that Alice has a right not to be killed. Understood in this way, the reason not to kill Alice does not disappear, it simply loses its special status as decisive. Accordingly, it can be argued that, even in cases such as *Rising Stakes*, there is a distinct deontological loss.³⁸ What is unfortunate about threshold situations, on this line of thought, is that

³⁸ Though he does not use the term, W.D. Ross (1930) tends to characterize threshold situations as leaving traces of moral residue. Michael Walzer (1983) also characterizes threshold situations in this way, arguing that although we should not be utilitarians we must, unfortunately, get our hands dirty—the dirt being, in such cases, deontological residue.

one is required to do wrong. Thus, it may be objected that it is not open to McClane to regard *Rising Stakes* as good news.³⁹

Understood in this way, the Moderate's characterization of a threshold situation is a tragic dilemma where no matter which choice we make, there will be some sort of loss that we cannot avoid. If I must choose, for example, between joining the French resistance and taking care of my ailing mother, then no matter which choice I make, I will leave one of two values unrealized.⁴⁰ In a similar manner, the Moderate might be understood to claim that even though it is permissible to side with consequentialist considerations, their commitment to deontological standards remains unrealized. Hence, although McClane kills Alice permissibly, he leaves a trace of deontological residue.

It is not clear, however, that this line of response is open to the Moderate. What makes the choice in the above example a tragic one is that there is a loss in one value that cannot be made up for by a gain in the other. If I decide to take care of my ailing mother, then I cannot make up for the loss of not joining the resistance because the values in question do not admit of a common measure. To put the point another way, if there is a sense of "French resistance residue" then this is because the two values are incommensurable.

Noting the above point, we can see that the Moderate's characterization of threshold situations is different from the above example in two respects. First, if the choice between the French resistance and my mother is a choice between incommensurable values, then there is no determinate answer as to what I should do. And this, obviously, is a very different sort of claim

³⁹ It is not clear that the idea of moral residue really supports the infringement/violation distinction. As John Oberdiek (2004) argues, the fact of residue does not, in itself, support the infringement/violation distinction, since the sense of residue in such cases can be non-deontological. In *High Stakes*, for example, even a consequentialist can note that there is a sense of residue—"something left to be desired"—insofar as there is the loss of life.

⁴⁰ The example belongs to Jean-Paul Sartre (1975, 295-296).

from the one that the Moderate makes. The Moderate claims that above the threshold, killing Alice becomes *the right thing to do* and this is not the same as claiming that there is no right answer. Secondly, we can note that the Weighing Model commits the Moderate to the idea that the values in question are commensurable. In weighing the disvalue of killing Alice against the prospective consequences of not killing Alice, the Moderate assumes that these values can be compared to one another. As Anthony Ellis explains, “For [Moderates], the considerations must be commensurable, for the claim that a normally prohibited act is in certain circumstances right is supposed to be *justified* by a reference to its consequences” (863).⁴¹

A commitment to commensurability casts considerable doubt on the Moderate’s claim to find deontological residue in threshold cases. If the two values are, in fact, commensurable, then we should think that the gain in consequential value makes up for the loss of a deontological value. Suppose, for instance, that I am weighing options A and B against one another in terms of how much fun they are. If I choose option B, because it is more fun than option A then there is no option A residue: choosing option B makes up for the fun that would be gained by option A.⁴²

The above claim about value incommensurability and value residue may be controversial to some, so I will not rest my argument on it.⁴³ Instead, to refocus, consider what the idea of deontological residue might mean for my argument about McClane’s practical outlook. The Moderate might argue that McClane’s commitments are different from the way that I have

⁴¹ Anthony Ellis (2007) argues that because the values are incommensurable, the point at which the threshold is posited must be arbitrary.

⁴² The significance of deontological residue can be questioned in other ways. If we accept the idea of deontological residue in *High Stakes*, then it is not clear why we should not also accept the idea that there is consequentialist residue in *Low Stakes*. And, if we accept this, then the Moderate’s claim that *High Stakes* is worse than *Low Stakes* appears to be entirely ad-hoc. Both would be unfortunate situations—insofar as neither commitment is fully lived up to—but neither situation would be worse than the other.

⁴³ It might be argued that the values could be commensurable and still leave a trace of residue. Williams (1981) seems to think this. I am not sure how to understand this claim. For more on this issue see Raz (1986; 1999), Anderson (1993), and Wiggins (1997).

described them. Specifically, the Moderate might argue that the Moderate has a commitment to promoting welfare and a commitment to not *infringing* rights. On this line of argument, the fact that there is a trace of moral residue indicates that McClane has failed to live up to his deontological commitment.

It is not obvious, however, that it is open to the Moderate to characterize McClane's practical commitments in this way. Even if the infringement of a right leaves deontological residue, it is not a sense of residue that can structure McClane's practical commitments. A commitment to not permissibly infringing rights is tantamount to a commitment to not doing, or to avoiding, the permissible. And this, it seems to me, cannot be coherently thought of as a practical commitment: we cannot be committed to avoiding the *permissible* promotion of the good. Instead, McClane can only be committed to avoiding the *impermissible* violation of rights. Thus, accepting the moral residue claim does not change the structure of McClane's practical commitments. Although McClane may infringe Alice's right, it is a permissible infringement that does not carry any normative significance.

These considerations put us in a better position to appreciate the significance of the Absolutist's authority explanation. As previously argued, the Absolutist views McClane as having a protected reason: a reason not to kill Alice and a reason not to consider consequentialist considerations. Accordingly, for the Absolutist, there is no sense in which McClane can tip the balance of reasons one way or the other. If constraints mute the normative force of states of affairs, then it they cannot be re-introduced at any level. It is only by excluding consequentialist reasoning altogether that constraints remain decisive in all situations.

It may be objected that protected reasons can be outweighed by the value that they exclude. For example, suppose that in the tennis example, I play to win and this gives me reason

to ignore my opponent's feelings when I am in a position to spike the ball. However, it does seem that playing to win can be outweighed by the fact that my opponent is in tears—this would give me reason to abandon the idea of playing to win. Here, the protected reason would be outweighed by the considerations that it excludes. As such, the Moderate deontologist may argue that constraints can exclude states of affairs from the deliberative process up to a certain point. Reaching a deontological threshold is akin to finding your tennis opponent in tears: once the competing considerations are strong enough, they cannot be excluded. If this is the case, then *Rising Stakes* will not be a problem.

This will not work. In the tennis example, the protected reason is only outweighed because the rationale for treating it as protected has been defeated. If I adopt the aim of playing to win because I believe it will lead me to have more fun in the long run, then my opponent's tears will defeat this rationale. Playing to win—in order to have more fun—becomes futile, because it is not fun to play against a sobbing tennis partner. The rationale can be defeated by considerations of having fun, *only because* the rationale itself is rooted in fun-maximizing considerations. When I play to win for the sake of maximizing fun, I presuppose the importance of having fun in such a way that fun-time considerations can re-enter my deliberative process. Thus protected reason can be outweighed, but only because the rationale for adopting the aim has been defeated.

The problem, for the Moderate, is whether or not they can tell a similar story *without* presupposing the authority of consequentialist standards. As the tennis example demonstrates, excluded considerations can re-enter the deliberative process—but only if the rationale presupposes the importance of having fun. Any rationale for deontic constraints that permits exceptions in favor of states of affairs will have to presuppose the importance of states of affairs

and, once this presupposition is made, it is difficult to explain why the importance of states of affairs only enter the deliberative process in *High Stakes* but not in *Low Stakes*.

To elucidate, consider how a simplified form of rule-consequentialism might explain the existence of a threshold. If the rationale for the constraint is that “the best way to minimize killings overall is to minimize one’s own killings”, then it is easy to locate a threshold: once the rule fails to be optimal, it can be abandoned. However, although the rule-consequentialist may have reason to respect rule-based constraints, they also have reason to hope that the rule can be broken. If we should respect constraints because doing so minimizes overall killings, then the rationale for constraints is in deep tension with the prohibition: it would be better, according to the rule-consequentialist’s own standards, if we could minimize killings even further by living without constraints.

The Moderate deontologist is in the same position as the simplified rule-consequentialist. If the Moderate deontologist formulates a rationale for deontic constraints that presupposes the importance of consequentialist standards, then there will always be a reason to hope that the balance of reasons tips in favor of consequentialism.

Meeting the Authority Challenge

The Absolutist—with their absolute exclusion of states of affairs—has the basic framework to make the authority of constraints intelligible at all levels. The basic conceptual structure of protected reasons provides the blueprint for Absolutism and the Authority Claim: constraints, for the Absolutist, remain decisively connected to an All-Things-Considered ought, because the normative force of competing considerations has been muted. This highlights the overlooked significance of the Absolutist position: it is only by occluding states of affairs entirely that they never gain traction.

This does not, of course, let the Absolutist off the hook. While I have outlined the basic structure of an Absolutist framework in terms of protected reasons, I have not detailed the source of authority that grants constraints their exclusionary character. At an abstract level, it is not difficult to understand the basic functioning of constraints in exclusionary terms, but to fully explain the Authority Claim, it needs to be shown why constraints have the exclusionary character that they do. Furthermore, a plausible authority explanation ought to diminish charges of fanaticism.

Certain Absolutists seem to carry out this latter task by appealing to an ethical framework that we no longer have access to. Anscombe's defense of Absolutism in "War and Murder", for example, presupposes a Catholic background. Absolutism, according to Anscombe, makes sense from a Catholic perspective, because Catholics must trust that God is responsible for all consequences while knowing that they can only do their best within the limits of obedience (61).⁴⁴ If Anscombe's view is the best that the Absolutist can do, then it may be the case that Absolutism is not a real option for us given the difficulty, and perhaps the undesirability, of resurrecting a divine authority. From what has been argued, it would be entirely plausible to conclude that Moderate deontologists ought to give up on deontology and embrace consequentialism.

It should be noted that the arguments of this chapter are not meant as a full defense of Absolutism. My aim has been more modest. In particular, by outlining the problem of a shifting authority, I hope to have shown that the authoritative nature of constraints can only be maintained within an Absolutist framework. In light of this, the problem of authority should be understood as a *challenge* to Absolutism rather than its refutation. While the resurrection of a divine authority

⁴⁴ Geach (1969) makes a similar argument for absolutism.

may be off the table, formulating a secular account of moral authority may be within the grasp of modern, secular non-consequentialists.⁴⁵ Restoring Absolutism remains a possibility for those reluctant to accept consequentialism.

Whether or not the task of rehabilitating Absolutism is worthwhile will depend, in part, on how strong one finds the pull of the Authority Claim. For my own part, I would submit that, given what is at stake, the challenge is worth trying to meet. As Nagel might put the point, if the authority of low-level constraints cannot be maintained, then we are only one step away from losing “the only barrier before the abyss of utilitarian apologetics for large-scale murder” (56).

In conclusion, I have argued that Moderate deontologists lose their grip on the Authority Claim in light of the fact that they are committed to the weighing and balancing of consequentialist and non-consequentialist reasons. The problem, for the Moderate, is that the weighing process does not rule out problem cases where we are permitted to tip the balance of reasons in favor of consequentialism. The possibility of tipping the balance, I have argued, dissolves the Moderate’s explanation of the Authority Claim. From the Moderate’s perspective, low-level constraints are not authoritative considerations, but unfortunate impositions that prevent us from acting on consequentialist standards. Moreover, once deontologists lose confidence in the authoritative character of restrictions, it is unclear why they should not also lose confidence in their existence.

Faced with the choice between Moderate deontology and accusations of religious zealotry, non-consequentialists have tended to opt for the Moderate view. However, by outlining the authority problem plaguing Moderate deontology, I sought to have shown the overlooked

⁴⁵ Although they do not discuss the idea of Absolutism specifically, non-consequentialists such as Stephen Darwall (2009) and T.M Scanlon (1998) both attempt to articulate the authority of morality in interpersonal terms. Whether or not an interpersonal sense of authority makes Absolutism a more plausible position remains an open question.

appeal of Absolutism. By offering a secular account of deontological authority, non-consequentialists may be able to maintain the Authority Claim and avoid charges of religious fanaticism. This latter task was beyond the scope of this paper, but the problem of an evasive authority gives us reason to investigate further.

Chapter 5

Threats, Fairness, and the Moral Relevance of Causal Placement

Consider the following:

Threat: Jake-the-threat is falling, through no fault of his own, towards Finn. If Jake-the-threat lands on Finn, Jake-the-threat will be unharmed and Finn will be killed. Finn can pull out his ray gun and evaporate Jake-the-threat before he is hit.

Bystander: A villain is shooting at Finn. Finn cannot escape the shots, but he can block them by grabbing Jake-the-bystander and using his body as a shield.

Many non-consequentialists believe that it is permissible to kill in *Threat* but not in *Bystander*. Michael Otsuka (1994) and Jeff McMahan (2009) have (separately) challenged this view by arguing that although we can be held responsible for killings that flow from our agency, we cannot be liable to harm because of lethal bodily movements over which we have no control. Both bystanders and threats, according to McMahan and Otsuka, lack any sort of responsibility that would render them liable to harm. Accordingly, McMahan and Otsuka argue that we should accept:

The Moral-Equivalence-Thesis: The killing of a threat and the killing of a bystander are, other things equal, on par as far as permissibility is concerned. (Otsuka, 76; McMahan, 170) ⁴⁶

⁴⁶ Henceforth, please assume that all bystanders and threats are “innocent” unless otherwise noted.

There is, however, a widely held intuition that it is worse to kill a bystander than it is to kill a threat. Building on the work of Warren Quinn (1993), Jonathan Quong (2009) argues that the intuitive difference between killing threats and bystanders turns on a distinction between opportunistic and eliminative agency. When we kill bystanders we use them as mere survival tools, but when we kill threats we simply eliminate them. Killing a threat, or so the argument goes, involves a less pernicious form of agency than the killing of a bystander. According to Quong, the distinction between opportunistic and eliminative agency reveals that there are two different ways of killing someone as a means to save your own life. For ease of exposition, I will refer to this as,

The Means Distinction: When we use someone to perform an action that we could not perform without them—i.e. when their presence is casually necessary to the action being performed—we use them as a mere means.

I argue that although the Means Distinction seems *prima facie* plausible it is not a sufficient explanation for the permissibility of killing threats. My argument against the Means Distinction is two-fold. Most non-consequentialists, including McMahan and Otsuka, accept that the Means Distinction carries some moral significance, but I argue that this is a mistake: we do not have any reason to believe that opportunistic killings are, in general, worse than eliminative killings. Following this, I argue that even if we accept the Means Distinction, there are threat-type scenarios in which there is no intuitive difference between killing a threat opportunistically and killing a threat eliminatively.

In the second half of this chapter, I argue that although we should accept the Moral-Equivalence-Thesis, this should not lead us to overlook significant asymmetries between parties in a threat-type situation and ordinary bystanders. Threat-type cases ought to be understood as

posing questions of fair burden distribution and that one's place in the causal architecture can bestow one with especially cumbersome duties of aid. Specifically, I argue that, unlike ordinary bystanders, both parties in a threat-type situation—both threats and the threatened—are subject to:

The-Principle-of-Cost-Mitigation: we are required to bear heavy burdens when we are either (A) sustaining a lethal sequence of events, or (B) when we would, by saving ourselves, cause a lethal side-effect. This is true even when we are not responsible for the fact that we are in a threat-type situation.

The-Principle-of-Cost-Mitigation, I argue, is supported by considerations of fairness. Generally, when two agents face a burden—for which neither of them are responsible—they ought, if they are able, to mitigate the costs of the situation as best as they can. Additionally, when the costs of the situation cannot be mitigated—when only one agent can escape the threat—the beneficiary ought to be determined by some sort of fair outcome procedure. In this vein, I argue that a prohibition on killing threats is consistent with matters of fairness.

The Moral-Equivalence-Thesis

The argument for the Moral-Equivalence-Thesis begins by assuming that it is impermissible to kill bystanders. According to McMahan and Otsuka, it is impermissible to kill bystanders because they have not done anything to forfeit their rights not to be killed; bystanders are off limits because they are not responsible for the relevant threat. Therefore, they have not done anything to forfeit their right not to be killed. Following this line of argument, someone's

liability to harm is contingent upon whether or not they are responsible for the relevant threat. Moral responsibility, on McMahan and Otsuka's view, is a necessary condition for liability.⁴⁷

McMahan and Otsuka argue that if we accept this claim—that responsibility is a necessary condition for liability—then we ought to accept the Moral-Equivalence-Thesis. Neither threats nor bystanders bear any responsibility for the situation at hand; thus, they cannot be liable to harm. Of course, it is true that in *Threat*, it is Jake-the-threat's body that constitutes the threat, but according to McMahan and Otsuka, this does not make a moral difference. Merely constituting a threat does not entail that one has forfeited their right not to be killed. In Otsuka's words:

It is natural to think that one bears special moral responsibility for such killings that flow from one's agency. No such special responsibility, however, plausibly attaches to lethal movements over which one has no control, as in the case of a Threat. (79)

Given Jake-the-threat's lack of responsibility, it cannot be the case that he is liable to be killed. While it is true that threats endanger in virtue of their positioning within the local causal architecture, this fact does not—according to McMahan and Otsuka—carry any moral significance. Since it is impermissible to kill a bystander, and since there is nothing morally relevant that distinguishes threats from bystanders, it is also impermissible to kill threats.

The Means Distinction

⁴⁷ It is worth noting that although responsibility may be a necessary condition for liability, it is a sufficient but not a necessary condition for permissible harming. In the standard *Trolley* case, for instance, our pulling the trolley switch may permissibly harm the people on the right side of the track, but they are not liable (they have not lost their rights) to harm. Additionally, it should be noted that moral responsibility is distinct from culpability (blameworthiness). Someone may be morally responsible without being blameworthy. I may, for example, be epistemically justified (but mistaken) in thinking that you are a dangerous serial killer. Because I am epistemically justified, I am not blameworthy if I attack you, but I am, nonetheless morally responsible for the fact that I pose an unjust threat.

Recently, building on the work of Warren Quinn, Jonathan Quong has challenged the Moral-Equivalence-Thesis by noting the distinction between *opportunistic* agency and *eliminative* agency (Quinn 344).⁴⁸ To explain, consider the different modes of agency at work in *Bystander* and *Threat*. In *Bystander*, Finn grabs Jake-the-bystander and uses him to accomplish a goal that he could not accomplish if Jake were not there. In this sense, Finn uses Jake-the-bystander as a survival tool insofar as he *exploits* Jake's presence for his own benefit; Finn *profits* from Jake's presence in the sense that he is made better off by Jake's being there. In contrast, in *Threat*, Finn kills Jake-the-threat *eliminatively*: he merely reacts to Jake's presence by removing him. In eliminating Jake-the-threat, Finn neither exploits nor profits from the Jake-the-threat's presence: if Jake were not there Finn would not be under threat in the first place.

Quong argues that the distinction between opportunistic and eliminative agency corresponds to a distinction between killing someone as a means to survival and killing someone as a *mere means* to survival. Under opportunistic agency you view your victim as a mere means to an end because you are using them as a tool in order to do something that you could not do without them. In contrast, under eliminative agency you does not *use* your victim as a mere means; rather you simply remove the threat that they present.⁴⁹

Quong makes the Means Distinction more precise by noting that in opportunistic agency the victim's presence is causally necessary to the agent defending himself:

⁴⁸ Helen Frowe (2009) gives an argument that is roughly similar to Quong's argument. To avoid muddying the waters I am concentrating on Quong's particular argument.

⁴⁹ Quong's argument for self-defense also rests on the claim that the right to self-defense is grounded in the agent-relative value that one's life has for one's self; i.e. it is because my life is especially important to me that I am allowed to shoot an innocent threat. For present purposes, I am concentrating on the Means Distinction because it forms the crux of Quong's argument: even if we accept that the agent-relative value of my life gives me a permission to ignore people in need, it is not obvious that it gives me a liberty right to kill innocent people. As Quong himself notes, "Even if our life has special agent-relative value, the permission to kill innocent people in self-defense seems stronger and more controversial than the permission to let innocent people die if rescuing them would be too risky"(523).

X uses Y, in the special sense that I am interested in here, when Y's presence is causally necessary for the successful execution of X's action. Put counterfactually, without Y's presence, X cannot perform the required action or achieve the relevant goal. Y is, we could say, a tool, which X needs to successfully complete some action or attain some goal. (525)

Of course, threats are also causally necessary for self-defense in that they make the action possible. In *Threat*, for example, Finn cannot defend himself if Jake-the-threat is not there. However, to clarify, the relevant counterfactual is not whether the agent's presence makes self-defense an eligible goal but whether their presence makes the *already* eligible goal of self-defense possible. In *Bystander*, Finn has the goal of self-defense whether Jake-the-bystander is there or not; his presence, should Finn choose to use him as a shield, is something welcome and convenient. This is not the case in *Threat* because without Jake-the-threat's presence there is no threat to be eliminated.

Why should we think the Means Distinction important? The attraction of the Means Distinction lies in the suggestion that there is something dubious about viewing your victim as a mere pawn to be commandeered in the pursuit of your survival goals. Quong finds the distinction important because it "captures part of the Kantian idea that there is something particularly wrong about treating another person as a mere means to your own ends" (525). Similarly, Quinn argues that opportunistic agency involves a "kind of disrespect" that is incompatible with a view of human agents as separate persons (190-193). Under opportunistic agency we make the additional presumption that our victim is there to serve *our own* goals and it is this additional presumption that marks opportunistic agency as especially disrespectful. In contrast, when we engage in eliminative agency we do not need to view our victim as tool and this seems to make it easier to

justify than opportunistic agency. As Quong explains the point, “An innocent threat cannot claim, “You were just using me” if they are killed in self-defense because, in such cases, there is no sense in which the agent “takes advantage of, or exploits the presence of” the threat for their own gain” (526). Finn does, of course, kill Jake-the-threat as a means to survival—killing him enables Finn to survive—but his mode of agency does not require him to make the *additional* presumption that Jake is there to be used for his benefit.

The Means Distinction seems to reveal an important discrepancy between killing bystanders and threats and thereby demonstrates that the killing of threats and bystanders is not obviously morally equivalent. If we accept the Means Distinction, then the burden of argument seems to lie with proponents of the Moral-Equivalence-Thesis.⁵⁰

I am not persuaded that the burden of argument has been shifted, because I find it implausible that opportunistic killing is actually worse than eliminative killing. It is worth noting that eliminative agency, at least as it pertains to *Threat*, should not be confused with killing someone as a side-effect. Eliminative agency is a form of *intentional* killing and this fact, I think, should make us suspicious of the claim that opportunistic agency is worse than eliminative

⁵⁰ The Means Distinction has been criticized as being too permissive. It is not obvious, for instance, that eliminative agency is always permissible. As McMahan points out, if we are running away from a maniac and we need to cross a bridge that is obstructed by a bystander, it is obviously impermissible to eliminate the bystander by pushing him off of the bridge (171). This is a case of eliminative agency, but it is obviously impermissible. Quong argues that, contrary to first appearances, killing obstructers is actually an instance of opportunistic killing. Obstructers, according to Quong, have a claim to the physical space they are in and, by moving them you are taking advantage of their property claim. In Quong’s words: “Since your survival depends on things over which X has a rightful claim, to take them and thereby kill X would be to shift the harm of death from yourself onto X by using X’s entitlements against them. This, I believe, would be to use X as a mere means to your own survival, and as such it should be deemed impermissible” (530). Furthermore, “We treat someone as a mere means to our own survival if we kill them when we could not survive without them or something else to which they have a rightful claim” (532-533). I doubt that rights of self-ownership extend to property rights, but for present purposes I will set this problem aside. For a separate critique of this aspect of Quong’s argument, see Jason Hanna’s “The Moral Status of Nonresponsible Threats” *Journal of Applied Philosophy*, Vol. 29, No.1, 2012.

agency. Racists, for example, may not view minorities as useful tools, but as problems to be exterminated. As Derek Parfit notes, “Though Hitler treated the Slavs in his conquered Eastern territories as a mere means, that is not how he treated the Jews” (2011, 228). Unpacking Parfit’s observation, we can note that the fact that Hitler saw the Jews as things to be eliminated rather than as useful tools does not seem to make a difference as to how we think about the moral badness of concentration camps. Although the Means Distinction has an attractive Kantian resonance—and there does seem to be something bad about viewing people in this way—it is hard to believe that viewing people opportunistically is somehow worse than seeing people as mere things to be eliminated.

Viewing someone as a mere tool disrespects his or her moral status, but viewing someone as a mere eliminable seems no less disrespectful. Consider the claim that opportunistic agency involves an especially dubious presumption about seeing others as subservient to your own goals. Under eliminative agency we do not assume that others are useful tools that aid us in our own goals, but *we do assume* that our own goals outweigh their claim not to be killed. The charge “you were just using me!” has a certain moral force to it, but so too, I think, does the accusation “you just eliminated me as though I didn’t even matter!” Opportunistic agency seems reprehensible, in part, because it is a failure of recognition: when we use people as tools, we deny their status as independent, self-governing, agents. That said, if eliminative agency requires us to view people as things to be removed, then it also belies a failure of recognition: when we eliminate people we destroy their ability to set goals for themselves and we treat them in the way that we would treat a rock or a sack of flour. If eliminative agency can be as disrespectful as opportunistic agency, we need to know why it is not disrespectful in threat-type cases.

The Irrelevance of the Means Distinction

In what follows, I will assume for the sake of argument, that it is permissible to kill falling threats. The Means Distinction posits that the killing of threats is permissible, because such killings do not contain an opportunistic aspect. If this is true, then inserting an opportunistic aspect into a threat-type scenario should reverse the permissibility of killing the threat. It is not clear, however, that this is so. Consider the following:

Double-Threat, Jake is falling through no fault of his own, towards Finn. A rock is also falling, at a different speed, towards Finn, such that Jake will hit Finn at T1 and the rock at T2. Jake and the rock each constitute a lethal threat. Finn has two guns. He has a ray gun, but the ray gun is not powerful enough to destroy the rock. He also has a “momentous-force” gun; this gun will cause whatever it is aimed at to move at an extremely high velocity. If Finn uses his momentous-force gun, he will drive Jake’s body into the rock at such a speed that the rock will explode into harmless bits. Using the momentous-force gun will kill but not evaporate Jake. After destroying the rock, Finn can use his ray gun to evaporate Jake’s corpse. If Finn does not use the momentous force gun, Jake will survive the fall and evade the rock.

Recall Quong’s counterfactual test: X uses Y opportunistically when it is the case that without Y’s presence, X cannot perform the required action or achieve the relevant goal. Finn’s action, in *Double-Threat* fits this description: if Jake were not there it would not be possible for Finn to save himself by destroying the rock. Thus, according to the Means Distinction, killing in *Double-Threat* should be impermissible: if Finn uses the momentous force gun, he uses Jake *opportunistically* in order to eliminate the rock. It is important to note that Finn cannot see the elimination of the rock as a mere side-effect of his action; if Finn does not break the rock, the goal of self-defense cannot be met and it is, in this way, a necessary component of Finn’s action.

The Means Distinction, it seems to me, posits a difference where there is none; I see no significant difference between *Threat* and *Double-Threat*. Killing threats opportunistically does not seem to be any different from killing threats eliminatively. Suppose, for instance, that Finn had a ray gun powerful enough to break the rock. In this case, proponents of the Means Distinction would deem it permissible for Finn to separately evaporate Jake and the rock. But it is difficult to believe that whether or not one has a ray gun or a momentous-force gun can affect the permissibility of self-defense. Pointedly, the momentous force gun does exactly what the ray gun would do: it kills Jake. The only difference between the two guns is that the momentous force gun kills Jake by producing a collision between his body and the rock, while the ray gun evaporates him. Assuming that both deaths are equally painful, it is hard to believe that Jake should object to being killed in *Double-Threat* but not in *Threat*. Whether or not the killing of Jake brings about an additional benefit of destroying a rock seems irrelevant to the question of permissibility. Whatever normative significance the Means Distinction has seems to be impotent in this case.

Proponents of the Means Distinction might object to this case on the following grounds: in *Double-Threat*, there is an opportunistic aspect to the killing, but the killing is *not merely* opportunistic. When Finn kills Jake opportunistically, he does so only because he will also eliminate his body. As such, it might be argued that although Finn treats Jake as a means he does not treat him as a *mere* means.

This response is too hasty. It is not clear what significance turns on the distinction between using someone opportunistically and purely opportunistically. Suppose, for example, that I use a bystander as a shield, but I do so in such a way that I mitigate her damage but, in doing so, sustain more self-damage. In this case, I obviously treat my human shield as a means, but I do not treat her as a mere means because I show some concern for her well-being. If it is

wrong to use people opportunistically, I doubt that the opportunistic element must be unadulterated.

It might be argued that killing in *Double-Threat* is permissible on grounds of Secondary Permissibility. Frances Kamm argues that, in certain cases, an otherwise impermissible act can be rendered permissible. Kamm explains the principle as follows:

...it is impermissible to push an innocent bystander into a trolley that will crush his leg in order to save five other people from the trolleys killing them. However, suppose the alternative is to redirect the trolley away from the five and toward that very same person, thereby killing him. Redirecting it toward him would ordinarily be permissible and is, suppose, something we would do if we could do nothing else. However, since it is in his interest to have his leg crushed rather than to be killed, secondarily it becomes permissible, I think, to push him into the trolley, an action that was not, in the first instance, permissible. (26)

On this line of thought, the permissibility of self-defense in *Double-Threat* might stem from the fact that it would be permissible, counterfactually, to kill Jake eliminatively. In this sense, it might be argued that because it is permissible for Finn to kill Jake in an eliminative sense, it becomes permissible for Finn to kill Jake in an opportunistic sense. By invoking the Principle of Secondary Permissibility, the permissibility of *Double Threat* can be explained as an exception to the Means Distinction.

In response, it can be noted that secondary permissibility applies when the secondary act *would be better for the victim*. In Kamm's words "...what was originally a greater moral evil can become permissible secondarily, *because it involves less harm to the same people*" (170). In the

trolley case that Kamm considers, for instance, it is in the interests of the victim to suffer broken legs rather than death and this is why it is secondarily permissible to break his legs. The principle of Secondary Permissibility only applies when the secondary act would, in some way, be better for the victim than the first-order permissible act. To put the point another way, we cannot exchange a first-order permissible act with an act that is worse for the victim. But if this is the case, proponents of the Means Distinction cannot appeal to Secondary Permissibility in order to explain the symmetry between *Threat* and *Double-Threat*: there is nothing better, for Jake, about being killed opportunistically rather than eliminatively. Of course, there does not seem to be anything worse about being killed opportunistically in *Double-Threat*, but to admit this would be to give up the Means Distinction.

What makes the killing of a threat permissible, according to the Means Distinction, is the absence of an opportunistic element. The problem, however, is that *Double-Threat* does not seem to be different, in any important respect, from *Threat*. Killing a threat opportunistically seems no different from killing a threat eliminatively. I would submit, then, that the Means Distinction does not explain the intuitive difference between killing threats and bystanders.

Causal Placement and Fairness

In this section I will proceed to argue that though McMahan and Otsuka may be correct to claim that we may not intentionally kill non-responsible threats, we should not accept the claim that threats and bystanders are morally equivalent across the board. Despite what Otsuka and McMahan claim, it seems enormously plausible that one's position within the causal architecture can make a difference to the burdens that one can be expected to bear. Compare the following:

Options 1: Jake-the-threat is falling, through no fault of his own, towards Finn. But, Jake can twist to the left so that he will land on Finn in such a way that he will suffer broken legs, but neither he nor Finn will be killed. If Jake does not maneuver his body he will be unharmed and Finn will be killed.⁵¹

Options 2: A trolley threatens Finn. Jake-the-bystander can jump in front of the trolley so that he will suffer broken legs, but Finn will be saved.

Comparing *Options 1* and *2* illustrates that the duties that we can be expected to bear depends, in part, on our placement within the situation. Intuitively, in *Options 1*, Jake-the-threat ought to burden himself. Although Jake-the-threat bears no responsibility for his lethal movements, it does seem that he is required to suffer broken legs. Comparatively, in *Options 2*, it seems implausible that morality requires Jake-the-bystander to voluntarily break his legs. The cases are similar for the fact that Jake and Finn are in a threat situation. Yet despite this common denominator, it *does* seem plausible that morality will demand more from Jake-the-threat than Jake-the-bystander. It is worth underscoring that this is not an ordinary duty of aid: we are not typically required to suffer a cost as burdensome as broken legs. Thus, the asymmetry between threats and bystanders runs deeper than it initially appears to: threats can be subject to especially burdensome moral requirements that bystanders are exempt from.

McMahan and Otsuka would, I think, accept the claim that Jake-the-threat is required to burden himself. Otsuka, for example, considers the following case:

Flagpole: You are holding a flagpole upright. You come to realize that if and only if you do nothing but continue to hold this flagpole, a falling threat will be impaled on the pole

⁵¹ Frances Kamm considers a similar case in *Creation and Abortion* (1992, 48-54).

but you will be unharmed. If you drop the flagpole, then the threat will land on you, killing you, but surviving the fall himself. (89)

Otsuka argues that in *Flagpole* one is sustaining a lethal sequence of events and that since there is no significant moral distinction between sustaining and initiating a lethal sequence of events, it is impermissible to continue holding the flagpole (89). Following this line of thought, Otsuka seems committed to the claim that Jake-the-threat should twist to the left in *Options 1*. If Jake-the-threat allows himself to continue falling, he is sustaining a lethal threat to Finn. And, if there is no difference between sustaining and initiating a lethal sequence of events, Jake should be required to heavily burden himself for Finn's sake.⁵²

Threatened parties also seem to be unlike ordinary bystanders. To wit:

Options 3: Jake is falling, through no fault of his own, towards Finn. If he lands on Finn, Jake will be unharmed and Finn will be killed. Finn can step to the left so that Jake will hit the ground and die. Or, Finn can position himself in such a way that Jake's falling on him will result in Finn suffering a broken arm, but both parties will survive.

In *Options 3*, Finn can step to the side and cause Jake's death as a side-effect or he can save his own life and Jake's by taking on the cost of a broken arm. Intuitively, Finn is required to suffer this cost in order to prevent Jake's death and this makes him unlike an ordinary bystander.⁵³

⁵² Notably, Otsuka's analysis of *Flagpole* commits him to the view that Jake-the-threat would be required to kill himself if he were able to. Do threats have a duty to die? Frances Kamm claims that although falling threats may be obligated to burden themselves with something like broken legs, they cannot be obligated to kill themselves (1992, 48). For present purposes, I will remain agnostic on this question but I will assume that Jake would have the duty to suffer something as severe as broken legs.

⁵³ This, of course, is not to say that Finn would be required to die for Jake's sake.

Given that the duties being considered here are not ordinary duties of aid—we are not ordinarily required to suffer broken legs or arms—it is worth explaining why Jake and Finn seem to have the duties that they do. Consider *Options 1*. The fact that Jake’s agency is involved changes our description of what is happening: if Jake is able to twist his body he threatens Finn as a person-who-is-falling rather than as a falling-object-who-happens-to-be-a-person. Put another way, having the option to maneuver moves Jake into the realm of agency: he is no longer simply a falling person, but a person who is falling *and* sustaining a threat by choosing not to maneuver his body. Jake is no longer a mere passenger in his body (as he is in *Threat*); he becomes the driver.⁵⁴

We might think that the presence of choice transforms Jake’s falling into an intentional killing: if he continues to fall, rather than maneuver, he is aiming at Finn’s death. More specifically, if Jake can only survive by landing on Finn, then it seems fair to say that he is aiming at Finn’s death insofar as he would be using Finn’s body as a means to break his fall. However, the verity of this seems to depend on the particularities of the situation. Suppose, for instance, that Finn is standing on soft foam and that Jake can safely land on this foam. In this situation, Finn’s death would only be a side-effect of Jake’s safe landing. That said, both cases seem equally hard to justify: falling on someone—even if we are not using their death as a means to survival—is akin to continuing to drive on a road even if you know that you will hit a pedestrian. The fact that the killing is foreseen rather than intended does not seem to affect the

⁵⁴ The question arises as to how burdensome the costs of the self-burdening thesis might actually be. Might Jake be required to suffer the loss of a limb or paraplegia? Might Finn? I suspect, though I do not have room to argue, that Jake would be required to suffer such costs, but that Finn’s costs may be slightly lower. Finn seems to occupy a position that is in between a bystander and a lethal threat. Nevertheless, for present purposes, it will suffice to note that the costs in *Options 1* and *3*—broken legs for Jake and a broken arm for Finn—are much higher than the costs that an ordinary bystander would be required to bear.

permissibility of continuing to fall; either way, Jake's survival would be the *cause* of Finn's death and this seems to be the point of significance.

A similar analysis can be made with respect to Finn. In *Options 2*, if Finn steps to the left, he must see his survival as causing the death of Jake. Finn would not, of course, be intentionally killing Jake nor would he be sustaining a lethal sequence of events in the same way as a falling person. Nevertheless, he would be surviving at the cost of Jake's life. Finn's survival, in other words, creates a lethal side-effect: in saving himself, Finn shifts the lethal burden onto someone else. Although it is not necessarily impermissible to shift a lethal burden, it does seem plausible that we are required, if we are able, to reduce the prospective harm of our actions.

We can encapsulate our intuitions about *Options 1, 2, and 3* as follows:

The-Principle-of-Cost-Mitigation: We are required to bear heavy burdens when we find ourselves (A) sustaining a lethal sequence of events, or (B) when we would, by saving ourselves, cause a lethal side-effect. This is true even when we are not responsible for the fact that we are in these situations.

Comparing *Options 1* and *3* reveals that causal positioning is quite significant with respect to the duties that one can be expected to bear within a threat-type situation and that the assumption made by McMahan and Otsuka about the irrelevance of causal positioning is not one that we should accept. Additionally, the discrepancy between *Options 1* and *2* reveals that liability, with respect to threat-type cases, is not the only salient issue. Although neither Jake nor Finn can be deemed responsible for creating the situation that they are in, it is wholly plausible that they are subject to special moral requirements simply in virtue of their positioning. We can be required to diminish a threat (that we pose or face) *even if* we are not causally responsible for the fact that we

are in the position in which we now find ourselves and *even if* doing so comes at a significant cost to ourselves.

The Principle-of-Cost-Mitigation seems to be true simply because it is fair. Neither Jake nor Finn are responsible for creating the situation that they are in and, as such, it only seems fair that they mitigate the cost of the situation as best as they can. Generally, when we find ourselves part of an unfortunate situation with another person, we ought to work together to lessen the burden that we face. More precisely, it seems plausible to characterize the duties in *Options 1* and *3* as cases of cost mitigation: situations where two agents are faced with the task of dividing a significant burden between them. To support this recasting, consider how neatly our intuitions about *Options 1* and *3* map onto our intuitions provoked by straightforward burden splitting cases. For example:

Gas Mask: Jake and Finn are working in a laboratory when a lethal gas starts to leak.

There is only one working gas mask.

In *Gas Mask*, Jake and Finn can only divide the burden between them so that one of them will live. Prima facie, the gas mask ought to be distributed according to some sort of fair outcome procedure such as flipping a coin. Both parties have equal reason to want to survive and it only seems fair that they each have an equal chance to possess the mask. If Jake obtains the gas mask fairly, then it seems permissible for him to keep it and it seems impermissible for Finn to take it from him by force.

The task of division can be made less simple by slightly altering the case:

Tablet: Jake and Finn are working in a laboratory when a lethal gas starts to leak. There is a tablet that, if ingested, will provide full immunity from the effects of the gas.

Alternatively, taking half of a tablet will prevent death, but will not prevent chronic health problems.

In *Tablet*, it seems that Jake and Finn ought to split the tablet. If they do so, they will each suffer a significant burden, but neither of them will die. Additionally, it seems wrong for either of the agents to refuse to share the tablet. In this situation, splitting the tablet is a way of dividing the burden between them.

Gas Mask and *Tablet* illustrate the general parameters of the Principle-of-Cost-Mitigation. Where there are options for distributing a burden between both agents these options ought to be pursued. And, if an agent can significantly lessen the burden of another agent—even at a significant cost to herself—then they ought to do so. Notably, the intuition provoked by *Tablet* is the same intuition attached to *Options 1*, and *3*: in each case, it is possible for one agent to prevent death for another, but only by suffering a substantial burden herself. This seems, moreover, to be a matter of fairness: since neither agent is responsible for the situation at hand, it seems fair that they alleviate the misfortune.

What does this tell us about the original *Threat* case and the right to self-defense? If we think of *Threat* as a burden splitting situation then a prohibition on killing threats can be seen to accord with fairness. If Jake obtains the gas mask fairly, it is clearly impermissible for Finn to use lethal force: in losing the coin toss, Finn is unlucky but he has no complaint that would legitimate the use of lethal force. Keeping this point in mind, it is easier to see why it would be wrong for Finn to kill Jake in the original *Threat* case. Assuming that both Finn and Jake have an equal chance of being the threat, Jake—insofar as he happens to occupy the least harmful position—seems akin to an agent who *already* has the gas mask. Of course, if he is able to lessen the burden for Finn in some way—even at a substantial cost to himself—then he is required to do so. But, in

a *Threat* case where there are no options for re-distribution, the burden has already been distributed at random.

We can bring the above thought into sharper focus by thinking about the case from an ex-ante perspective:

Ex-Ante Threat: Jake and Finn know that tomorrow they will end up in a *Threat* scenario.

They do not know who will be falling or who will be on the ground. Each party has a fifty percent chance of being the threat.

Since both parties have an equal chance of ending up in the least harmful position, it seems safe to say that the burden has been distributed at random. Of course, no coin toss has actually been performed, but given the randomness of the situation this fact does not seem to matter; neither party has been denied a fair chance at escaping the burden. From an ex-ante perspective, there would be no reason to grant the person on the ground the right to use lethal force. Shooting the falling person would be akin to shooting Jake for his gas mask after he has won the coin toss. Thus prohibiting the use of lethal force accords with fairness.

In conclusion, this chapter has sought to accomplish two tasks. First, I have argued that the Means Distinction fails to overturn a prohibition on killing innocent threats. Killing people opportunistically, or so I have argued, does not seem to be morally worse than killing people eliminatively. Additionally, I have argued that even if we accept that the Means Distinction carries some normative significance, it does not seem to be operational in threat-type cases; killing a threat opportunistically does not seem to be significantly different from killing a threat eliminatively. If there is a significant difference between killing threats and bystanders, the Means Distinction fails to uproot it. In the second half of this chapter, I have argued that threat-

type cases ought to be recast as matters of fair burden distribution. Both parties in a threat situation—threats and the threatened—are subject to especially burdensome duties that ordinary bystanders are exempt from. In line with this view, I have argued that a prohibition on killing threats is consistent with the value of fairness.

Bibliography

- Anderson, Elizabeth. 1993. *Value in Ethics and Economics*. Cambridge: Harvard University Press.
- Anscombe, G.E.M. 1958. "Modern Moral Philosophy," *Philosophy*, 33(124): 1-19.
- .1981. "War and Murder". *Papers of G.E.M. Anscombe Volume III: Ethics, Religion and Politics*. University of Minnesota Press.
- Brand-Ballard, Jeffrey. 2004. "Contractualism and Deontic Restrictions." *Ethics* 114: 269-300.
- Brennan, Samantha. 1995. "Thresholds for Rights." *The Southern Journal of Philosophy* 33 (2): 143-168.
- Darwall, Stephen. 1986. "Agent-Centered Restrictions from the Inside Out." *Philosophical Studies* 50: 291-319.
- . 2007. "Reply to Korsgaard, Wallace, and Watson." *Ethics* 118 (1): 52-69.
- . 2009. *The Second-Person Standpoint*. Cambridge: Harvard University Press.
- . 2010. "But it Would be Wrong." *Social Philosophy and Policy* 27 (2): 135-157.
- Ellis, Anthony. 1992. "Deontology, Incommensurability, and the Arbitrary." *Philosophy and Phenomenological Research* 52 (4):855-875.
- Enoch, David. 2011. "Giving Practical Reasons." *Philosophers' Imprint* 11 (4)
- Feinberg, Joel. 1980. *Rights, Justice and the Bounds of Liberty*. Princeton: Princeton University Press.
- Foot, Philippa. 1978. "Morality as a System of Hypothetical Imperatives." *Virtues and*

- Vices* Oxford: Oxford University Press.
- . 1985. "Utilitarianism and the Virtues." *Mind*. (374): 196-209.
- Frowe, Helen. 2008. "Equating Innocent Threats and Bystanders" *Journal of Applied Philosophy*. 277-289.
- Geach, P.T. 1969. *God and the Soul*. London: Routledge Press.
- Hanna, Jason. 2012. "The Moral Status of Nonresponsible Threats" *Journal of Applied Philosophy*. 19-32.
- Hume, David. 1985. *Enquiry Concerning the Principle of Morals*, in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Ed. L.A. Selby-Bigge, 3rd ed., rev. P.H. Nidditch. Oxford: Clarendon Press.
- Hurley, Paul. 2011. *Beyond Consequentialism*. Oxford: Oxford University Press.
- Kagan, Shelley. 1998. *Normative Ethics*. Boulder: Westview Press.
- Kamm, Frances. 1989. "Harming Some to Save Others." *Philosophical Studies* 57: 227-260.
- . 1992. *Creation and Abortion*. Oxford: Oxford University Press.
- . 1992. "Non-Consequentialism, the Person as an End-in-Itself, and the Significance of Status." *Philosophy and Public Affairs*. (21): 354-389.
- . 2007. *Intricate Ethics*. Oxford: Oxford University Press.
- . 2011. *Ethics for Enemies: Terror, Torture, and War*. Oxford: Oxford University Press.
- Kumar, Rahul. 1999. "Defending the Moral Moderate: Contractualism and Common

- Sense." *Philosophy and Public Affairs* 28: 99-122.
- . 2001. "Contractualism on Saving the Many." *Analysis* 61 (2): 165-170.
- . 2003. "Reasonable Reasons in Contractualist Moral Argument." *Ethics* 114:6-37.
- Lebar, Mark. 2009. "Virtue Ethics and Deontic Constraints." *Ethics* 119 (3): 576-580.
- Lippert-Rasmussen, Kasper. 1999. "In What Way Are Constraints Paradoxical?" *Utilitas*.
586-615.
- . 2009. "Kamm on Inviolability and Agent Relative Restrictions." *Res Publica*. 15
(2009): 165-178.
- MacIntyre, Alasdair. 2007. *After Virtue*, 3rd Edition. University of Notre Dame Press.
- McDowell, John. 1979. "Virtue and Reason." *The Monist* 62: 331-350.
- McMahan, Jeff. 2009. *Killing in War*. Oxford: Oxford University Press.
- Moore, Michael. 2008. "Patrolling the Borders of Consequentialist Justifications: The
Scope of Agent-Relative Restrictions." *Law and Philosophy*. 27 (1): 35-96.
- Mill, John Stuart. 1987. *Utilitarianism*. Ed. Alan Ryan. London: Penguin Books.
- Nagel, Thomas. 1979. "War and Massacre." *Mortal Questions*. Cambridge: Cambridge
University Press.
1986. *The View from Nowhere*. New York: Oxford University Press.
- . 1995. "Personal Rights and Public Space." *Philosophy and Public Affairs* (24): 83-
107.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Nussbaum, Martha. 2001. "Comment." *Goodness and Advice*. Ed. Amy Gutmann. New
Jersey: Princeton University Press.
- 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):325-246.
- . 2007. “Culpability and the Definition of Deontological Constraints.”
Law and Philosophy. (27): 105-122.
- . 2008. “Specifying Rights out of Necessity.” *Oxford Journal of Legal Studies*. 28 (1): 19.
- . Otsuka, Michael. 1994. “Killing the Innocent in Self-Defense”. *Philosophy and Public Affairs*. 74-94.
- .1997. “Kamm on the Morality of Killing”, *Ethics*, (108): 197-207.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- . 2011. *On What Matters: Volume One*. Oxford: Oxford University Press.
- Quinn, Warren. 1991. *Morality and Action*. Cambridge: Cambridge University Press.
- Quong, Jonathan. 2009. “Killing in Self-Defense”. *Ethics*. 507-537.
- Rachels, James. 1970. “On Moral Absolutism.” *Australasian Journal of Philosophy*. 3 (48): 338-353.
- Rawls, John. 1995. “Fifty Years after Hiroshima.” *Dissent*: 323-327.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.
- .1999. *Engaging Reason*. Oxford: Oxford University Press.
- . 2003. “Numbers With and Without Contractualism.” *Ratio* 16 (4):346–367.
- Ross, W.D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Sartre, Jean-Paul. 1975. “Existentialism is a Humanism.” *Existentialism from Dostoyevsky to Sartre*. Ed. Walter Kaufmann. New York: Meridian-New American.

- Scanlon, T.M. 1998. *What We Owe to Each Other*. Harvard: Harvard University Press.
- . 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Harvard: Harvard University Press.
- Shafer-Landau, Russ. 1995. "Specifying Absolute Rights." *Arizona Law Review*. (37): 209-224.
- Scheffler, Samuel. 1982. *The Rejection of Consequentialism*. Oxford: Clarendon Press.
- . 1985. "Agent-Centered Restrictions, Rationality, and the Virtues." *Mind*. (375): 409-419
- Smilansky, Saul. 2003. "Can Deontologists be Moderate?" *Utilitas*. 15 (01): 71-75.
- Sreenivasan, Gopal. 2010. "Duties and their Direction." *Ethics*. 120 (3): 465-494.
- Tadros, Victor. 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press.
- Thomson, Judith Jarvis. 1990. *The Realm of Rights*. Harvard: Harvard University Press.
- .1986. "Some Ruminations on Rights." William Parent (ed.) *Rights, Restitution, and Risk* Cambridge, MA: Harvard University Press.
- . 1986. "The Trolley Problem". *Rights, Restitution, and Risk*. Harvard: Harvard University Press. 94-116
- Wallace, R. Jay. 2007. "Reasons, Relations, and Commands: Reflection on Darwall." *Ethics* 118: 24-36.
- . "The Deontic Structure of Morality." (Unpublished Manuscript).

- Walzer, Michael. 1977. *Just and Unjust Wars*. New York: Basic Books.
- . 1983. *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.
- Watson, Gary. 2007. "Morality as Equal Accountability: Comments on Stephen Darwall's the Second-Person Standpoint" *Ethics* 118 (1): 37-51.
- Wiggins, David. 1997. "Incommensurability: Four Proposals." Ed. Ruth Chang. *Incommensurability, Incomparability, and Practical Reason*. Cambridge: Harvard University Press.
- Williams, Bernard. 1973. *Utilitarianism; For and Against*. Cambridge: Cambridge University Press.
- . 1981. "Conflicts of Values." *Moral Luck*. Cambridge: Cambridge University Press.