

UCLA

UCLA Electronic Theses and Dissertations

Title

Mitigation, Deliberation, and Moral Assessment

Permalink

<https://escholarship.org/uc/item/5b2452sb>

Author

Cvrkel, Tiffany

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Mitigation, Deliberation, and Moral Assessment

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Philosophy

by

Tiffany Teeman Cvrkel

2012

© Copyright by
Tiffany Teeman Cvrkel
2012

ABSTRACT OF THE DISSERTATION

Mitigation, Deliberation, and Moral Assessment

by

Tiffany Teeman Cvrkel

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2012

Professor Barbara Herman, Chair

Many philosophers have explored what it is to commit a moral wrong and how to understand the violation of moral obligations. Many more have written about the nature of blame and agency. But much is left to say about the most common kind of moral wrong, the type of wrong that populates our everyday interactions. These are the wrongs performed by agents who are in the grip of anger or fear. Or who undervalue their obligations. Or who are intoxicated, or impaired, or coerced. When we morally assess wrongs like these, we don't get clean black and white declarations of moral success or failure. We get a lot of muddy gray. We get mitigating circumstances.

It is easy to consider mitigating circumstances rare and exotic – cases that exist in courtrooms and police procedurals. But this is a mistake. Moral mitigation, cases where the circumstances of an agent's moral performance make it appropriate to reduce the amount of blame she deserves for her wrong act, might be the norm and not the exception. And so figuring out how mitigation works will tell us how to get moral assessment right in the real world.

I will argue that mitigating circumstances can best be understood as circumstances that affect the deliberative part of action, and that successes and failures in the component parts of deliberation must be taken into consideration during moral assessment. In Chapter 1, I offer an introduction to the context of the project, discussing the scope and aim of this dissertation. Here I offer a number of preliminary remarks to set up the discussion to follow.

Chapter 2 focuses on the moral assessment process itself. I discuss the nature of the moral assessment process and suggest that approaching the moral assessment of action from the outside in – from the obligations that bind the agent all the way back to the deliberative parts of intention-formation – will offer a number of advantages when assessing action. This method will allow us to make room for a common sense notion of mitigation that includes reduction in blame, and it will also make way for us to draw moral distinctions between actions that differ in their inner-deliberative content and process.

In Chapters 3 and 4, I expand the picture of moral assessment I offered earlier, arguing that mitigating circumstances will turn out to be circumstances that cause or reflect distinctive kinds of performances in the inner-deliberative content and process. I will explore how this approach can handle a number of intuitively mitigating circumstances, such as coercion, adolescence, and intoxication.

And finally, in Chapter 5, I will situate this picture within the larger landscape of moral assessment, orienting this approach to mitigation alongside excuse and justification. I conclude with an appeal to consider this expansion of the moral assessment process as a means to handle complex and nuanced moral cases.

The dissertation of Tiffany Teeman Cvrkel is approved.

Alexander Jacob Julius

Julie Ane Tannenbaum

Pamela Hieronymi

Barbara Herman, Committee Chair

University of California, Los Angeles

2012

For Edward

TABLE OF CONTENTS

Abstract of the Dissertation	ii
Acknowledgements.....	vii
Vita.....	x
Chapter 1: The Big Picture	1
Chapter 2: Reductions and Redescriptions	13
Chapter 3: Moral Assessment and Deliberative Irregularities.....	46
Chapter 4: Mitigating Circumstances	70
Chapter 5: Excuse, Justification, and Mitigation.....	104
Conclusion	125
Bibliography	127

ACKNOWLEDGEMENTS

To begin, I owe much gratitude to my advisor, Barbara Herman. Barbara is not merely a philosopher. It is probably more accurate to call her a scholar without limit, because her interest and insight are so much greater than a single discipline could accommodate. She is voraciously curious and engaged with the world in and outside of philosophy, and she taught me to be curious, to find philosophical inspiration in all facets of life. This is part of what makes her such a rare and gifted philosopher, and part of what makes her such a rare and gifted adviser. Wherever my work took me, she was one step ahead, seeing where I wanted to go long before I got there. I am grateful for her advice, support, and suggestions. I am humbled having gotten to work with her.

I would also like to thank A.J. Julius. There is no human being more generous with his ideas and feedback, and with his time and support. No matter the topic of our discussion, he teaches me something with every conversation. I am a better philosopher – and person – for having worked with him. Much of Chapters 2 and 3 are a direct result of A.J.’s questions and comments, though I have no doubt that my discussion does not do his questions justice.

Pamela Hieronymi, Julie Tannenbaum, and Andrew Hsu have all helped me in ways big and small. Each of them has made my work better in significant ways. I owe them great thanks.

There are a number of people who have commented on early drafts of this work, especially early versions of Chapter 5. I owe special thanks to David Goldman, Jorah Dannenberg, Lewis Powell, and members of the UCLA Ethics Writing Workshop. This work was also made better by audiences at California State University, Northridge’s colloquia, where I

presented early versions of Chapter 4 and Chapter 5. Their insightful comments and suggestions were invaluable.

Many people have helped with comments, advice, and suggestions. The following list is incomplete, but I want to thank Janelle DeWitt, Jeff Helmreich, Rob Hughes, Diane Kierce, Ari Lev, James Rocha, Ingrid Steinberg, Jessie Summers, and Steve White. The people I have accidentally omitted deserve both my apologies and my gratitude.

I would also like to thank the staff of the UCLA Philosophy Department. They have guided me through the logistics of graduate school with kindness and wisdom, and I am so grateful. Kristin Olson in particular has played the role of teacher, therapist, and super-hero for me, more than once.

The undergraduates at UCLA have also changed who I am as a teacher and philosopher. I have gotten to work with remarkable students, and their enthusiasm and dedication have pushed me to grow in both my teaching and my writing. My students have let me share my embarrassingly goofy love of philosophy with them, and they have given me so much more in return.

I am blessed with an amazing support system. Thank you to my sister Sarah, who has spent countless hours on planes to Los Angeles to see me through this process. From the day that she set up a color-coded filing system to manage my graduate school applications, to the day she started planning for my hooding ceremony: she has been my sounding board and confidante.

For my beautiful husband Edward, there are no words that are sufficient. He is my partner in crime and philosophy, the reason why I get up in the morning. I could not have done this without him. Even my darkest days are joyful because he is in my life. He is my support, my inspiration, and the best conversationalist in the known universe (and probably beyond).

Finally, I would like to thank Laurence Bonjour, the man who changed the course of my life, one philosophy class at a time. I have never met a better teacher or mentor, and I entered academia because he taught me the power of the classroom. Over a decade ago, he shared his love of philosophy with a lost undergraduate. My life is forever better for it.

VITA

Education

University of California at Los Angeles, C. Phil, Fall 2008
University of California at Los Angeles, M.A., Spring 2008
University of Washington at Seattle, B.A. with distinction, Winter 2004

Employment

UCLA Molecular, Cell, and Developmental Biology	Instructor, 3/2012-12/2012 Teaching Fellow, 2011
UCLA Philosophy Department	Instructor, Summer 2008-2011 Teaching Fellow, 2005-2011

Presentations

CSUN Colloquium, April 2010, "Moral Responsibility, Excuse, and Mitigation"
APA Pacific Division Meeting, March 2010, Commentator, "On Swenson's Defense of Nonsocratic Skepticism"
UCLA Ethics Workshop, March 2010, "Moral Responsibility, Excuse, and Mitigation"
UCLA Ethics Workshop, February 2009, "Children, Adolescents, and Moral Capacities"
Albritton Society, June 2008, "Becoming Good: Aristotle, Virtue Acquisition, and Moral Luck"
UCLA Ethics Workshop, March 2008, "The Difficulties in Aristotle's Theory of Virtue Acquisition"
Northwest Student Philosophy Conference, May 2003, "Of Sickness and Health: The Pursuit of Wellness in Contemporary Medicine"

CHAPTER 1: THE BIG PICTURE

In both legal and moral contexts, we speak of mitigation. Yes, she killed him, we might say, but she had just found him in bed with another woman. There were mitigating circumstances. These particulars – whatever comes after that “but” – are supposed to play some role in our assessment of the moral status of her act. We add clauses like these when we think there are circumstances that alter the landscape of assessment. Perhaps these particular circumstances make it such that the agent deserves less blame for her act, or that she deserves less punishment. Perhaps they change the nature of the act itself. We speak these qualified statements when we think that the circumstances surrounding the act effectively mitigate something, but it is still an open question as to what these circumstances are, or why they alter moral assessment.

Mitigation’s distinctive feature is that it demands reduction in responses like blame or praise without eliminating them entirely. When we speak of mitigating circumstances, we do not mean circumstances that make moral and legal responses entirely inappropriate; mitigation is not elimination. Much has been written about what being a normal morally responsible agent in normally functioning circumstances looks like, someone who is paradigmatically deserving of full blame for any and all morally wrong acts that she’s committed. And much has been written about excuses, circumstances that prevent an agent from being morally blameworthy – at all – for a certain act, or responsible as an agent full-stop. In this dissertation, I aim to offer discussion of the cases in between.

One of the first questions that we can raise about mitigation is whether understanding its nature in terms of reduction is foundationally accurate. Take the following cases: suppose an agent walks in on her lover in bed with another woman. Enraged, the agent picks up a nearby

vase, strikes her lover on the head, and kills him. The killing – the murder! – is performed intentionally. She hit him with the vase because she wanted him dead. The agent in this case is not incapacitated by her rage, though she is certainly feeling a great deal of anger. The scenario I have in mind is not a case of compulsion, or a case of being fully incapacitated, or overcome, by strong emotion. Let me stipulate that this agent does not have grounds for excuse; she lacks a condition that will let her off the moral hook entirely. She has the relevant moral knowledge, and she believes and understands that murder is wrong. She deliberates, she acts intentionally, and it is accurate to say that she murdered her lover. She is no doubt deserving of blame, in some amount, for her wrong action.

For contrast's sake, let's consider another murder. In this case, another agent also kills her lover. But in this case, she has simply grown tired of her lover, and she knows that she's the beneficiary to his enormous fortune. Motivated by the acquisition of wealth and the ending of a lackluster relationship, she coldly premeditates, investigating methods that will prevent her from getting caught. She acquires the ingredients for a poison, assembles them, and drugs her lover's drink.

Both the Betrayed Lover and the Calculating Lover have committed murder. Both are morally responsible for their actions, and both no doubt deserve blame for those actions. And yet... I suspect that most of us think that there is something interesting about the circumstances of the Betrayed Lover that makes her deserving of less blame than the Calculating Lover. Legal distinctions between first- and second-degree murders are grounded in just such intuitions. We think there is something about the Betrayed Lover case that makes the moral assessment of her action importantly different from the assessment of the Calculating Lover, either because the

action is less wrong or because she deserves less blame. We might think that something in her circumstances is mitigating, while the Calculating Lover has no corresponding claim.

But it is an open question whether we should understand these two cases as two entirely different types of wrongs, each justifying the assessment of some amount of blame for their respective agents, or whether we should think of these as two performances of a single type of moral wrong – murders, in these cases – where one agent has special circumstances that serve to reduce the amount of blame she deserves and the other does not. I call this second approach the “common sense” notion of mitigation. In Chapter 2, I explore this question, arguing that thinking of mitigating circumstances as circumstances that force redescription and not reduction makes it difficult to properly handle the special moral features of cases such as coercion. I will offer an approach to moral assessment that includes room for the common sense notion of mitigation by working from the outside in, starting with the obligations that bind an agent, considering the intention of the agent, and then evaluating the morally relevant deliberative details that led to the intention formation. This approach looks at the plot of the violation, so to speak, letting us compare how particular instantiations of different types of wrongs might be considered better (where the agent deserve less blame) or worse (where the agent deserves more blame). This concept of “better” and “worse” will allow us to compare actions to one another within moral categories, making sense of the notion of reduction.

In Chapter 2, I also break the moral assessment process into stages. I will suggest that we must first look to the agent; we check if she meets the necessary and sufficient criteria for agency, and we also look to see what obligations bind her. The second step will become enormously important for the assessment of deliberation. Once we’ve assessed the status of the agent, we turn to the fullness of the action. Using the outside-in approach, we start first by

identifying which obligation(s) the agent violated. And then we turn to the deliberative and intention-formation part of her action. Why did she perform the act she did? At this stage, we can evaluate her deliberation in three ways. First, we look to see if she took the demands of her moral obligation to be both worth consideration and as appropriately reasons-generating. Next, we evaluate the deliberative path to her intention. We must ask: what did she take to be relevant to her deliberation, and what did she choose to deliberate about in relation to the content of her obligation? In other words, we check if her deliberative process reflects an acknowledgement of the obligations that bind her, and if she populates her deliberation with only appropriate considerations. Finally, after we assess what she takes as deliberation-worthy, we evaluate how she actually deliberates. Does she weigh her considerations correctly? And if not, in which way did she err, and why?

In this assessment process, there are several places where it is appropriate to assign blame. At the first level, we can ask whether the agent took her obligations seriously enough to deliberate about them at all. If she performs a wrong act because she simply didn't take her moral obligations to be reason-worthy and fails to include them in her deliberation at all, or if she enthusiastically disrespects her obligations through the entire deliberative process out of malice – and assuming she is a normally functioning agent – then we can assert that she deserves maximal blame for the wrong. But if she did treat her obligations as deliberation-worthy, then we must move to a more nuanced form of assessment. Namely, we must investigate whether the deliberation surrounding her obligation contained the appropriate respect for the object of her obligation. We do this by evaluating the rest of her deliberative process, to see if she populated her deliberation with only appropriate considerations, and if she gave proper respect to the source of her obligations by only weighing them against other morally relevant considerations.

And finally, we move to the content of her deliberation. Did she weigh her considerations appropriately, or did she undervalue her obligation? If she did undervalue her obligation, what exactly did she overvalue?

In other words, as part of the moral assessment process, we look to see if there is the presence of what I will call morally relevant deliberative irregularities. These deliberative irregularities may take a number of forms, but they all share a feature in common. They are cases where there is asymmetry in the three parts of the deliberative process that I described above – at least one success, coupled with at least one failure. In Chapter 3, I will argue that moral obligations demand we acknowledge the status of the object of our obligations at each stage of our deliberation and intention-formation process, and this acknowledgment will require us to regulate our deliberation appropriately. Blame attaches to the violation of this obligation, the absence of respect, and with cases of deliberative irregularity, this respect is not absent at every stage. The agent will deserve blame for the particular failures, but not the particular successes, that constitute the entirety of her action. And so the presence of these irregularities will indicate that the agent has a claim to mitigation. The irregularity itself isn't the cause of mitigation directly; the fact that the agent only fails to regulate what she is obligated to regulate at certain parts of the process is what does the moral work. But these deliberative irregularities are evidence of the mitigating circumstances.

In the second half of Chapter 3, I illustrate how this works with a circumstance that many find intuitively mitigating: coercion. I offer a case where a person has taken a vow to keep certain things secret and away from dangerous hands. Perhaps she is responsible for protecting the identity and location of political refugees from oppressors who seek to do them further harm. She is bound by the obligation to keep that vow and to protect those who would otherwise be in

danger. The coercer has forced her into a choice between keeping her vow and losing her life, and while she is morally bound by her promise, she also has morally relevant reasons to preserve her life. The same would hold if the coercer threatened to kill her family, or people across town. These threats carry distinctly moral weight. Let's stipulate that the right moral decision is for our agent to keep her secrets,¹ but she deliberates, and as a result of coercion, she fails to do the right thing. Are her circumstances mitigating? I claim that coercive circumstances have mitigating force because they alter the moral landscape in a very particular way. If a meaningful threat was made and the case of coercion is genuine, then the agent now must weigh two morally relevant considerations. The preservation of her life has moral worth, and by considering whether to save her life or keep her vow, she is weighing two reasons for acting that are appropriate to weigh. In this case of coercion, her deliberation is populated only by appropriate considerations, even if she mis-weighs those considerations in her deliberative process. And so she has a claim to mitigation. It is important to note that I will not be claiming that the agent's lack of wholeheartedness justifies mitigation in these cases, at least, not the mere lack of wholeheartedness. I will argue that it is instead relevant *why* the agent is not wholehearted. It is morally appropriate for an agent to be torn between conflicting moral demands, and since this is the way coercive circumstances affect this agent's deliberation, her coercive circumstances are mitigating. And so we will have an example of mitigation emerge.

I will argue that story of mitigation is ultimately a story about moral assessment. My claim is that the demands of moral obligation occur not just at the level of intention, but also at the level of deliberation and intention-formation. A moral obligation does not simply concern our

¹ This is far from clear for all cases of promise-keeping, and in fact, it may turn out that the morally correct choice is to spill the secrets in some circumstances, depending on the nature of the agent's promise and the secrets being kept. But for my purposes, I am imagining a case where a genuine moral wrong has been performed, and the agent does indeed deserve some amount of blame.

action; it concerns how we come to our actions. And thus, it only makes sense that we will be vulnerable to moral assessment throughout each part of the deliberative process, including both the content of the deliberation and the choice of which considerations to deliberate about. And deliberation does not always take place in a unified, harmonious way. If agents are vulnerable to the assignment of blame for the character of their deliberation, then an appropriate way to understand mitigation will be as a response to non-complete failures. To see why, it will be helpful to say a bit more about deliberation.

There are many different theories as to the shape and nature of deliberation. Nevertheless, I take certain features to be basic. We often say that people act for reasons, and this process will no doubt include a variety of desires, principles, beliefs, and judgments. As part of genuine action, an agent goes through some process of forming an intention, and at some point during intentional action, she must determine what counts as a reason to act. As part of a normal deliberative process, the agent assesses facts about her situation, weighing means-ends connections and the like, and considers candidates for intention. Not all considerations are reasons, of course, and so there must be a process – however it gets cashed out – of what makes desires, principles, beliefs, judgments, and so on, count as reasons for the agent. Some of these might be reason-giving as a means to fulfill an end, and some, presumably, will be reason-giving as ends in themselves. If I am deliberating about how to get coffee, my belief that the café across the street has coffee will, combined with my desire for coffee, give me reason to start walking across the street.

How I settle on what counts as reasons, and which reasons move me to action, is controversial. We might say that an agent picks a reason to act, or is pulled by the strongest reason, the reason she thinks best, the reason that reflects her values, or something of the sort.

We might say that she wills because of her reasons, that there is something above and separate from these desires, principles, emotions, and so on, and it is that faculty that generates action.

We might say that her actions are simply the product of her psychology, and her beliefs, desires, and mental states will simply determine that she is responsive to reasons of a certain sort. All of these possibilities are dependent on notions of activity and passivity in mental life, and questions about how our intentions, motives, and actions are shaped. And while these are interesting and important questions, the specifics of this part of the deliberative process need not be fully explained for us to understand their relation to moral assessment. So I will say this: by whatever means, an agent is moved to act by reasons, and she carries out her intention in the world via her act. This is part of what it is to be an agent performing intentional action.

As I've discussed earlier, however, there is another piece of the deliberative process, and our moral assessment is neither complete nor accurate if we focus assessment only on the final stages of intention-formation, the intention itself. Before determining which considerations will move her to action, the agent first decides whether the moral demands are worth deliberating about at all, and if so, which sorts of considerations belong on the deliberative field² along with it. I will argue that she is vulnerable to moral assessment for her performance of this process, because this performance is inherently a form of moral judgment. Placing considerations on the deliberative field is a way of showing (or not showing) moral respect, and in fulfilling the requirements of a moral obligation, in exactly the same way as other parts of intention-formation.

In the process of establishing her deliberative field, the agent takes some considerations to be relevant to the question of whether or how to act, and those considerations become

² I borrow this language from Barbara Herman because I think it captures nicely the notion that there are things that an agent brings to bear when deliberating – things that are on the field – and things that are not. My fondness for Nabokov is a fact about me, but it plays no role in whether I cross the street for coffee. In this instance of deliberation, that fondness is off the field. When I consider what book to read on a long flight, this fact will quite possibly be called up on my deliberative field for further consideration.

contenders for the reasons that might move her. When I am deliberating about whether to walk across the street, I consider my belief that there is coffee across the street and my desire for coffee as relevant considerations. There might also be other considerations I consider relevant. Perhaps I am teaching a class in five minutes, and my desire to be on time for my class, combined with my belief that it takes more than five minutes to get coffee from the café, is also relevant to my deliberative process. All of those considerations are on my deliberative field, and some combination of them will eventually move me to act (or not to act).

That said, I possess an enormous number of other beliefs, values, principles, desires, and inclinations that play no role in this particular deliberation, because part of the deliberative process is limiting the deliberative field to *relevant* considerations. I consider my desire for coffee when I form reasons for walking across the street, but I don't consider the name of the third Emperor of Japan or my desire to revisit Tanzania. One of the great challenges of artificial intelligence is the problem of capturing the uniquely rational process of judging which of almost endless considerations are relevant to a given question. Normally functioning agents do this routinely, and usually seamlessly, as part of our normal deliberative process. It is a necessary part of being a rational agent, and a necessary part of intentional action.

Assessing which considerations are first reasons-generating, and second, which ones the agent considers relevant -- which she thinks belong on the deliberative field -- are ways we can make judgments about whether deliberation is going well. If someone is seriously considering the density of plutonium on the deliberative field when she is attempting to decide whether to buy chocolate or vanilla ice cream, I believe that we can confidently say that something has gone wrong. I want to be very clear about the point I will make. It is certainly true that things which are not relevant to a potential action might "pop in our heads" as it were. We're quirky creatures,

and memory and attention work in complex ways. If someone is deliberating about whether to buy chocolate or vanilla ice cream and the density of plutonium simply comes to the agent's thoughts, this does not make her deliberative process faulty, or imply that she finds such a consideration relevant. The distinction I aim to capture is about what we take as relevant to the question of whether or how to act, and it is this piece upon which it is appropriate to morally assess the agent. This is the question of what one deems motive-worthy. One can pay attention to things without taking them as relevant and without them being part of her deliberative field. Considerations are placed on the deliberative field because the agent has judged that they have some influence on the question of whether or how to act, and each of them, minus other reasons that the agent takes to be more pressing, could serve as a the motive that moves the agent to action. Considering something on the deliberative field is to grant it a special status; it is to consider it choice-relevant. I will argue that this distinction does a great deal of work when we consider the moral assessment of the components of action. And so, let's turn our attention back to moral deliberation and moral assessment.

The deliberative process plays an important role for all stages of moral assessment. On most views of moral responsibility, having the capacity to perform this process is a prerequisite for considering someone a full-fledged agent. In his article "Freedom and Resentment," P.F. Strawson lists examples of circumstances that he considers inhibitors of reactive attitudes, circumstances where we should instead react to people with the objective attitude. He lists, among others: "He's only a child", "He's a hopeless schizophrenic", "His mind has been systematically perverted". ... [And so on. These cases present] the agent as psychologically abnormal – or as morally underdeveloped. The agent was himself; but he is warped or deranged, neurotic or just a child." What's clearly linking these examples is some notion that the person in

question is not a normally functioning deliberator, either at all or at this moment. Most of us take it as foundational that there is a necessary connection between one's status as an agent and, at the very least, the capacity to deliberate, and absence of such is typically clear grounds for moral excuse. This is no doubt what grounds legal mitigation to members of classes like the severely intellectually disabled. But moral assessment requires more than the mere capacity to deliberate. In properly assessing intentional action, we look to the deliberation that is contained within the action. Ethics demands the meeting of obligations created by the special status of our moral peers, and our relations to one another. These obligations have normative force over the content of our deliberations. It is permissible to spend time with people because we care about them, and because we enjoy it. It is not permissible to spend time with people for the sake of mocking them in our heads. The action might look the same from the outside – we're spending time with people – but the difference in reasons, the shape and content of our deliberation, makes one permissible and the other not. And so we might say that good moral deliberation requires not just acting from reasons, but acting from the right reasons, however those get cashed out.

In Chapter 4, I offer a variety of cases to illustrate how this approach to mitigation works in the real world. I discuss impairments that create instability in deliberation, particularly adolescence and intoxication. I argue that conditions such as these undermine part of the deliberative process, and so what looks like deliberative performance might actually be closer to a tic or seizure. These conditions are mitigating because they block the agent from being able to perform a piece of morally required deliberative regulation, and since we only morally assess genuine parts of the action-formation process, these phantom displays – which are not genuine parts of action – ought not be assessed. This will mean that people in these circumstances will deserve blame for the failures they genuinely perform, and not for the ones they do not.

I also discuss cases of what I call passion-inducing circumstances. These include cases of extreme quasi-sensory experiences, such as cravings or addiction, and I will argue that cases of urgency will also fall into this category. While circumstances like adolescence and intoxication impair deliberative capacity, circumstances such as these corrupt deliberation by hijacking deliberation. Passion-inducing circumstances demand attention and prioritization independently of the agent's judgments about how to regulate her deliberation, and thus they sever the relationship between agent and regulation. Since we assign blame to the agent for her moral failures during the assessment process, the assignment of blame will be inappropriate when those links are broken. And finally, I will discuss cases like the Betrayed Lover I introduced above. I will argue her circumstances are mitigating because she herself is a victim of a wrong, and her resentment is justified accordingly. I will argue that inappropriate expressions of appropriate desires and emotions can create deliberative irregularities similar in kind to coercion: they are failures to weigh appropriate considerations on the deliberative field correctly.

Chapter 5 brings a discussion of the neighboring concepts of excuse and justification. I will argue that conventional approaches to mitigation as simply lesser versions of excuse aren't successful without additional tools, and I will make the case that my method of looking for discrete moments in the deliberative and intention-formation process that show the presence of moral success and moral failure, will help fill that gap.

CHAPTER 2: REDUCTIONS AND REDESCRIPTIONS

My aim in this dissertation is to map out a picture of moral mitigation, but maps are useless without navigational context. So before focusing sharply on the particulars of mitigation, it will be helpful to get a larger lay of the land, situating the discussion to come. This chapter will focus on how we should understand the nature of mitigation within the context of moral assessment. In the first part of this chapter, I will offer some preliminary remarks on my project and try to make the case that approaching moral assessment from a “moral ledger” view is neither trivial nor without value. In the rest of this chapter, I will consider a question about the nature of mitigation, particularly how and whether the moral concept should properly be understood as a reduction of blame or a redescription of action. This will require discussing the process of moral assessment and raising a question about how we should use action description for the purposes of moral assessment. By the end of this chapter, I will be ready to begin the extended investigation into moral mitigation itself.

I.

Tackling the scope and content of moral assessment in its entirety would require dissertation-length discussions of blame and moral responsibility. These are not the tasks I aim to take up. But there is quite a bit to say about moral assessment that does not require first solving contentious foundational debates about the nature of ethics, moral responsibility, or compatibilism. First, some terminology. In this work, I will primarily be using the word “blame” to express the outcome of moral assessment. This is not the only possible outcome of the moral assessment process, of course. One can potentially find agents or actions morally neutral or deserving of praise, for example. The area I am investigating, however, is the spectrum of moral

wrongs, so my focus will be more limited. I would also like to speak briefly about what I mean by the notion of blame. Recent work on blame has included some appeals to separate the notion of blame from blameworthiness, defining blameworthiness the amount of blame that an agent deserves for her wrong action and blame as a form of interaction, a method of social communication and interpersonal relations.³ It is no doubt true that those who have mitigated blameworthiness will be deserving of less blame than those who do not have the same mitigating circumstances. But on some views, it is not true that all circumstances that mitigate blame stem from mitigated blameworthiness. For example, many people believe that overwhelming guilt or remorse for a past wrong justifies a reduction in blame for the agent, independently of what the agent deserves. People who believe that guilt or remorse are mitigating do not believe they mitigate because they reduce the blame deserved by the agent – indeed, part of the very conditions of guilt or remorse is an understanding of just how deserving of blame the agent is. The distinctive feature of this type of case is that even though the agent is deserving of full blame in the sense that she was fully responsible for the wrong act at the time, something about her future self suggests that we might be justified in reducing the amount of blame we do in fact give her. In this project, I set aside the role that blame might play outside and in addition to a factual assessment of the agent’s assignment of blame through the moral assessment process, and this necessarily limits the application of the theory of mitigation I will offer. I focus instead entirely on cases where circumstances will reduce the amount of blame an agent *deserves*, justified by the condition and circumstances of her action itself. An agent is “deserving of blame” when she is appropriately vulnerable to certain moral assessments, specifically that she is in the relevant sense the kind of creature who is open to assessment in this particular case (or is open to assessment in general), and the further assessment that she has committed a moral wrong. In

³ See T.M. Scanlon’s *Moral Dimensions* for a recent example of this approach.

other words, to deserve blame, as I use the phrase, refers to being the appropriate target of moral judgment, and in particular, the judgment that one has genuinely violated a moral obligation that binds her.

Making judgments about appropriate moral targets and the performance of moral wrongs will require some discussion about the nature of blame, and the standards of appropriateness. R.J. Wallace writes: “At issue are the conditions in which it would be appropriate to hold someone morally responsible for a particular act, in the sense of viewing the action as morally blameworthy. To hold someone responsible in this way, I have proposed, is to be subject to a reactive emotion because one believes the act to have violated a moral obligation one accepts, or it is to believe that the person’s having violated such an obligation would make it appropriate for one to be subject to such a reactive emotion.”⁴

I imagine that Wallace describes his project in terms of “appropriateness” for the same reason I use the word. There are different kinds of mistakes people can make regarding assessments. If I am performing an assessment of a runner in a race, I can fail to start the clock when she starts, fail to notice when she crosses the finish line, or fail to keep time accurately. But it is a failure of a different sort to try to determine running time for someone cheering on the sidelines. That person is not in the race. The notion of appropriateness captures both of these kinds of errors. It is appropriate to assess a given amount of blame when the conditions of desert merit it, and not appropriate when those conditions do not. The same is true for assessment of agents. If mitigation is a reduction or lessening of the amount of blame that an agent deserves for a wrong act, then mitigating circumstances will be circumstances that alter the appropriateness of blame and moral assessment. Thus, determining what makes certain conditions appropriate (or not) will be a part of understanding mitigation.

⁴ R.J. Wallace, *Responsibility and the Moral Sentiments*, p. 133

Wallace claims that appropriateness depends upon the presence of certain powers within the agent. He calls these the “powers of reflective self-control,” and it will be helpful to explore what he has in mind. Wallace argues that these powers of reflective self-control are a necessary condition for moral responsibility. This notion is loosely inspired by Scanlon’s “reflective, rational self-governance,” but for ease, I shall use Wallace’s formulation. Wallace divides the powers of reflective self-control into two parts: “(1) the power to grasp and apply moral reasons, and (2) the power to control or regulate [...] behavior by light of such reasons.”⁵ We can think of these as cognitive powers and affective powers, where the first aims to capture proficiency with the tools of moral deliberation, and the second aims to capture the ability to act on the fruits of one’s deliberative process, to be efficient agents in the world.

First, cognitive powers. Wallace writes that these cognitive powers “enable one to understand and to apply the justifications expressed in moral principles.”⁶ This capacity no doubt includes the necessary tools for deliberation in general: the ability to make accurate means-ends connections, notice relevant considerations, weigh considerations appropriately, formulate conclusions that cohere in some way with the other beliefs, judgments, attitudes, and actions of the agent, and so on. But it also requires that the agent be capable of uniquely moral deliberation, and thus we must add the capacity to grasp moral reasons, understand relevant moral principles, recognize when the application of those principles is called for, identify the appropriate targets of moral respect, and so on. The cognitive powers that Wallace has in mind seem to involve a fluency in moral principles. This will involve both knowing the accurate formulation and content of those principles – which we might call their grammar – and knowing which circumstances require which moral justifications – which we might call their context.

⁵ R.J. Wallace, *Responsibility and the Moral Sentiments*, p. 157 Other philosophers have offered similar conditions for agency. For more on approaches to this question, see Chapter 5’s section on excuse.

⁶ Wallace, p. 159.

Wallace suggests that this moral and deliberative skill is necessary for moral responsibility, and these skills are prerequisites for agency and moral assessment. He does not, however, believe they are sufficient. An agent who can be rightfully held accountable for her actions must also be able to translate her cognitive performance into action. Thus, Wallace classifies the other half of these powers as affective powers, the capacity to “regulate one’s behavior by the light of moral reasons.”⁷ These are powers of the will. To be accountable, an agent must not only have fluency in moral justifications and principles, she must also have the ability to form intentions appropriately from the results of her deliberative process. She must be able to will what she ought will. This involves some capacity to reflect on and regulate her performances, the ability to act. Without either of these powers, the cognitive or the affective, Wallace argues that it would not be fair to consider the person genuinely accountable. If this picture is right, then we have a starting point to consider standards of eligibility for moral assessment, and the beginning of a picture of appropriateness.

Once we establish eligibility for assessment, we must turn to the assessment process itself. All assessments are evaluations of sorts, the comparison of objects to a standard (or standards), and moral assessment is no exception. Thus, we have the following questions: what are we assessing, and why? The answers to these questions inform the process that assessments take. Some argue that the ultimate purpose of moral assessment is to form accurate judgments about the character of an agent. The moral assessment process would then be a method of grounding a claim that an agent has a vice or a “shortcoming” in character.⁸ This will, in a certain sense, be obviously true. If I morally assess the performance of an agent and determine that she deserves X amount of blame for her performance of a wrong action, then I have established

⁷ Wallace, 159.

⁸ For an example of this approach, see Robert Nozick in *Philosophical Explanations*.

that she is the kind of actor who has performed this morally wrong act. This is no doubt morally interesting, but extraordinarily limited in scope.

I mean by moral assessment something importantly different than character assessment. Accurate assessment of character – even of morally relevant character traits such as virtues or vices – will necessarily include much more than the performance of a particular wrong action. In Chapter 3 of this dissertation, I will consider the case of Jerry Sandusky, a man who was convicted of multiple counts of child rape. What does the moral assessment of these actions tell us about his character? It tells us something. The fact that this person deliberately harmed children in such a brutal way tells us that he had, at least on these occasions, a morally horrific intention and a willingness to act on it. But it does not tell us much more than that. Is he, in general, an honest person? One who handles anger appropriately? How did these intentions, and the deliberation they contained, fit within the larger context of his other dispositions? Without far more data, we simply don't know. Assessing the virtues of one's character will involve looking at the fruits of moral assessment for many, many of the agent's particular actions, and all of this data will be relevant to the character assessment process. But without a well-performed moral assessment of particular actions, we have no way to gather this data.

That said, there may be intentional actions that reveal nothing about an agent's character other than the fact that she performed that action once, under very particular circumstances. Literature on legal mitigation is filled with references to actions that are “out of character,” which is to say that they are not representative of the overall moral dispositions of the agent in question. It is not that the moral assessment of these out of character actions reveals *nothing* about the character of the agent, but by their nature, it is clear they reveal little that is generalizable. If the ultimate purpose of moral assessment was to simply gather additional data

for character assessment, the assessment of these peculiar out of character actions would not be as valuable as actions that revealed something deeper and more generally true of an agent.

People have offered many other theories about the purpose of moral assessment. Scanlon, for example, argues that moral assessment allows us to appropriately adjust the relationships we have with others,⁹ and Strawson flips the question around, arguing that our notion of moral assessment comes from what we do in relation to one another.¹⁰ Others still might argue that the purpose of moral assessment is to help us become better choosers in terms of our own actions, and in morally assessing ourselves and others, we reinforce our own moral education and improve our habits. These positions may capture something interesting about the usefulness of moral assessment, but these approaches fail to emphasize something basic and important. We might think there's a parallel to be made between moral assessment and the study of physics, for example. It's true that a deep understanding of physics allows us to be better engineers, to build structures with stronger and more efficient designs. But engineering is not the most basic reason we do physics. We investigate the truth of the physical world we live in, even when that truth doesn't have direct application, and we do such investigation because there is non-trivial value in finding out what is true in the world. We investigate physics because physics, in and of itself, is *worth* investigating.

In the discussion that follows I will be appealing to a very particular approach to moral assessment, an approach focused only on getting the theoretical relationship between the constitutive parts of action and the assessment of blame correct. The project I undertake in this dissertation is to increase the accuracy of moral assessment by exploring how and why mitigating circumstances alter the amount of blame that an agent deserves. This means the space

⁹ T.M. Scanlon, *Moral Dimensions*.

¹⁰ P.F. Strawson, "Freedom and Resentment."

that this project occupies is prior to and independent of these practical applications, depending on what is often called the “pointless grading” approach to moral assessment.¹¹ People use words like “pointless” to communicate a belief that doing ethics in this manner is, perhaps, fruitless. While I believe my project will be useful to people who have theories about the social meaning and function of blame, it is true that I am invoking a notion of what it is to deserve blame that is much more limited. I am not yet willing, however, to view this as a problem. Underneath the criticisms of the moral grading approach seems to lurk a theory that the only pursuits worth conducting are those that bring about some additional good result, independent of the pursuit itself. I am firmly committed to the belief that this is an inappropriate way to look at ethics (and no doubt countless other inquiries). I argue that we morally assess the status of certain action performances not because of the additional benefit that such knowledge might bring, but because the importance of the subject makes the assessment worth performing, independently.

As people constrained by limited hours in the day and limited days in our lives, our choice of where to focus our efforts is directed by the value and importance we place on those pursuits. We can argue about whether certain pursuits are worth the attention we give them, but one of the ways we acknowledge an important task is by giving it appropriate attention. The moral assessment process is valuable because it, in itself, is important. The truth of what actions are and are not morally permissible, how much blame an agent deserves for a given action, those are all important. They are important because they are the expression, in the deepest way, of who we are as moral creatures. The assessment of an agent’s action is an instantiation – for both the object and the assessor – of our unique qualities as moral agents. Moral assessment is one of the

¹¹ See Scanlon, *Moral Dimensions*, p 126 for the use of this phrasing, or Gary Watson making the same charge against viewing moral assessment processes as merely performing the work of a “moral clerk.” (Watson, *Responsibility and the Limits of Evil*, in *Agency and Answerability*, pp. 226-227.) For a more complete version of this approach, see Jonathon Glover, *Responsibility*, p 64. Glover’s picture of “moral accounting” has significant similarities to what I am trying to capture in my discussion of moral assessment.

ways we understand the truth of the world around us. This is a worthy pursuit, regardless of outcome.

It is worth mentioning an additional concern that I believe often motivates hostility towards the moral accounting approach. That concern focuses on a worry about entitlement and inappropriate judgment. When philosophers consider “pointless” moral accounting, there is at times a tendency to describe this approach as petty and tattletale-like, with the practical usefulness of moral assessment for relationships and the like serving to provide a justification for performing this otherwise unsavory task. This intuition seems to imply that the act of “keeping score” places one above and separate from the object of the assessment in a problematic way, and it implies that people who perform such “scorekeeping” are committing a wrong towards the agent they are “scoring.” I am sympathetic to this intuition. If I followed a colleague around all day shouting out grades as she made her way through the world (C+ for that omelet preparation! A- for that faculty meeting!), I *would* be wronging her in a significant way. In the first place, it is not my within my authority to evaluate her performance of her tasks. She did not consent to this form of relationship with me, and I am inserting myself into her life in ways that she rightfully rejects. But it is not the assessment of her tasks that is problematic here. It is the singling out of *her* that is inappropriate, along with what I do with those judgments (namely, publically call them out to her attention). There is an enormous difference between being invited to try someone’s omelet and noticing it is not the world’s best, and taking it upon oneself to become the city’s kitchen police.

It is morally unacceptable to become the ethics police. This is not because there is something inherently problematic about performing moral assessment, something that can be overcome by some further usefulness. It is wrong because policing without appropriate authority

is wrong. It is involving the singling out of people and involving them in power dynamics without their consent. But the “moral accounting” approach to moral assessment is not a form of policing, and the ledger is metaphorical. One does not involve an agent in an inappropriate, nonconsensual relationship by understanding the moral status of their action. Assessing an action is not performing an action *on* the agent. It is not saying she is less a person if she committed a wrong, any more than saying a ray of light is less a ray if it is visible on the blue spectrum. Understanding the status of the action is morally neutral.¹² This discussion is worth having, because focusing too much on the projected “purpose” of moral assessment can overshadow our discussion of the assessment process itself. When we don’t keep in mind the fact that the ultimate purpose behind the moral assessment process is to *get it right*, to pay respect to the importance of the pursuit, then we dishonor the fact that moral accounting is worthwhile in its own respect. Please note: I am in no way arguing that the fruits of the moral assessment process are not important and worth investigating in their own right, and if I have failed to convince you of the inherent value of understanding what makes a moral ledger accurate as a project worth investigating in its own right, well, then this dissertation will be useful as preliminary work for those concerned with the applications of an accurate moral ledger. But it my sincere hope that I have made a case that the approach I will take has value in itself. And now, it’s time to talk about mitigation.

¹² For a similar approach to the neutrality of moral assessment, including the assessment of blame, see Pamela Hieronymi’s “The Force and Fairness of Blame.”

II.

Many of us believe that mitigation involves some notion of reduction, or lessening of blame, and indeed, that this reduction is the foundational piece of what makes mitigation, mitigation. I will call this the common sense view of mitigation. If this picture is true, we will need some notion of moral quantification to make sense of it. Earlier in this chapter, I mentioned that my project is concerned with determining the amount of blame that an agent deserves for performing a particular (in this case, morally wrong) act. An accurate assessment of deserved blame depends on facts about both the action and the agent's moral status and obligations, and these facts will determine how much blame it is appropriate to ascribe to an agent. I will discuss facts about the agent and the context of her action in Chapter 3, but for now, I will raise a question about how to describe the boundaries and nature of a wrong action for the purpose of moral assessment. Such a description will require capturing both the agent's internal deliberation and how she instantiates her intention in the world. It is only once we have a handle on what the action *is* that we will be able to look at which of its features are morally relevant and how those features affect the assessment process.

When it comes to questions of moral assessment, many philosophers rightfully focus on the connection between action and the agent's reasons for acting, arguing that the reasons behind one's action determine whether that action has moral worth,¹³ or that what it is to be morally responsible is to be open to praise or blame based on how the agent settles the question of whether to ϕ ,¹⁴ or the notion that an action must be intentional in order to be morally assessed, where intentional means something like "done for reasons." These approaches all focus on the

¹³ This, for example, is one of the cornerstones of Kant's ethical theory as described in the *Groundwork*. Kant writes: "An action from duty has its moral worth not in the purpose to be attained by it but in the maxim in accordance with which it is decided upon," p. 13.

¹⁴ See Pamela Hieronymi's "The Will as Reason."

deliberative process that takes place prior to and during the act itself, and thus, if we are interested in assessing action, we need a method of action description that will allow us to capture the relevant features of this deliberation. In particular, we need a method of action description that will allow us to identify the morally-relevant reasons that will impact our assessment. Let's take a particular action: me pouring myself a cup of coffee. How should this action be described?

There are nearly infinite ways. I can say that I just poured myself a cup of coffee. I can say that I rose from my desk and poured myself a cup of coffee because I am a lover of coffee, and a friend just bought me a rare and delicious coffee blend. I can say I poured myself a cup of coffee because I enjoy the taste of this blend. I can say I poured myself this coffee because I want to try this batch before I next talk to my friend, so I can tell him how much I enjoyed it. I can say I poured this coffee because it's 2:30 in the afternoon and I could use a little pick-me-up. I can say I poured this coffee as an excuse to step away from the computer for a moment and rest my eyes. I can say my partner wanted to try this blend too and I promised to leave enough for him. I can say, while pouring, I may not have left enough for him to fill his mug (and then I felt guilt). I could go on.

When we teach Kantian ethics to our undergraduate students, one of the first things our students do is attribute extraordinarily simple motives to agents. Action X was done from the motive of greed. Action Y was done from the motive of sympathy. And for someone just dipping their toes in Kantian waters, this is a helpful approach to the complexity of ethics. But Kantian theories don't apply only to (flat!) fictional characters, and we agents are not so simple. Everything I wrote above describing my coffee-pouring is a true description of the deliberative piece of my action, and it is far from complete. The particulars of our deliberation differ from

moment to moment, and from action to action, and so if we were to give a complete description of the entire deliberative process as part of every action description – and if every piece of the deliberative process was relevant to action description and assessment – it would reflect the fact that the deliberation contained within every action is, in essence, different from almost every other action. If moral assessment of action depended on this complete description of action, we would have to start from scratch in approaching each action in order to assess it. But this isn't what we do, and moreover, it's not what we should do. Namely, we sort actions based on certain characteristics, pieces of description, according to our purposes. We use moral categories and sort actions accordingly. And since we are working with action for the purpose of moral assessment, we sort based on morally relevant characteristics.

Return your attention to the two murders that I introduced in Chapter 1. Let's say that the Betrayed Lover who murdered her partner noticed, during the act, that her partner – the murder victim-to-be – just got a haircut. This thought forced her to deliberate briefly, and in the way we sometimes reflect on the strange and ill-timed thoughts that come to mind, she considered whether or not to mention it.¹⁵ Ultimately, she chose to disregard it, focusing on her vengeance and the violent act at hand. That makes this bit of deliberation a part, albeit a small one, of the overall deliberative process that was contained in her act of murder. It also means it is true that a full and accurate description of the deliberative process of this action would include that she chose to murder instead of discussing her victim's hair. I suspect most of us wouldn't judge her

¹⁵ I want to be clear about the case I have in mind here. I am not claiming that a thought “popping into one's mind” necessarily makes it the case that the intruding thought is part of one's deliberation. That view might be true, or it might not be depending on our understanding of what constitutes deliberation. But I do think it is uncontroversial that our process of *addressing* those thoughts is part of our deliberative process. At that moment, the murderer is deliberating on the question of whether, and how, to act. She considers the addition of a claim about her victim's hair and decides against it. This internal decision is no doubt part of the complete description of her deliberative process leading to, and shaping the content of, her action.

more or less blameworthy for murdering her partner if she hadn't dismissed the thought, or if she had decided to mumble something about his hair as she attacked him. The thought was part of the entire deliberative story of her action, part of the particulars that make up a complete metaphysical description, but is it the kind of detail we need when we perform a moral assessment?

It might be suggested that this detail is just part of the fluff around the edges of deliberation, the chaff that exists as part of any particular deliberative process. We might think the extra details unique to this particular murder simply don't matter for the purposes of moral assessment, any more than the fact that the murder was committed at 1:52PM instead of 1:51PM, or with a vase instead of a bookend, or in the bedroom instead of the living room, matter to moral assessment. Perhaps it is the case that there are pieces of deliberation that, while part of the entire metaphysically complete action-description, do not matter for the purposes of moral assessment. But why should this be the case? It seems that if we are holding people accountable for the particular action they took, and the action is defined by the particulars of their deliberation, all of this deliberative detail will be part and parcel of making the action what it is, and so all of this deliberative detail will be relevant when it comes to moral assessment. We might try to suggest that this fluff around the edges of deliberation can be dismissed as somehow not relevant to assessment, but we cannot merely assert this position. Such a dismissal will require a principled theory of categorization that will tell us just which pieces of deliberation are relevant and which are fluff.

This problem becomes pressing once we move to the process of moral assessment. How to understand the boundaries of an action – getting to the appropriate action description – tells us how to understand just what the agent is doing, and consequently, what we are morally assessing.

Doing the “arithmetic” of blame assignment only works if we understand how much blame each potential action is “worth,” or to be more precise, how much blame the agent deserves based on the performance of this particular action. And while this is a problem for all moral assessment, the problem of appropriate action description affects the mechanics of mitigation in a very foundational way.

If we are to consider the complete deliberative picture as being part and parcel of the boundaries of an action, then the notion of mitigation as a kind of reduction or lessening becomes, at best, problematic. We can see this more clearly with our two murders. Here we have two women, one of whom committed murder out of anger at her lover's betrayal, the other who committed murder out of boredom and a desire for financial gain. We might think that the amount of blame earned by the performance of “murder” is X, and while the Calculating Lover is blameworthy for the full X amount, the Betrayed Lover’s circumstances make it such that she is now deserving of X-5 (for example) units of blame. On this, the common sense view, mitigation is a subtraction trigger, and our moral assessment might go like this: determine how much blame an agent deserves for a particular moral wrong with no special circumstances. Assess whether she does indeed have special circumstances. Determine the subtractive weight of those special circumstances. And then simply do the arithmetic. The result is the mitigated amount of blame appropriate to assign to the agent.

But is this performing moral accounting properly? If approaching moral assessment at the level of the complete description of deliberation is correct, then mitigating circumstances don't actually reduce anything; they simply mark the presence of an entirely different act. The Betrayed Lover didn't do something equally morally bad as the Calculating Lover. She deserves less blame because she is fully blameworthy for a less-bad act, and not because she is less than

fully deserving of blame for an equally bad act. If this is true, critics of the common sense picture of mitigation might argue that we are making a very basic mistake. What we often take to be a kind of blame reduction, mitigation in common usage, is nothing other than mistaking one action for another, and mitigating circumstances, as we often think of them, don't actually do what we think they do.

It does seem intuitive to think that the woman who murdered her lover in a fit of rage is doing something very different than the one who acted out of greed, even if their actions are both properly classed as murders. One might argue that we should understand these actions as members of two different moral categories altogether, one an example of “hot” murder and the other an example of “cool” murder.¹⁶ It is true that the betrayed lover committed what we colloquially term a “crime of passion” while the premeditated murder did something with advanced calculation. But the argument still holds with any two actions within the same moral category. Take the cold, Calculating Lover case I described earlier, where the woman poisoned her lover out of a desire for financial gain. Consider two versions of this case. In the first case, the woman went out of her way to find the quickest, most painless poison she could get her hands on. She fully intended to murder her lover, but she chose not to make him suffer. Contrast that with a version of this case where the murder did no such research, not caring one way or another if her poison method caused pain or anguish for her victim. Both of these are clear cases of the genus “cold-blooded murder,” and the question remains: how should we morally assess them? Should we think of the amount of blame justified by the performance of “cold-blooded murder” as Y, assess the non-caring murder as being deserving of Y amount of blame, and say that the thoughtful murder is deserving of Y-10, where 10 is the amount of reduction/mitigation she deserves for trying to eliminate her victim’s suffering? Or should we say that the non-caring

¹⁶ For more discussion on this terminology, see Chapter 4.

murderer deserves Y amount of blame for her particular action, inclusive of her particular deliberation, and the thoughtful murderer deserves Z amount of blame for her particular action, inclusive of her particular deliberation? These two actions are different in their particulars. Should we think they are therefore different kinds of wrongs, where one wrong is worth Y amount of blame and the other is worth Z amount of blame?

If we are to understand these two murders as totally different kinds wrongs, with different amounts of blame attached, it seems there is no place for a “reduction” to enter the picture. It would be silly to say that a rabbit is a “reduced” elephant, even though they are both mammals; it misunderstands what it is to be a rabbit and an elephant. If this critic is correct, and action and blame assessment work in the same way, then saying that the betrayed lover's circumstances are “mitigating” can't mean that they serve to reduce deserved blame. They are just the circumstances that mark the presence of an entirely different kind of action -- a moral rabbit, not a moral elephant. Since the details in their deliberative processes are different, we are left to assess them as entirely different kinds of actions. And as different kinds of actions, they should be assessed differently. This eliminates the common sense view of mitigation.

This picture seems to be what Wallace has in mind when he describes cases of duress and coercion.¹⁷ Let me return to the coercion case I described in the introduction. Say that this agent – call her the Secret-Spiller – is threatened with death if she doesn't turn over the names and locations of the political refugees. I will stipulate that turning over this information is morally impermissible. The agent made a pledge not to, the release of the information puts innocent people in harm's way, and so on. She deliberates, and with the coercer's threat as her reason for acting, she releases the information.

¹⁷ Wallace, p. 146.

I've chosen to describe this case using the hallmark features of coercion. In this case, the Secret-Spiller's motive stemmed directly from the threat issued by a third party, a threat issued with the clear intent of influencing her actions. And in this case, she did indeed perform the morally wrong action.¹⁸ By any definition, however, she was coerced into doing it. Cases of coercion are particularly interesting because they are defined by a set of circumstances that influence deliberation in a particular way. That is to say, when we speak of a coerced act, part of what we mean by the label is that the coerced agent's motive for performing the action stems in a meaningful way from a threat made by the coercer, and that the coercer delivered this threat for the purposes of controlling the coerced agent's behavior.¹⁹ These unique features of the agent's deliberation are necessary conditions for qualifying as a case of coercion.

In Chapter 5 of this work, I will discuss Wallace's picture of justification and excuse in greater detail. For now, the important part to mention is that Wallace claims that his notion of excuse also speaks to cases of duress and coercion. To hold the Secret-Spiller blameworthy for the attitudes normally reflected by the action "turn over information about political refugees" would be holding her accountable for a choice that she didn't actually make, and on Wallace's account, that is simply not fair.²⁰ Wallace would argue that she is *not at all* blameworthy for turning over this information, because that's not the action she actually chose. She instead

¹⁸ It is contentious whether cases of coercion require that the coercee perform the wrong action, or whether the presence of the threat is sufficient. I intend to leave this question aside and focus on the most non-controversial cases, since the scope of coercion will only matter if coercion is a category that might have an impact on moral assessment. It is this latter issue I will be taking up in this chapter.

¹⁹ The full scope of what is and is not coercion is interesting and worth its own discussion, but for my purposes, I aim to consider a non-controversial case where someone committed an action that is indeed morally wrong. There might well be cases where coercion justifies the performance of an action, making the agent not deserving of blame at all. I am not sure if these cases will end up qualifying as genuine coercion, but for my purposes, I set them aside and focus on cases where we stipulate that a wrong has occurred. For more on the scope and nature of coercion, see Nozick's "Coercion," Frankfurt's "Coercion and Moral Responsibility," O'Neill's "Which are the Offers You, Can't Refuse?" and Ripstein's "Authority and Coercion."

²⁰ Whether it is accurate and appropriate to use the notion of "fairness" here is controversial. See Hieronymi's "Force and Fairness of Blame" for more discussion on this subject. Regardless of Wallace's terminology, however, we can at minimum understand him to be saying that it would be inappropriate to hold the agent blameworthy for the bald action "turn over state secrets" without any additional modifying description.

intentionally performed the action x-rather-than-y, or to be more specific, turn-over-information-to-avoid-death. The second action reveals a very different choice than the first. In other words, Wallace argues that coercive circumstances are excusing (or depending on the case, mitigating) because they change the shape of the action that the agent is actually performing. The presence of this particular deliberation-altering circumstance makes it such that the agent is not performing x, she is performing x-rather-than-y. The amount of blame appropriately attached to each action is different, because they are entirely different actions.

With this picture, Wallace has a view of coercion that is consistent with the objection to the common sense view that I describe above. The coercive circumstances that serve as excuse or mitigation don't do so by reducing or eliminating the amount of blame that the agent deserves. They work by altering the nature of the action, particularly by changing the shape of the deliberation. We might think that on Wallace's view, excusing circumstances merely block the assignment of blame to an agent for an action that she did not choose. So in the case of the Secret-Spiller, the presence of coercion makes it inappropriate to hold her blameworthy for the simple act of turning over information. She is still fully blameworthy for her action x-rather-than-y (in this case, turning-over-information-to-avoid-death). And if this second action violates a moral obligation, then she has no excusing condition to inhibit her blameworthiness. Wallace seems to think that excuses of this kind block blame assignment for an incompletely described action but leave the agent fully open to blame assignment for the accurately described action. This latter action may or may not turn out to be morally acceptable; the mere presence of coercive force or duress will not automatically excuse the assignment of blame for any x-rather-than-y decisions. It is not entirely clear whether turning over such sensitive information to avoid death does or does not actually violate a moral obligation, but it is clear that an action like kill-

an-innocent-bystander-to-avoid-being-tickled isn't going to be morally acceptable, even if it is a case of (extremely mild) coercion. But the redescribed -- the more fully described! -- action will still result in a unique moral assessment because it is properly considered a different action.

We might think: why then is this not true for *all* action? As described, we might think that cases of coercion are nothing more than cases where we get the action description, at first glance, wrong. Once our action description is accurate, our moral assessments go back to functioning like normal. This is no more interesting than any other kind of mistake in assessment. I thought that agent murdered that man. I was wrong. It was only an accident. Assessment adjusted. I just had the action description wrong. If this is the view that Wallace is pushing, we are left with a problem for the common sense view of mitigation. Yes, cases of coercion, by definition, include deliberation of the sort "should I do x-rather-than-y?", but if this is the feature that forces us to adjust our assessment by readjusting our description of the action, how does this not affect all cases where at first glance we thought actions were morally similar, until deeper investigation reveals the differences in deliberative particulars? If the particulars of deliberation change the coerced agent's action into an entirely different sort of action, why not think that this is true for all circumstances that affect deliberation? When we think about mitigating circumstances, why not think that we're *not* looking at a reduction in blame deserved, but instead the deserving of full blame for an entirely different act?

I have so far approached this problem from the angle of moral mitigation, but the problem of action description and moral assessment cuts deeper. After all, the vast majority of human actions involve complex and nuanced deliberation, and it will almost never be the case that the deliberation behind an act of murder, for example, will look as simple as answering the question "do I intend to murder?" in the affirmative. There are countless complexities to our

deliberative and intention-forming processes. If accurate moral assessment of coercion cases requires us to acknowledge that the coerced agent is doing x-rather-than-y instead of merely x, then why do we need not acknowledge that she is really doing x-rather-than-z-rather-than-a-rather-than-b-rather-than...? And why doesn't this acknowledgment become mandatory for all actions, not just for cases of coercion? If the presence of coercion alters moral assessment in the way described above, that is, it indicates the presence of a different-than-previously-identified act, then it is really just an instance of a much larger problem. If we must point to the particulars of deliberation to accurately describe the action for moral assessment, then we don't yet have a principled reason to stop at the form "x-rather-than-y". As it stands, we would need a principled reason to stop anywhere before the accurate and complete description of the deliberation underlying the formation of the intention, the accurate description of the fullness of the agent's choice. Without such a principled reason to stop elsewhere in the process, we're faced with a regress back to the complete action-description – at least, of the complete shape of the deliberative process – in order to perform accurate moral assessment. And to paraphrase Carl Sagan, this has us recreating the universe to bake an apple pie from scratch.²¹

Here is where this discussion leaves us. We have a picture of what the moral assessment looks like: the assessment of the actor and the status and context of her agency, and the assessment of the action performed. But the assessment of the action depends on determining what the action *is*, since we cannot determine how much blame the agent deserves for a particular act if we don't understand the nature of the act performed. I have introduced two ways of approaching this assessment of action. The first possibility follows from Wallace's treatment of coercion. On this view, we are holding agents blameworthy for the fully described action they committed, x-rather-than-y-rather-than-z and so on. On the second view, the common sense

²¹ *Cosmos*, p. 218.

view, we are saying that for some actions, the appropriate circumstances will allow us to say that they are members of a moral category, and yet they do not justify the assignment of blame normally associated with members of this moral category. These circumstances serve to reduce blame. They serve to mitigate.

The difference between these two pictures concerns the fundamental nature of mitigation. If indeed the common sense picture is correct, we will need to be able to give a theory that will explain why it is appropriate to first talk about actions in moral groupings, and second speak about some of those actions as being morally better (justifying the assignment of less blame) or morally worse (justifying the assignment of more blame) particulars of their category.

III.

In the last section, I left off with questions about the common sense view of mitigation, including the question of whether the moral assessment process has room for a notion of mitigated (where mitigation means “deserves reduced amounts of blame) cases of premeditated murder, for example, and if so, how to understand the notion of reduction in this context. In the following sections, I aim to answer those questions, arguing that we have resources that can make sense of a notion of blame reduction, and thus, it is both possible and advantageous to save the common-sense view of mitigation.

The concern with the common sense view of mitigation revolves around action description for the purposes of moral assessment. In particular, the concern is whether we should assess actions as individual particulars independently of moral type, or whether we can understand actions as members of moral categories first, individuals second. The first approach requires us to assess each x-rather-than-y-rather-than-z deliberation as its own kind of action, and

the second allows us to compare members of a certain moral category as particular tokens of a type. If we are to defend the common sense notion of mitigation, we will need to appeal to a theory of moral categories that can justify the second approach to assessment.

If there are almost infinite ways to describe action, then there are almost infinite ways to divide action into categories. It is possible to slice action by focusing on almost any part of action description, but my goal is to offer a plausible way to divide action for the purposes of moral assessment, to carve principled types out of the whole which would allow for meaningful moral assessment at the level of type instead of at the level of token. My goal is not to argue that all murders, for example, will look the same when they are morally assessed. There will no doubt be a difference in the blame deserved by different agents who commit different murders. But my goal is to suggest that there is something to be gained in the process of moral assessment by classifying something as a murder, and thus we have some reason to carve up action description accordingly, instead of merely evaluating all the particular details of this murder or that murder.

Now, it will certainly be true that the sorting of action into moral types will make our moral assessment easier. After all, if I know something about the moral type “murder” and I know that this action is accurately classed as a token of that type, then I know some things relevant to blame and blameworthiness before I go any further in my assessment. I have a shortcut, of sorts. But beyond ease in application, I will argue that the category-based approach to moral assessment that I will defend below respects the nature of the moral assessment process. I will offer a system of categorization that is based on the obligations that bind the agent, an approach to moral assessment that I describe as moving from the outside-in. I will argue that this system of categorization is effectively grounded in the nature and purpose of moral assessment, and that this approach will shed light on the nature of mitigation.

At the most basic level, a morally wrong action is wrong because it violates some obligation that binds the agent. The content of these obligations will vary from ethical theory to ethical theory, as will theories of justification for these obligations. But there are some noncontroversial cases. We have an obligation not to torture other people for fun, let's say, or an obligation to come to someone's aid when they are in great need and it costs us very little. If I choose to kick my colleague Red for entertainment, my action is morally wrong, at its base, just because I have an obligation not to kick Red for my amusement. Whatever the theory that grounds it, the command "don't kick Red for fun" binds me. It's worth noting here, however, that the content of that obligation includes and depends upon facts about both the subject and the object of the obligation. I have an obligation to refrain from kicking Red, but that obligation does not extend to things relevantly different from Red, like this stuffed animal. In fact, if I took this stuffed animal and kicked it until its stuffing fell out, I am not violating a moral obligation at all. So it is not the kicking that is the moral problem; it's the kicking of Red, the kicking of someone who is the appropriate object of a moral obligation not to kick for fun.

If I were to start categorizing morally wrong actions based on the obligations they violated, I very well might be vulnerable to the same kind of regress problems we started with. Our obligations seem particular in form. If I promise to meet Red for lunch, then I have an obligation to Red, and that obligation doesn't exist without the particular circumstances that created it. So if I am trying to sort morally wrong actions by the obligations they violate, and obligations exist because of particular facts in the world, between real and individual agents, I am left with a problem. It seems that sorting actions based on which obligations they violate merely pushes the problem back a step, and now, instead of having to describe an action "from scratch" in order to morally assess it, I have to describe the particular obligation "from scratch"

instead. This method of categorization doesn't add anything interesting to the process of moral assessment, nor does it make any advancement for the common sense view of mitigation.

But this is not quite the right way to think of obligations. While obligations are created by relationships and circumstances, they are not dependent on particular arbitrary features contained within the instantiations of these relationships and circumstances. Fundamentally, the obligations we have towards one another will depend on the relationships between us. I have certain moral obligations towards my husband that I don't have to a stranger down the street, others I have to my students but not my husband, and still others I have towards all of my fellow moral peers. The moral relationship I stand in with my husband dictates certain particular obligations that bind me; I've made vows to him that I will be blameworthy if I break. But the fact that I made those vows on a certain date, to a particular person, is not foundational to their nature as obligations. In fact, if I had made the same vows in the same place at the same time to a different person, I would still be bound in the very same way, just directed at a different moral object. I could counterfactually stand in this moral relationship with countless other people instead of the one I am with, and still be bound in the same way.

We might think of a moral obligation as having a rope tied from our wrist to another person's wrist. It is true that this particular rope is tying our particular wrists. But this rope, with these knots, could be tying someone else's wrists instead. The form of the rope connecting two strangers bound by a promise would be exactly like ours. Making a promise to someone creates a moral obligation because both the agent, in this case me, and the object of this obligation, the person I make a promise towards, have a certain more status, and it is this status that allows her to ground my moral obligation, and me to be accordingly bound. It is certainly true that a particular obligation will be created by a particular promise, for example, to a particular person.

But the nature of obligations makes it such that in a counterfactual universe, anyone with the same moral status could have grounded the obligation that arises when I make a promise, and anyone else counterfactually in my circumstances would be bound in the same way as I am when promising.

This makes moral obligations importantly immune to certain kinds of concerns about regression to particulars. They will be grounded in certain circumstances – I don't have obligations of loyalty to someone I did not promise them to, for example – but those circumstances do not include the fact that the person I happened to make the promise to was this person. They couldn't, because if moral obligations functioned in such a way, we would be stuck with outcomes like "It is a moral wrong to break a promise to Red, but it is not a moral wrong to break a promise to Green, all other things equal." We find outcomes like this ridiculous because we recognize that Red and Green (and all other members of the moral community) are importantly equal, due the same fundamental respects as agents and sources of moral obligation. To treat them as not equal in this way misunderstands the nature of what grounds ethical obligations. The nature of obligations lends them to a basic type/token distinction. We have intelligible types of obligations that would bind any agents who occupy the relevant circumstances, and we have particular circumstances between particular agents that create tokens of these types.

Here's where that leaves us: since we can understand the nature of moral obligations independently of the particular objects that ground those obligations (that is, we can say that the same obligation would hold for Red, as it would for Green, as it would for all other agents in the relevant circumstances), it is possible to offer a system of categorizing moral obligations, based on the relationships and circumstances involved. But there is an additional benefit to approaching

moral assessment in this way. Moral assessment cannot happen in a vacuum. All actions take place in a moral landscape, and the very same action, with the very same deliberative shape, should be morally assessed differently if the agent's circumstances are relevantly different. If I spend my morning helping my neighbor in her garden because working in the ground is rough on her joints, I am performing a kind act, not deserving of any blame. If I perform the same action against the context of a standing promise (let's say, I had promised to donate bone marrow that afternoon to save my sick relative), I am now deserving of blame for a violation of a promise. My deliberation in these versions of the act can look identical, but in one case, I am bound by an obligation that I am neglecting to include in my deliberation (and I am blameworthy for that omission). It is not just the boundaries of our deliberation and action that matter to moral assessment. It is the backdrop of obligations that this deliberation is situated within. Approaching the moral assessment process from the perspective of binding obligations captures this feature of moral assessment and allows us to assess the action in its proper context. This is why I call this approach to moral assessment from the outside in. Facts about the agent's deliberation and intention-formation process can only be properly assessed if we first understand how the agent is situated in the world, the moral context that this action exists within. Without this perspective from the outside, looking at the obligations that bind the agent as she makes her decisions, we will not be able to properly morally assess agents who fail to deliberate at all about what they should. Certain types of failures will only come to light when approached in this way.

The detail work of the sorting will no doubt be messy and difficult, but I can make some claims, hopefully, without invoking controversy. We might say that we have an obligation not to exploit our moral peers, for example. We might have an obligation not to break our promises. We might have an obligation to aid those that are in dire need when it costs us little. And so on.

And because I can understand obligations independently of particular instantiations, grounded by circumstances that could conceptually exist between any two agents of the proper moral status, I can start to morally categorize action as well. To be specific, I can begin to speak of a category of action as the collection of actions that violate the moral obligation not to exploit, the collection of actions that violate the moral obligation to aid, and so on. And with this step, we can make some genuine progress in moral assessment and towards defending the common sense view of mitigation.

Earlier, I discussed a problem for the common sense view of mitigation revolving around the choice of action description for the purposes of moral assessment. The problem took the following form: if we assess people as deserving of blame because of the deliberative part of their actions, and the deliberative part of any action includes countless particulars, minutiae that force us to look at the entire deliberative process in a descriptively complete sense, then moral assessment would not allow room for the notion of reduction. If this is true, we cannot compare apples to apples (or murders to murders), and we cannot say that some circumstances *reduce* the amount of blame for a member of a moral category.

With the outside-in approach, we do not begin the moral assessment by looking at the action in isolation. We start by looking at which obligations bind the agent at the time of the morally wrong act and filter the pieces of her action through this lens. This approach allows us to focus only on the pieces of deliberation relevant to the violation of obligations, sorting actions into categories according to which moral obligations they violate (or fulfill). This approach reorients our moral assessment project. Instead of inventorying all the cars on the road to get the count of blue ones, we simply count up the blue and forget the rest.²² To put this back into moral

²² Julie Tannenbaum raised an excellent question about whether such a system of assessment really does have any advantage on the practical level. After all, if one is looking for only blue cars, doesn't one still need to examine each

terms, an alternative assessment process might look like Wallace's take on cases of coercion. As I discussed earlier, Wallace argues that coercion cases are fully excusing, because the agent is not actually doing *x at all*, she is doing *x-rather-than-y*. If this is true, then there is no reason we should stop at "*x-rather-than-y*" instead of "*x-rather-than-y-rather-than-z*," and so on, and no reason to think this affects only cases of coercion. Because a significant number of our actions contain deliberative content that looks like "*murder-instead-of-therapy*" or "*breaking-a-promise-instead-of-going-on-that-awkward-lunchdate*," if we apply Wallace's reasoning about coercion, performing "*murder-instead-of-therapy*" is a fully excusing condition for the moral assessment of "*murder*," since those actions should be rightfully considered two entirely separate moral kinds. This leaves no room for the common sense view of moral assessment.

In contrast, starting the moral assessment process by first evaluating which obligations bind the agent, sorting the actions into categories according to which obligations they violate, and then (and only then) assessing the action helps us preserve the common sense view of mitigation. It lets us compare moral wrongs within a given category, and makes sense of the concept of reduction. Beyond this, however, this method of moral assessment gets to the root of what we're doing when we perform a moral assessment. It is not accidental that blameworthy actions violate moral obligations – that feature is what *makes* them wrong. This categorization gives us context for understanding the source of moral wrongness. This advantage will become

car to check if it is blue? In other words, does performing moral assessment from the outside-in actually reduce the need to check the status of all the deliberative particulars contained within an action? My response is that the outside-in approach to moral assessment will still require a sorting of deliberative particulars into the classes of morally relevant (that is, bearing on the fulfillment or lack of fulfillment of the obligations that bind the agent) and not morally relevant, and this sorting will indeed require looking at all the deliberative particulars. The advantage, however, comes after this point. Once we establish that a particular piece of deliberation is not morally relevant, no further consideration of that piece is required. Metaphorically, we might say that once I sort the cars into "blue" and "not-blue," I no longer have to worry about all the color distinctions I could make within the category "not-blue." This approach to moral assessment is really just one way to capture the assessment of all and only morally relevant features of action while setting aside features of action that do not have moral bearing.

clear when we attempt to morally assess a complicated case. Let's take a case that is dominating media coverage as I write this: the case of a Gerald Arthur "Jerry" Sandusky.²³

Sandusky was an assistant football coach for Pennsylvania State University for over thirty years. In the late 70's, Sandusky founded a charity that focused on helping Pennsylvania's underprivileged and at-risk youth populations, with a special focus on foster children. The charity, called The Second Mile, became hugely successful, in large part because of Sandusky's continued association with Penn State's popular football team.

Sandusky's retirement from coaching came soon after campus police began an investigation of charges of inappropriate behavior, prompted by a mother's complaint that Sandusky sexually abused her child in the Penn State showers. In 2002, a graduate assistant walked in on Sandusky raping a young boy in showers of the Penn State athletic facilities. In 2009, a grand jury was convened, and in 2011, they indicted Sandusky for sex crimes against children. In 2012, Sandusky was convicted of 45 of the 48 charges against him. As of this writing, Sandusky is awaiting sentencing, but given Pennsylvania's minimum sentencing laws, he will certainly spend the rest of his life in prison.

Public opinion towards Sandusky's crimes is dramatic. In particular, there seems to be a general consensus that Sandusky's actions were *especially* bad, even for a child molester. When I asked Pittsburgh-area residents²⁴ whether we should consider Sandusky morally worse – whatever that will amount to – than someone who snatches children from a shopping mall and sexually abuses them, my conversationalists uniformly agreed that Sandusky's abuse was worse.

²³ The details I describe about this case come from the following sources: Viera, Mark and Jo Becker (November 14, 2011). "Ex-Coach Denies Charges Amid New Accusations". The New York Times. Report: Former coach Jerry Sandusky used charity to molest kids (Harrisburg Patriot-News, November 6, 2011). Freeh Sporkin & Sullivan, LLP (2012-07-12). "Report of the Special Investigative Counsel Regarding the Actions of The Pennsylvania State University Related to the Child Sexual Abuse Committed by Gerald A. Sandusky."

²⁴ Personal interviews, June 19-23, 2012.

They pointed out that Sandusky met his victims through the charity that he founded. They pointed out that the victims were primarily foster children, or other children in especially vulnerable situations, and this made them easier prey for Sandusky. They point out that Sandusky used his connections with both the charity and the university to enable his hunt for new victims. They argue: these circumstances make Sandusky's moral wrongs worse than the shopping mall predator.

How might we understand this intuition in terms of moral assessment? Yes, the circumstances of the abuse are different in these cases. Sandusky gained access to his victims through his work with the charity, and the shopping mall predator gained access to his victims through knowledge about where isolated children can be found at a shopping complex. The acts are different in those ways. But how can we understand the claim that Sandusky's actions are *worse*? The wrongness of abusing children is not dependent on the circumstances of access to the children. Each of these two instances of child abuse is *maximally* bad. It simply doesn't make sense to say that one of these two acts of premeditated child rape is morally worse, deserving of more blame, than the other act of premeditated child rape – it is insulting to even try to make this argument. But by looking first to the obligations that bind the agent, and then to the pieces of action relevant to these obligations, we can start to make sense of the intuition that Sandusky's actions do indeed earn more blame than the shopping mall predator. And in that sense, we can say his actions were morally worse.

If we morally assess Sandusky in the manner I am suggesting, we will see that he is bound by several obligations to the children he worked with. First, yes, he is bound by an obligation not to harm those children, not to abuse them, in sexual ways or otherwise, and he is bound by the same obligation to all children (and adults) that he stands in any relation with. But

this is not the only obligation that binds him. In choosing to start a charity focusing on especially vulnerable children, and in choosing to relate to those children as a mentor and protector, Sandusky made an additional commitment to the children in his care. We might think, not only is he bound by the obligation not to harm those children – the obligation he has not to harm *any* children – he is now also bound by the positive obligation to promote the well-being and flourishing of the children under his care. These children entered a relationship with him based on this promise, and this moral obligation stands at the core of their relationship.

If we start our moral assessment from the outside-in, using the moral category approach, we can see that Sandusky had two strong moral obligations towards the children he abused, one not to harm and the other to proactively promote their flourishing. With these obligations in mind, we can look to the particulars of his acts. He raped several children under the age of twelve. This is a maximal violation of the obligation not to harm those children, and for the purposes of moral assessment, we will focus on the deliberation relevant to those actions. But he also broke his commitment to those children (and the parents that placed their trust in him). If Sandusky used his position as a mentor to gain access to vulnerable children for the purposes of abusing them, then this same action will violate both of the obligations that bound him.

This is not the case for the shopping mall predator. This abuser is bound by the same general obligation that bound Sandusky, the moral duty not to harm or abuse any children (or adults). But he lacks the second positive obligation to the particular children he abuses. He does not stand in a special relation to them, and he has not promised to keep them safe, or to promote their well-being and flourishing. So when we look at the assessment of blame, both Sandusky and the shopping mall predator will deserve full blame for the violation of the general moral obligation not to abuse children, and that makes each of them deserve some X amount of blame.

But Sandusky's action also violates another obligation that he does not share with the shopping mall predator. Sandusky is also deserving of full blame for his broken promises of protection and the promotion of flourishing of his victims. In his one act of raping that ten year old boy in the Penn State showers, he deserves some X amount of blame for the violation of abusing the child, and some Y amount of blame for betraying the trust of the boy and the boy's parent, and for failing to promote the well-being of the child. Both the shopping mall predator and Sandusky are maximally deserving of blame for child abuse, but only Sandusky for the second violation. The shopping mall predator is blameworthy X , then, and Sandusky is blameworthy $X+Y$, where $X < X+Y$. This does not mean that Sandusky's violation of the obligation not to harm children is worse than the other man's violation, but because of the particular obligations that bound him and the particulars of his action, Sandusky deserves more blame (because his one action makes him blameworthy for two violations of moral obligation). This is the sense in which we can say that Sandusky's action is *worse* than the shopping mall predator's.

That said, there are going to be cases where we want to use the concept of "worse" in a different way. When we discuss the Betrayed Lover and the Calculating Lover from Chapter 1, many of us still have the intuition that the latter is worse than the former in the sense that the agent in the latter deserves more blame than the former. And we believe that this is true because it seems there is something mitigating about the circumstances of the former that serve to reduce blameworthiness. Earlier, I called this kind of intuition an example of the common sense picture of mitigation, and it is time to explore what role mitigation plays in the moral assessment process.

CHAPTER 3: MORAL ASSESSMENT AND DELIBERATIVE IRREGULARITIES

The focus of this dissertation is mitigation of blame, and the story of mitigation will ultimately be a story about moral assessment. We worry about the mitigating force of certain circumstances not when we ourselves consider how to act, but when we're morally assessing actions already – or hypothetically – taken. Discovering where in this process mitigation is located will require looking closely at the pieces of moral assessment, and investigating each in turn. In this chapter, I will focus on the moral assessment process and the location of mitigation. This will involve looking at how the assessment process should go, and which factors will be relevant to the provisioning of blame. Finally, I will introduce my theory of mitigation's location and nature, making the case that this theory can both explain and justify our common intuitions about mitigating circumstances.

I.

In Chapter 2, I discussed preconditions for agency, particularly Wallace's notion of cognitive and affective powers. We will refer to these conditions of agency (or a similar formulation of these conditions) when we perform moral assessment. The moral assessment process can be broken down into parts. The first part focuses on the status of the actor's agency. Does she have a global or local condition impairing her agency? Is she an appropriate target of moral assessment at this time (or at all)? Is she overcome by mental illness, a child, or drunk? The answers to these questions might cause us to stop, or adjust, the assessment here. If the actor doesn't meet the criteria for agency, either generally or at the relevant moment, then no further moral assessment is appropriate. If someone does not meet the criteria for agency, we can evaluate her action (if it is even accurate to call her performance action) in all kinds of ways, but

the evaluation won't be a moral one. This is also the stage where we locate what Wallace calls powers of reflective self-control. The first part of moral assessment is fundamentally an evaluation of the agent's status qua agent. Here, we first look for criteria of eligibility. Does the actor meet the minimum standards for agency? Is this the right type of creature to morally assess at all? There will be times we find criteria that exempt the person from moral assessment. The person might be asleep, for example, or in a coma. She might be an infant, or overwhelmingly intoxicated. These people don't meet the necessary eligibility conditions for agency.

But this is not where we stop the assessment of the agent. There may be circumstances that affect the status of the actor's agency, without eliminating it. These are circumstances that impair the agent's ability to perform qua agent, and thus they might affect the amount of blame appropriately assigned to the agent. These circumstances range from mild intoxication to passionate cravings, and beyond. The feature they have in common is that they impair either cognitive or affective powers, and at this stage in moral assessment, those impairments might play a role in moral assessment.²⁵ But this does not yet say what kinds of circumstances we are looking for. Wallace's notion of affective and cognitive powers are meant to give us necessary conditions for agency, but the project of mitigation requires we look at these conditions in a more nuanced way.

But there is another part to this stage of moral assessment, and it is often overlooked. Evaluation of the agent qua agent does not only concern the particulars of the agent's capacities. It also concerns the context of the actor's agency. The hallmark of moral action is that it is action performed relative to obligation. Before we get to the stage in assessment where we evaluate what it is the agent does, the particulars of her action, we must first consider what the context of her action demands. Agency is not merely a matter of capacities. It is also a matter of being

²⁵ See Chapter 4 for an extended discussion on impairments.

situated in a particular way with morally relevant objects, both fellow agents and those whom we owe something towards. And so beyond what capacities the agent possesses, we must know too what burdens bind the agent in question.

Looking at the agent's capacities and acting context is only helpful, however, if we know what we're looking for. As I mentioned earlier, moral assessment demands we look for the necessary conditions of agency, in Wallace's terms, cognitive and affective powers. And we look to see what obligations bind the agent. But for the accurate and nuanced moral assessment that mitigation requires, we must look for something else. I call these features deliberative irregularities, and they can indicate grounds for mitigation.

II.

To understand the role and force of deliberative irregularities, we need to return to the question of what justifies the assessment that an agent *deserves* some amount of blame. If I am claiming that blame is the result of desert, then I am claiming that there is some fact about the agent and her action that stands as the bases for deserving this blame.²⁶ This basis, or fact about the agent and her relation to her act, is what we aim to accurately capture with moral assessment. For the purposes of moral assessment, part of the story that grounds desert of blame is how (and in what way) the agent violated a moral obligation towards a moral target. But in order for a common sense notion of mitigation to make sense in the context of moral assessment, we need to understand how someone can deserve more or less blame for a violation of a given moral obligation. Sandusky is an example of a case where a single action resulted in additive blame; he deserved the full amount of blame for violating two different moral obligations, and thus, for his one action, he deserved more blame than someone whose action only violated one obligation.

²⁶ For more on the notion of desert, see: Joel Feinberg, "Justice and Personal Desert."

But the concept of mitigation forces us to make clear the reverse kind of case: someone who is deserving of less than full blame for the violation of a single moral obligation. Understanding how this works will require a discussion of what conditions make it the case that an agent *deserves* blame in the first place.

At its core, the fulfillment of a moral obligation requires a regulation of one's actions (both at the level of deliberation and at the level of performance) out of acknowledgement of the status of the object of the obligation. When done successfully, this acknowledgment is in itself built into action – it requires active and intentional choice on the part of the agent – and thus its genuine performance can never happen by accident, or come from creatures who aren't capable of intentional action. This is what conditions like Wallace's cognitive and affective powers are meant to exclude: creatures that are either locally or globally incapable of the kind of action regulation that is constitutive of fulfilling moral obligations. Moral acknowledgement is often spoken about in terms of respect, and ethicists have parsed the content of respect in various ways. Different philosophers also have different positions on what grounds this demand for acknowledgment, but whatever the foundation, moral obligation is a way of relating to the object of obligation, requiring us to take the status (and the demands that flow from that status) of moral peers as reasons-generating, including it in our action formation and performance in the appropriate way.

When it comes to moral acknowledgement, there are many ways to go wrong. Respecting one's moral peers requires the regulation of deliberation and action formation, but this regulation is not simply on or off. If we think there is one way to meet the demands of our obligations, there are countless ways to fail. One can fail to show appropriate respect by showing respect in the improper amount (if I owe you \$10, and I repay you \$5, I am acknowledging the existence of

my debt but not paying it appropriately), by showing an absence of respect (I forgot I owe you any money), or by showing active disrespect (I was lying when I told you I'd repay the \$10 I owe, and I'm going to steal \$10 more when you're not looking).

Blame is deserved when the agent fails to fulfill a moral obligation that properly binds her, and the state of affairs -- the fact about the agent -- that justifies an assessment of moral failure is an absence of the required regulation. A moral wrong is committed when an agent forms an intention that violates a moral obligation. Making note of that failure tells us the agent is blameworthy in some amount. But it is the particular shape of the failure that will determine how much blame the agent deserves. We can speak of "better" or "worse" moral failures by identifying where and why the failures of regulation take place. Failures that stem from foundational disrespect, or malice, will be different in kind than failures of other sorts, for they show an active and intentional dismissal of the obligation's object. They are considered wrongs, wholehearted, and complete. Failures that involve partial, but not complete, regulation will be failures of a different sort.

What it is to regulate one's action formation in light of moral obligation is to perform correctly at several stages of the process. It requires respecting the object of the obligation in granting her the appropriate status by 1. including the obligation she grounds on the deliberative field, treating it as a reasons-generating fact,²⁷ 2. populating the remainder of the deliberative field with only considerations of the same status, and 3. respecting the object of the obligation in the content of the deliberation, by not letting the obligation she grounds be undervalued by a lesser consideration.

²⁷ For more on treating moral peers as reasons-generating, and the second-person relationship as a source of normative reasons, see Stephen Darwall's *The Second-Person Standpoint: Morality, Respect, and Accountability*.

Let me pause momentarily and discuss why I locate the success and failure of meeting these obligations within deliberation. One of the challenges I discussed in Chapter 2 was how to understand action description for the purposes of moral assessment, and this challenge by and large concerned the description of deliberation. This challenge exists because a singular focus on intention, at least in terms of what the agent takes herself to be doing (or how she answers certain questions, or what she wills, and so on), does not come near to capturing the fullness of what it is she does. The deliberative process is both the unspoken context of intention, and intention's origin story. It tells us the deep story that exists behind intention, including what the agent considered and discarded, what the agent failed to consider at all, and how she engaged with the content of what she considered. I focus on these pieces because they are part of what must be regulated in the fulfillment of moral obligation. Correct deliberation is part of what obligation requires, and so we are vulnerable to moral assessment for the shape of deliberation. We think of moral assessment centering on intention, but I want to suggest that proper evaluation and understanding of intention will require deep description of deliberation. I believe this is what Wallace is trying to gesture at with his discussion of coercion. Wallace believes it matters that the coerced agent did x-rather-than-y, because he thinks that x-rather-than-y displays a different moral attitude towards the object of the failed moral obligation than strictly x. And he is right. But as I mentioned in Chapter 2, this is not only true for cases of coercion. Most of our actions take the form x-rather-than-y-rather-than-z and so on, where the "rather than" string is the content of the deliberation that got us to our intention. The way to properly morally assess is not merely to focus on the surface intention, the x. It's to include the "rather than" string when this string contains morally relevant deliberation. The deep story of an agent's deliberation, how it is

she got to her intention, matters because moral acknowledgment or respect does not merely happen at the surface.

There will certainly be pieces of deliberation that are not relevant to moral assessment, of course, like the woman who briefly considered whether to mention a trivial observation during the murder she was committing. But we can capture pieces of deliberation that are morally relevant when we look to the constitutive story of the intention to violate or fulfill the moral obligation. Using this outside-in method, working from the external context of the obligations that bind the agent all the way to the deep deliberation that was part of the intention-formation process, will allow us to essentially start the story at the conclusion, tracing the plot all the way through. We will be able to see that the decision not to mention the haircut was not a “rather-than” contender; it was an “in-addition-to” contender to the murder. And so starting from the woman’s intention to murder and working our way back to uncover the full picture of her deliberative process, we will see that this consideration was not part of the story of how she formed her intention. Since the agent was not bound by an obligation to speak about haircuts (or to avoid speaking about haircuts), we can dismiss this deliberative detail as not being relevant to our immediate moral assessment. It was concurrent deliberation, not constitutive. An outside-in approach to moral assessment will help us distinguish the difference, and the ability to trace the “plot” of an intention, as it were, the formative circumstances, will allow us to begin to make distinctions about “better” and “worse” violations of a common type. I have tried to make the case that the evaluation of deliberation in moral assessment is necessary to capture the genuine moral worth of the action. But I have not yet made the case for why we should think that mitigation will be located within deliberation. Now it is time to do just that.

As I mentioned earlier, moral obligations require that we regulate our actions at both the level of deliberation and the level of performance in order to show the appropriate acknowledgement of the object of the obligation. For actions that merit the assignment of some amount of blame, we know that something has gone wrong for the agent, and at the very least, she has not regulated her action performance appropriately. Since we are worried about mitigating circumstances that reduce the amount of blame deserved for this violation, the only place to look for potential candidates to justify mitigating circumstances is at the level of deliberation that makes up the morally relevant details of the action. To be clear, I mean by deliberation the parts of action both prior to and inclusive of the intention-formation process. I use the word “deliberation” to capture not only the mechanics of comparing and contrasting data and stringing together inferences, but also the judgment that goes into choosing how to compare, contrast, and string things together. I use this word to refer to the deep description of the intention, its origin story and the process of its formation.

If we are to find mitigation anywhere – some factor that alters the condition of how much blame the agent deserves for the violation of a moral obligation – it can only be located somewhere in the conditions relevant to the assignment of blame. This is because mitigating conditions must have the ability to alter the blame deserved for the performance of the action. Above, I offered that the conditions relevant to the provisioning of blame are the lack of appropriate respect or acknowledgement of the status of the object that grounds one’s obligations. And thus, the only way mitigating circumstances can have the power to alter our judgments about the amount of blame deserved is if they somehow affect the lack of appropriate respect or acknowledgement. There are only a few places this can happen.

When we morally assess action, there is a temptation to view action as a singular, continuous entity, but this is not the case. I speak of action morally succeeding in one way and failing in others because there are a number of discreet points where failure or success is possible. I identified these three discreet points above: 1. the inclusion of the obligation on the deliberative field (treating it as a reasons-generating fact), 2. the populating the remainder of the deliberative field appropriately (only with considerations of the same status), and 3. the proper managing of the deliberative content (weighing of demands of the obligation appropriately). These three steps are what it is to form the morally correct intention appropriately, and each of them represents a piece of required regulation. Moral obligations do not require us to do one thing; they require us to do all of these things, and the complete package will be successful moral action.

Blame is deserved when a failure to display respect appears in this process, but the *amount* of blame will depend on exactly which failures occur. Mitigation will occur when the agent failed to show appropriate respect at some, but not all, of the stages in this process. Blame is appropriately attached to each point vulnerable to assessment, and the agent deserves blame for each failure. But she does not deserve blame for where she succeeds. And so if a total failure – a lack of appropriate respect shown – is worth some X amount of blame, then it will be because each of the constitutive failures are worth some amount of blame, and added together, the total will be X . We can speak in these arithmetic-like terms because there is nothing magical about a total moral failure that makes it worth a greater amount of blame than the sum earned by its constitutive failures. There is no blame bonus for missing every possible mark. It is only worth what it's worth, blame-wise, because a total failure represents failures at each stage of the process. Blame is deserved for particular failures to regulate appropriately, and it will attach to

each point where regulation is required. We can say that a failure is *better* (or perhaps, less bad) when there are fewer points of failure, and *worse* when there are more. Or, to invoke the common sense notion of mitigation I introduced earlier, we can say that a less-bad wrong has mitigation.

We often speak of mitigation as something that serves to reduce the amount of blame deserved for an action, and now we can say why. Mitigation occurs when the deep description of the action shows that there is at least one portion of the deliberation, the intention-formation process, which is regulated appropriately. The agent deserves less blame because she did not fail at that test-point. The action as a whole will still be wrong, and she will still be deserving of blame for the points where she did fail, but she will not deserve as much blame as someone who fails on all counts of the action.

I call these partial failures deliberative irregularities – times when the sub-parts of the deliberation and intention-formation process are not morally consistent. Deliberative irregularities will occur when appropriate acknowledgement of moral demands exists at (at least) one point in the action-formation process, but failure to appropriately acknowledge exists at (at least) one point as well. Mitigation will obtain when the deliberative process contains these irregularities. Since there are three general places where appropriate acknowledgment demands regulation, we can see that mitigation will come in a few general types. First, the nature of the requirement to place the relevant consideration on the deliberative field necessarily comes first in the process, since deciding that something is worth considering must come prior to the particulars of how one considers it. But a failure at this stage is unique in that it prevents any later chance of success in regulation, and thus, a failure at this stage cannot present a case of deliberative irregularity. The reasons for this failure will matter to moral assessment, of course.

Having control over what considerations we take seriously in the action-formation process is one of the necessary conditions for agency, and certainly falls under Wallace's notion of cognitive powers. If someone fails to place the demands of a moral obligation on her deliberative field, fails to consider them reasons-worthy, because she has been drugged or because she is under the complete control of an evil scientist, for example, then our moral assessment process will result in a case of excuse and not a case of mitigation.²⁸ When there is an impairment that stops the process of action-formation before it starts, we are not looking at a case of genuine action. Of course, a moral mistake could also occur at this stage; we might have an agent who does not realize that she is bound by an obligation, and thus does not place it on her deliberative field. Here, we ask whether the type of her mistake is morally relevant. If the mistake is epistemic in nature, then the agent's failure to place the obligation on her deliberative field will either be excuse or justification (depending on how we cash this story out). If I see you waving at me from the ocean and wave back, not realizing that you are flailing your arms to indicate you are drowning, I have not disrespected you by not taking my obligation to come to your aid as reason-generating. With few exceptions, failures to show proper acknowledgment in this way will turn out to be cases of exception or justification, or blame-meriting failures.

Failures in the next stage of the process might well be contained within a deliberative irregularity. This would look like success at the junction of placing the obligation on the deliberative field, but failure to populate the deliberative field with only appropriate contenders. A failure here could also be part of a regular failure, of course, but should the agent succeed at the level of taking the obligation to be reasons-generation (and thus appropriate to consider on her deliberative field) but fails to populate her deliberative appropriately, then the agent has a claim to a minor amount of mitigation. Placing the demands of the moral obligation on one's

²⁸ See Chapter 5 for a continuation of this discussion.

deliberative field is at least paying respect to the reasons-generating force of the obligation – it is acknowledging the type of thing that the moral demand is, and the status of the object of the obligation. But failing at the level of populating the deliberative field is failing to show the proper amount of acknowledgement. I will discuss this circumstances below, when I turn to some examples of this theory in action, but a mistake of this kind is like acknowledging I owe you a debt but seriously considering folding the money into little origami flowers to decorate my desk instead of paying you back. Seriously weighing the possibility of folding the bills into flowers – placing it on the same field as the demands of my obligation – is fundamentally disrespecting the kind of thing a moral obligation is.

And the final type of deliberative irregularity will occur when one succeeds at the first junction, and fails at the third. (Success or failure at the second will determine whether the agent has a claim to more or less mitigation, and thus deserves more or less blame.) These are agents that take their moral demands to be reason-worthy, but who mis-weigh considerations, failing at the level of deliberative content. These are agents who seriously consider doing what they should, but they undervalue their moral obligation and thus fail to regulate their deliberation appropriately. Mitigation occurs at this point of this morally relevant deliberative irregularity. If the agent fails to consider her obligation at all, not bothering to place it on her deliberative field as worthy of serious consideration, then she is fully blameworthy for all of the violations that necessarily follow. But let's say the agent in question does take her obligation to be worth deliberating about, and she considers it a relevant contender to drive her motive. The question concerning this second check-point is now this: how seriously is she taking her obligation? The way to answer that question is by looking to see what other considerations the agent places as equally deserving contenders on the field. If she populates her deliberative field with morally

frivolous considerations, she is showing a fundamental disrespect towards her obligation, and the object of her obligation who grounds that action. She does not deserve as much blame as someone who does not even consider the obligation to be reasons-generating, but she still deserves blame for her failures. But there will be times, for example, when the agent populates her deliberative field with her moral obligation, and only other morally relevant considerations. This is appropriately acknowledging the obligation. If you place an athlete on a minor league team with other minor league players, you are displaying a judgment about where that player belongs. Populating the deliberative field in the appropriate way, with considerations of the appropriate level, is a way of displaying a judgment about what kind of consideration the moral obligation is, and what kinds of considerations deserve share its deliberative space. This kind of regulation just is part of the respect demanded by obligations.

A common example of this type of irregularity is a case to which I will soon return: coercion. In cases like these, where the agent has taken her obligation to be reasons-generating and has populated her deliberative field with only appropriate considerations, she has shown appropriate respect for two of the stages demanded of her. She does not deserve blame for either of these pieces. Her moral failing comes at the end, where she does not consider these factors appropriately. I will discuss these kinds of irregularities in more depth by looking at particular cases. It is worth noting that there are some common circumstances that often bring about these deliberative irregularities. Impairments that prevent the agent from being able to stably regulate a part of her deliberative process will often bring about deliberative irregularities, as will cases like moral overlap, where an agent mishandles the weighing of multiple moral obligations.

Even the three “check-points” we examine for deliberative irregularities can be failed in many different ways. I focus on these three points in the deliberation and intention-formation

process because they are three distinct moments in the process that must be gotten right in order for a moral obligation to be satisfied. But it is not the case that blame will simply be deserved in three blocks, where we can grade 0/3, 1/3, 2/3, or 3/3. Blame is deserved based on the nature and magnitude of the failure, and enthusiastic, wholehearted failures display a disrespect that will cause the agent to deserve more blame than someone who merely fails to show the appropriate respect. It is one thing to disrespect an obligation because you don't think it is important. This is a moral sin, and it will earn blame. But it is another point on the continuum of disrespect to actively and gleefully disrespect someone's status for the sake of disrespecting it. The kinds of failures at each of these stages, even with the presence of deliberative irregularities, will vary based on the nuances of the failures, and how those failures show (or don't) the appropriate acknowledgement of the obligation's source. Deliberative irregularities, then, don't tell us precisely how much blame we can knock off the performance of a particular wrong. But they do illustrate that there is mitigation, in some amount, available to the agent, and they signal that the moral assessment process must take these nuances into account. To see the fullness of this picture, I now turn to some cases.

III.

First, let's return to coercion. Many philosophers have offered theories about how to handle these cases. In *What We Owe To Each Other*, Scanlon offers an analysis of coercive cases.²⁹ He argues that in cases of coercion, agents are "forced to do something that they would not otherwise have chosen to do."³⁰ This fact, he claims, makes them responsible for their actions, but not blameworthy. Later, in *Moral Dimensions*, Scanlon writes, "It is one thing to

²⁹ T.M. Scanlon, *What We Owe To Each Other*, pp. 279-280.

³⁰ *Ibid*, p. 280.

inflict harm on another person gratuitously, another thing to do so (even unjustifiably) out of fear of harm to oneself. ... Changes ... (in blameworthiness) occur because coercion changes the reasons on which the agent acted.”³¹

There are several things that Scanlon could have in mind here. The first possibility centers on structural features of deliberation. One of the distinctive elements of cases of coercion is that the coerced agent does not have a primary end to commit the harm she performs. Instead, her desire to act wrongly is secondary, dependent on her primary end, to avoid the threatened harm. Scanlon phrases this by saying that she “would not otherwise have chosen” to perform the wrong in question. While this is no doubt true of cases of coercion, it is not clear how this feature could ground mitigation. We commit all kinds of moral wrongs for instrumental reasons, and yet it seems strange to say that the structural component of deliberation has power to reduce blameworthiness. If this were true, the Calculating Lover case I describe in Chapter 1 might also contain mitigating circumstances, since the murderer did not hold the death of her victim as her primary end. Her primary motive was financial gain. She just used the murder as the means to get to her desired end, and she “would not otherwise have chosen” murder. To think that this gives her mitigating circumstances seems deeply problematic, and counter to most of our intuitions. I suspect most of us think that this particular case of deliberation makes her deserving of even more blame than if she had just wanted her lover dead.

There is something attractive about this structural approach to mitigation, however, because it suggests a kind of lack of endorsement or wholeheartedness that is implied by the agent not otherwise choosing this course of action. But there are other problems with focusing on deliberative structure. Not only does this approach fail to distinguish between cases of coercion and cases of murder-for-profit (and others like it), it also cannot help explain our intuitions that

³¹ T.M. Scanlon, *Moral Dimensions*, p. 181.

some forms of threat are not adequate to count as coercive threats, even when the format of deliberation is as described. Additionally, many philosophers also suggest that there are some actions that are sufficiently morally bad that even the presence of coercion is not enough to adjust our moral assessment of the action. Judith Jarvis Thomson, for example, argues that while it may be permissible to kill someone who is attacking you, it is not permissible to kill an innocent third party to escape the threat of death from another.³² Scanlon's view lacks the resources to explain intuitions like Thomson's, particularly her contention that some moral wrongs are such that even genuinely coercive circumstances can't mitigate or excuse them. It seems unlikely that the deliberative structure is providing the mitigating force in cases of coercion. That said, I do think Scanlon is onto something interesting and true, but there is still much left to say.

Where we left off on Wallace's view, coercion *doesn't* work as a mitigating circumstance for the purposes of moral assessment. On this picture, the morally important feature of coercion is that the intention within the act in question is not plainly "x" and is instead something closer to "x-rather-than-y." I have argued that this is not a satisfying picture. Most philosophers agree that not all forced "x-rather-than-y" cases count as coercion, and in particular, if coercion has any special moral properties, those moral properties will not apply to cases where an agent tortures someone because she has been threatened with a poke in the ribs.³³ If, as Wallace suggests, coercion was merely a signal to redescribe the action in a more nuanced, accurate way, we wouldn't need all these caveats.

³² Judith Jarvis Thomson, "Self-Defense" *Philosophy and Public Affairs*, Vol. 20, No 4 (Autumn, 1991), 283-310.

³³ For some relevant cases on the legal limits of duress and reasonable firmness, see *State v. Van Dyke* (361 N.J. Super. 403, 825 A.2d 1163), a case where a woman was denied a duress defense because it was determined that the threat was not sufficient, and *Commonwealth v. DeMarco* (809 A.2d 256 Pa. 2002), where the court took into account the defendant's borderline mental retardation when considering his duress defense.

But the caveats are important because coercion is more than just a trigger for action redescription. While Wallace argues (rightfully, I believe) that the appropriate way to understand the intention in a case of coercion is “x-instead-of-y” instead of “x”, he claims that the agent is still fully responsible for the properly described action. This just pushes the problem up a level, for we are still left with the question of how to appropriately assess a coerced “x-rather-than-y” intentional act. I believe we now have the appropriate tools to do just that. I return to the case of the Secret-Spiller.

The way I described the case, the Secret-Spiller has performed a morally wrong action, and her primary reason for acting was the threat that some other person made against her life. Let’s morally assess her action. First, we look to the obligations that bind this agent. She is bound by her vow to keep this information safe, and we are granting that this is her moral duty in this case. This is the obligation she fails to uphold, and thus her action is properly categorized as a failure of this type. Once we know the standard we are holding her to – the content of her obligation – we can turn to the shape of the failure. In this case, we know that her motive for acting was the threat upon her life. She performed a particular wrong, but she acted as she did out of a desire for self-protection, a morally worthy end. The Secret-Spiller still failed to fulfill her obligation, but the way she failed will determine how much blame she deserves.

As I argued above, obligations bind us, in part, by requiring us to constrain the shape of our deliberation. Their moral force over us is not just in what we decide to do, but also in *how* we decide what to do. In this case, the particulars of this agent’s deliberation are morally relevant in several ways. First, it is morally relevant that she formed the intention to break her vow (or, to be more precise, an intention that necessitated the breaking of her vow). This is the piece that gets her action characterized as a moral wrong of this particular type – she violated her moral

obligation in this way. The second piece is that her intention was influenced heavily (and inappropriately) by another agent, who used threats in an effort to influence her action.³⁴ The third piece concerns not just the outcome of her deliberation, but also her process of deliberation. In forming her intention, she took her obligation seriously and placed it on the deliberative field, and she chose to weigh the decision to break her vow against the threat her coercer made against her life, and the preservation of her life has moral value. She populated her deliberative field with morally relevant contenders, and thus showed appropriate respect at this level as well. But she mis-weighed those considerations. So the assessment will look something like this: an agent who breaks her vows with fully malicious intent deserves blame (is blameworthy) for some X amount. This is the paradigmatic violation of this particular moral obligation, and we can assess a theoretical wrong of this type, determining how much blame an agent would deserve for committing a wrong of this sort with full disrespect and no access to mitigating conditions.

But the Secret-Spiller's deliberation does not look like that paradigmatic case. As the case is described, we can't say that the Secret-Spiller displayed anguish over the choice and is actively fighting the intervention of the coercer – we just don't know. But what we do know is that her deliberation does not contain malice, intentional disrespect, or a dismissal of her moral obligation as her motive for action. She takes her obligation as a contender for motive, and considers acting on the reasons it generates. But the coercer has altered the shape of her deliberation in such a way that her motives of breaking her obligation happens indirectly, via her direct motive to preserve her life. And so, she does not violate her obligation in the same way, or we might say as *badly*, as the malicious secret-spiller. If the latter is fully blameworthy for some X amount, the Secret-Spiller will deserve X-Y amount of blame. And so it will go for other cases

³⁴ The coerced agent is also the victim of a moral wrong, and this might play a role in our moral assessment. For more discussion of this issue, see Chapter 4.

of genuine coercion as well. The Secret-Spiller's case contains what I have called moral overlap. These are circumstances where an agent is bound by more than one moral demand, and an accurate moral assessment will capture not only the nature of her wrong and the content of her deliberation, but also how she chose to respond to this moral layering. Mitigation in this case reflects that she is has appropriately regulated her considerations to only morally relevant considerations, the best possible acknowledgement of the status of the source of each obligation. She just fails on the appropriate handling of the moral overlap. There will be many cases like this that do not involve coercion, but do contain mitigation for the same fundamental reasons.

It is worth noting that I am not speaking here about moral dilemmas. The nature of a genuine moral dilemma is such that we cannot say that one moral choice "overrides" another.³⁵ If an agent is faced with a genuine moral dilemma and picks one of her morally acceptable choices, it is not clear to me that she is blameworthy at all. The moral layering I have in mind is more of the everyday sort. We have an obligation to pick our friends up from the airport when we promised we would. But that obligation is easily, and properly, overridden if we are called upon to rescue someone from a fire instead. This is moral overlap, two (or more) moral demands pressing on us, and part of the process of moral assessment is to evaluate how the agent handles these moral layers.

I also mean to exclude cases of epistemic mistake.³⁶ By this I mean cases where a person does not realize that she has moral overlap, through mistakes in awareness or circumstance. I might not realize that I am standing you up at the airport while I rescue these people from the fire, because my watch is wrong, so I don't even consider the promise I made to you because I

³⁵ For more on moral dilemmas, genuineness, and the notion of "overriding", please see Walter Sinnott-Armstrong's *Moral Dilemmas*, Oxford: Basil Blackwell, 1988, Chapter 1.

³⁶ Assuming these mistakes aren't caused by negligence, in which case we can morally assess the negligence as part of the deliberative content of the act.

don't think it is relevant. Or perhaps I went merrily to the airport to fetch you, not realizing there was a fire, because my fire-fighter pager had died. Cases of epistemic mistake are typically going to qualify as moral excuse or justification, depending on the nature of the error. But such mistakes do not reveal the respect (or lack of respect) that an agent has towards her obligations and her moral peers, and so these are not the kinds of cases relevant to the assessment of blame. The moral overlap of the Secret-Spiller case is far more straight-forward. She has an obligation to keep this sensitive material safe, and she has an obligation to maintain her own well-being. She has been forced into a position where she must consider which of the reasons-generating considerations should ground her intention.

The intuition that many of us have about the mitigating force of coercion stems from what coercion fundamentally entrails. The distinctive feature is not merely structural similarities – those caveats I mentioned above establish that format isn't enough to establish the presence of coercion. Cases of genuine coercion also involve tension between morally relevant considerations. Cases of coercion are not merely cases where an outside agent alters the landscape of reasons for the coerced. If they were, we could say that I was coercing someone by offering to think more highly of her if she performs the action that I prefer. Instead, cases of coercion depend on the introduction of a threat, and the shape of that threat matters. As I discussed earlier, a threat that the coercer will wear an outfit that the coerced does not like is not the right sort of considerations to count as coercive. We think that coercion requires meaningful threats to qualify it qua coercion, and so the obvious question follows: what makes a threat meaningful? I claim that a meaningful threat is one in which it is appropriate to deliberate about in relation to the obligation that the agent is being coerced into violating. The notion of deliberative irregularities – in particular, the idea that a successful display of respect exists in one

stage of the deliberative process and a failure in another – helps justify the intuitions many people have about cases like coercion. We might think that Thomson is correct in claiming that some actions are never permissible, even with the presence of coercion, because there is simply nothing that could permissibly outweigh it on the deliberative field. To use Thomson's example, coercion then can never be a justification for murder, since on the deliberative field, there is no moral concern that should ever rightfully outweigh murder. If someone murders a person because she has been coerced, she will always be deserving of some amount of blame. Her blame will of course be mitigated – assuming the coercive threat is genuine and meaningful, it is a morally relevant consideration that she rightfully weighs against the performance of the murder. But choosing to murder will simply always be the wrong choice. This picture of mitigation captures the force of these caveats in ways that Wallace's and Scanlon's picture can't quite explain on their own.

We can see that this kind of mitigation also happens independently of coercion. Let's say a person has a deeply unpopular peer, a peer who is openly gay. The two live in a deeply homophobic community, and her peer is often bullied. She knows that she could bully him as well, and doing so would gain her a great deal of social status. If all goes as it should, the agent, despite having those beliefs, also judges that bullying her peer is morally impermissible. She knows that bullying disrespects the recipient's dignity as a fellow agent, and that she has may not permissibly use him as a mere means to her social climbing. She deliberates, and her respect for her peer becomes her motive for action. She does not bully him. But things don't always go as they should.

Let's say that the agent does indeed harass her peer. Without looking at her deliberation, we don't yet know how much blame she deserves for this act. Let's consider one possible

deliberative landscape, calling this agent the Social Bully. In this case, the agent knows that bullying is impermissible and she has as an obligation to respect the status of her moral peer. But this is joined on her deliberative field by several other considerations. She also has a desire to be invited to a very exclusive party, and she believes that the status she gains by harassing her peer would secure her an invitation. And so it is this desire for the invitation, coupled with her belief about the bullying being a means to this end, which moves her to act. The Social Bully bullies him, and she is no doubt deserving of blame for both her failure to populate her deliberative field appropriately (a desire to go to a party is not appropriate to weigh against an obligation not to harass someone) and her over-valuing of her desire and under-valuing of her obligation. She no doubt deserves blame for these failures. Let's contrast this with another bully, who we'll call the Concerned Bully. In the Concerned Bully's case, she also knows that bullying is impermissible and she has an obligation to respect the dignity of her moral peer. But this is joined on her deliberative field by a belief that her peer is harming himself and his family with his orientation, and a belief that bullying him would force him to change. It is this desire to help, combined with the belief that bullying is an effective means to help, that moves her. And so she bullies her peer.

Although both the Social Bully and the Concerned Bully performed the morally wrong act (horrifyingly so), I think there is a strong intuition that the Social Bully deserves more blame than the Concerned Bully, or to be more specific, that something is mitigating some (though it remains to be seen how much) of the Concerned Bully's blame. The Social Bully might be a case of paradigmatic wrongdoing, but something more interesting is going on with the Concerned Bully. She has still made an enormous moral mistake; even if we assume that her belief that bullying her peer will indeed help him in some way is genuine, it is not acceptable for her to disrespect him as a moral agent and take the action she did. She deserves blame for her

morally wrong intention. But I argue that her mistake is of a very particular kind. The difference between her failure and the Social Bully's failure happens at the level of the deliberative field. The Social Bully populated her field with the moral obligation she had to respect her peer, and her desire for a party invitation. She weighed them both as potential motives, finally acting from her desire. The deliberative field for the Concerned Bully looks very different; it looks like the coerced agent's. The Concerned Bully populated her deliberative field with the obligation to respect her moral peer, and with a desire to help him. The difference in this landscape is morally significant. She did not fail at the level of populating the deliberative field.

In constructing her deliberative field as she did, The Social Bully treated her neighbor as if he was of the same type and significance as her desire for the party invitation. She treated them as considerations of the same sort, and of the same value. In judging them to both be relevant to her intention – letting them both have equal status as potential guides to her will – she fundamentally discounted the moral status of her peer. Yes, she considered her obligation towards him, but not in the right way. She surrounded it by reasons that were not worthy of being on the same deliberative field for this particular deliberation, and in doing so, she disrespected him in a fundamental way. If someone were to seriously deliberate over whether a da Vinci or a child's crayon scribble represented more skill, she would be disrespecting the very notion of skill. And so too with moral obligations. To populate the deliberative field with considerations that do not belong in this company is to disrespect the object of the obligation, making a moral judgment about the value of her status. To consider the obligation so flimsy as to be in the same company as desires for party invitations is not to treat it seriously as an obligation at all. To fail to regulate the content of the deliberative field is to fail to fulfill the demands of the moral obligation. Such a sin fundamentally disrespects the nature of the agent, whose status serves as

the source of the obligation. And to treat the obligation so lightly is to do the same to the source of that obligation. The Social Bully did not simply overvalue her desire for the invitation. It should not have been considered in the first place.

But this is not at all what is happening with the Concerned Bully. Her deliberative landscape is populated by her moral obligation, and something else morally relevant: a desire to help her peer. She overvalued this desire and undervalued her obligation to respect her moral peer, and she deserves the appropriate amount of blame for this failure, but she succeeded with placing the obligation on the deliberative field with only with other relevant contenders.

The view of mitigation I offer explains why many of us think that the Concerned Bully is intuitively deserving of less blame than the Social Bully, and it explains why the same kind of moral sin – even if it is motivated the same reason – can lead to different amounts of blame. The deep deliberation and the intention-formation process might be different, and the motive that actually led the agent to act does not capture the entire nature of the action. These cases of bullying and coercion illustrate just one flavor of deliberative irregularity, but there are others. In Chapter 4, I will give a fuller defense of this picture, and I will apply this theory to several cases. I will argue that this approach to mitigation can ground many of our everyday intuitions about moral assessment, and more importantly, help us understand where mitigation fits into the larger moral landscape.

CHAPTER 4: MITIGATING CIRCUMSTANCES

In Chapter 3, I introduced the idea that moral mitigation can often be found in the presence of morally relevant deliberative irregularities. While Chapter 3 focused on one particular type of irregularity – failures with moral overlap, such as cases of coercion – there are other types. A fuller picture of mitigation will require examining the variety of irregularities, and seeing why these circumstances that cause them have mitigating force. This chapter will do just that, focusing on circumstances that many commonly believe are mitigating. In this chapter, I will consider several types of intuitively mitigating conditions: impairments such as adolescence and intoxication, passion-inducing circumstances, and actions performed by agents who themselves have been morally wronged. I will argue that my approach to mitigation will shed light on the moral status of these cases, and furthermore, understanding these cases will help illustrate why their characteristic deliberative features result in mitigation in the assignment of blame.

Before moving on, however, I would like to note that I will be discussing impairments, and this concept is necessarily thorny. In legal contexts, people often argue that the presence of impaired capacities reduces the defendant's culpability for a crime, or the amount of punishment she should receive.³⁷ This legal position is typically justified by the claim that it is morally inappropriate to hold someone with impaired capacities fully blameworthy for her crime because she is, in some sense, less than a full agent, and the implication is that such status will qualify as a mitigating or excusing circumstance. This needs some clarification. There are certainly impairments that prevent a person from deliberating at all. We might think that if a person is

³⁷ For examples, see 18 USC § 3592 - Mitigating and aggravating factors to be considered in determining whether a sentence of death is justified, Arizona Capital Sentence Guide A.R.S. §13-751(G)(1), and Section 19.2-264.4.B of the Virginia Code.

incapable of means-ends reasoning, for example, then such impairment would disqualify them from doing the bare minimum for deliberation, and this type of impairment will no doubt serve as a global excuse for the purposes of moral assessment. In literature on blame and responsibility, discussions of impairment often focus on conditions that undermine agency, such as the examples from Strawson that I offered in Chapter 1, circumstances that include being a child, for instance, or an extreme schizophrenic. In Chapter 2, I described Wallace's notion of powers of reflective self-control, the cognitive and affective powers that he argued are necessary and sufficient for moral agency.³⁸ Impairments that prohibit someone from possessing (or exercising) these powers will turn out to eliminate the foundation of agency itself. All forms of extreme mental illness will probably qualify, along with vegetative states and brain damage, as well as temporary conditions such as extreme intoxication, mind control, and so on. Circumstances like these will undermine agency entirely, acting as excusing conditions and making it such that it would be inappropriate to assess the person as blameworthy at all. And so we have one class of impairments: a collection of circumstances that make it inappropriate to assign blame, a collection of impairments that excuse rather than mitigate.

Notice, however, that these excusing cases have all been qualified by degree. It is no doubt true that "extreme" mental illness will be excusing, as will being a young child, but there are less extreme cases to address. A two-year-old obviously lacks the relevant prerequisites for agency, but what should we say about a 14-year-old, for example? Someone mildly intoxicated, or with a minor mental illness? The distinctive feature of these "minor" impairments is that they do *not* eliminate agency. We do not assume that a person is incapable of intentional action because she suffers from depression, nor should we. And so we have a category of conditions that many of us feel are intuitively relevant to moral assessment, but don't serve as excusing

³⁸ Wallace, pp. 154-194.

conditions. Here is what we cannot do: we cannot draw inferences about mitigation from the fact that extremes in certain conditions excuse. This is no different than trying to make an argument that drinking 20 ounces of water/hour will make you sick because drinking 3 liters/hour will kill you. And so any discussion of impairments and mitigation will have to include argumentation about which impairments are relevant to moral assessment, and exactly what effect they have on the appropriate provisioning of blame. This is the task I will take up in discussing several types of impairments. To begin, let's turn our attention to a "less extreme" version of an excusing condition that Strawson raised. If being a child excuses, what are we to think about the state of adolescence?

I.

The label "adolescent" is used in many ways. We might use the word to describe people within a certain biological age range, and we might use the term to refer to people in a certain developmental stage. And while it is true that those two often, and typically, correspond, they do not always do so. Humans mature at different rates, and it is entirely possible to find one nineteen-year-old who is in all relevant ways identical to an adult and another who is still firmly in the adolescent developmental stage. When I discuss the circumstances of adolescence, I am referring to the developmental phase. While this may present an epistemic challenge for the application of blame and punishment, this is not a problem when it comes to determining the appropriateness of blame. It is also true that epistemic issues make it very difficult to tell if someone is genuinely insane, for example, but it is noncontroversial that genuine insanity is an excusing condition for blame. I am concerned with the developmental stage of adolescence, whatever the age boundaries happen to look like for particular individuals, and how that impacts the appropriateness of moral assessment.

There has been much recent research into the condition of adolescence. Specifically, psychologists have tried to find whether there is some combination of factors that explain (or cause) the fact that adolescents, as a group, tend to exhibit more morally problematic behavior than adults. As a group, adolescents engage in extreme risk-taking, violence (towards others or oneself), and other socially disruptive behavior³⁹ at a significantly greater rate than the general adult population. This is behavior for which normal adults, barring some special story, would be considered fully blameworthy. Much of the contemporary psychological literature has focused on whether adolescents' behavior is caused by some deficiency (or deficiencies) in capacity, and whether such capacities might also be relevant to the question of responsibility and mitigation. Until fairly recently, much of the psychological research done on adolescents focused on their cognitive capacities. By this, researchers meant something like rational faculties: the ability to use means-ends reasoning, to foresee consequences, to use inferential and deductive reasoning, and so on. While no one thought that adolescents lacked this capacity out-right, many hypothesized that there was some distinctive immaturity in adolescents' cognitive capacities that (generally) prevented them from deliberating as well as adults. Researchers conducted experiments that compared adolescent cognitive capacities against adult cognitive capacities, and despite the fact that the cognitive capacities model was once regarded with great optimism, researchers have found no significant differences in cognitive capacity between adolescents and adults.⁴⁰

Some psychologists have also hypothesized that a lack of knowledge about actual risk and consequences of actions might explain these trends in adolescent behavior. Perhaps what explains the difference between adolescent and adult risk-taking behavior, for example, is the

³⁹ This is often referred to as "antisocial behavior" in the psychology literature.

⁴⁰ See Fischhoff, "Risk Taking: A Developmental Perspective."

fact that adults generally have more experience and familiarity with the actual real-world implications of their actions. And while this is certainly true (it is obvious on its face that adolescents, by nature of their limited life-spans compared to adults, have fewer experiences with real-world consequences), this alone does not explain the distinctive differences in behavior trends between adolescents and adults. Researchers have tested this hypothesis by spending time informing adolescents of the particular dangers of high-risk behavior, testing them on their understanding of the dangers, and then tracking how many of them performed that high-risk behavior. Their results showed that providing (and testing for) understanding of the dangers of high-risk behavior did not in *any* meaningful way reduce the incidences of adolescents engaging in that risky behavior.⁴¹ Simply giving adolescents more information (and testing their comprehension of that information in a laboratory) was not enough to change their behavioral tendencies.

Psychologist Laurence Steinberg has recently put forth a more promising theory about the source of adolescent antisocial behavioral tendencies, and his theory has revolutionized the field's views on adolescence. He notes the unsatisfactory results of the work done on adolescent cognitive capacities, and instead suggests that adolescence should be more properly understood as a state of immature psychosocial skills. He writes, "One reasonable hypothesis is that adults and adolescents [...] share the same logical competencies but that age differences in social and emotional factors such as susceptibility to peer influence or impulse control, lead to age differences in actual decision-making."⁴²

⁴¹ See Steinberg, "Cognitive and Affective Development in Adolescence."

⁴² Steinberg, p. 71.

Steinberg argues for a link between pubertal maturation and specific developmental changes in arousal, motivation, emotion, sensation-seeking, risk-taking, and reckless behavior.⁴³ He found that while development of cognitive skills tended to level off after mid-adolescence and do not track the distinctive behavior trends of the group, puberty seems to be a signifier of the beginning of a rapid maturation of capacities that aren't, strictly speaking, cognitive. His claims appear to be supported both by empirical studies⁴⁴ and by adolescent brain scans.⁴⁵ His picture of a typical adolescent, then, is a person who can deliberate well in some respects. She can give socially and morally appropriate answers to hypothetical situations posed to her under controlled circumstances, and she is able to competently draw appropriate inferences, use means-ends consequences, foresee consequences, and so on. But because her psychosocial capacities are still immature, her deliberation is, in a sense, fragile. When put in a "hot" environment where her psychosocial capacities are tested (she is exposed to peer pressure, for example), she deliberates differently – and usually worse – than she did in the lab. We might say of the typical adolescent that she is a good deliberator, as competent as a typical adult, except for when she's in circumstances that test her skill. According to Steinberg's picture, she meets some of the qualifications for adequate deliberation some of the time, but because of her psychosocial immaturity, she is prone to poor deliberation when faced with certain kinds of obstacles. In other words, an adolescent's deliberative experience is different than a normal adult's internal experience. Vulnerability to peer pressure is part of the picture (of course), but this immaturity is also represented by impulsivity, reduced future-orientation, poor emotional control, and unpredictability. While some of these features may (and certainly do) influence how "loudly" the adolescent experiences her desires, emotions, and so on, some of these features (like reduced

⁴³ Steinberg, p.72.

⁴⁴ Ibid., p. 69.

⁴⁵ Ibid., p. 70.

future-orientation) affect how she weighs considerations without a strong quasi-sensory experience attached. In the heat of the moment, it appears that adolescents are prone to acting in ways that are inconsistent with beliefs or judgments that they have about the dangers and future consequences of their actions.

At core, adolescence is a state where people do not have the ability to stabilize their deliberative process, either at the level of deliberative field, or at the level of deliberative content. People in this developmental phase are prone to impulsivity, at times unable to translate their beliefs and judgments into motives for acting, and they are vulnerable to social pressures over-exerting themselves in the deliberative process. Thus, we can consider adolescence as a type of impairment, characterized by deliberative fragility. This is not to say that an adolescent's deliberative process is prone to breaking completely; adolescence is not a stage where subjects find themselves compelled to perform some action that is completely foreign or external to them. We might think that cases of complete deliberative breakdown would more closely resemble something like Tourette's syndrome, where action is interrupted by physical and verbal tics that do not, in any way, stem from one's agency. Events such as these are a departure from deliberation altogether, a foreign interruption. They are examples for which the agent is not at all blameworthy, cases of local excuse. But this is importantly different that the condition of adolescence.

When I say that an adolescent's deliberation is fragile, I mean that all of the beliefs, desires, emotions, and so on, that are part of her deliberative process are, in a sense, rightfully hers. They originate with her, and they represent, in one manner of speaking, the values and judgments that she genuinely holds, and there are no deliberation-stopping interruptions. But the adolescent is not always in complete control over how her deliberative pieces get put together.

All the parts are hers, but she does not necessarily endorse their assembly. This impacts her deliberation in two ways. First, it can impact what is placed on her deliberative field at a given moment: what she takes to be relevant to her deliberation and which considerations she takes to be reason-worthy. And second, it can impact how she weighs the considerations on the field: what she decides will move her to act.

The unique feature of the adolescent is that this deliberative instability does not stem from a lack of commitment. Let's return to the topic of bullying I raised in Chapter 3. This is one of the moral wrongs especially common in adolescence, a moral sin that flourishes in this developmental stage. We can imagine the following case: an adolescent knows that bullying is wrong. In the cool, non-pressured environment of a laboratory, she will correctly answer moral questions about how to treat others. She is not inclined to bully in these cool environments, and she has no latent desire to bully for her own entertainment.⁴⁶ But once she is in a "hot" environment, with the full force of peer pressure upon her, the adolescent joins in and mocks her peer in order to increase her social standing with her popular friends. This adolescent's deliberation is importantly different than the agents I described in Chapter 3. Let's return to the Concerned Bully, who had a claim to (a small amount of) mitigation because she succeeded at regulating the content of her deliberative field appropriately, but failed in mis-weighting morally

⁴⁶ There are, of course, adolescents who do not share these traits with the one I am describing. There will certainly be adolescents who maliciously desire to do wrong out of a firm and stable commitment to harming others, perhaps because of ill will towards their victims. If these are the attitudes that the adolescent holds, she is fully blameworthy for the moral wrong she commits in bullying her victims, and she has no claim to mitigation. We should not think that adolescence is a mitigating condition for all wrongs that an adolescent might commit. It is far more accurate to think of the mitigating force of adolescence as similar in type to the excusing force of epilepsy. The epileptic has an impairment that predisposes her to episodes from which she is fully excused. If an epileptic has a seizure and strikes her nurse, we do not say she is blameworthy at all for this injury. But if the epileptic intentionally strikes her nurse, not while she is incapacitated by a seizure, she has no claim to excuse. The adolescent functions in much the same way. No one suggests that adolescents are *always* incapable of stable deliberation. There will be times the adolescent functions as well as a normal agent, and when she is not being affected by her impairment, she has no claim to mitigation. But the condition of adolescence is like the condition of epilepsy. Both are vulnerable to certain sorts of interruptions in normal agency, and mitigation will occur when the adolescent's deliberation is compromised by her impairment.

relevant considerations, inserting herself into a role of “helper” in a way that was entirely unacceptable. She got the first two parts of the morally required deliberation correct, but she failed in the final part, and she is fully to blame only for her failure at the end of the action-formation process. Although the bullying adolescent will have a claim to mitigation, it will be for very different reasons.

To understand the bullying adolescent’s claim to mitigation, we must look to the particulars of her deliberation. It looks like this: the adolescent has judgments and beliefs about the moral worth of her peer, the lower relative worth of fitting in with a popular crowd, and the impermissibility of bullying. In cool environments, she is able to express these judgments and beliefs and deliberate with them appropriately, coming to morally correct decisions. In hot environments, facing extreme peer pressure, she loses control over the structure of her deliberation, and the desire to fit in with the popular crowd unseats her judgment that it is of lower relative worth than respecting her peer. She is not *committed* to the belief that fitting in with the popular crowd is *more* valuable than respecting her peer. Instead, it is more accurate to say that she is vulnerable to this kind of consideration overwhelming her judgments. Just like a person allergic to ragweed will be vulnerable to respiratory distress in its presence, so too is the adolescent vulnerable to deliberative irregularity in the presence of certain triggers. Her deliberative irregularity stems from a kind of deliberative distress.

We can contrast the adolescent with a normal agent who also deliberates differently in hot and cold environments, someone I will call the Childish Adult. Let’s say that the Childish Adult also answers moral questions appropriately in lab settings and in cool situations, but when faced with hot, or challenging, environments (such as the influence of peer pressure), the Childish Adult gives in to her desire for social belonging and acts impermissibly, bullying her

peer. To appropriately morally assess the Childish Adult, we must look at the particulars of her deliberation. In the hot environment, she encounters social pressure to perform a morally wrong action, and while she also holds the judgment that bullying is morally impermissible (and that the action she is contemplating is indeed an act of bullying), she allows the desire to fit in to override her judgment about the moral status of bullying and the moral status of her victim. I use this language intentionally. The deliberation of the Childish Adult is fragile because she lacks a commitment to consider the content of her obligations appropriately.

The adolescent also lacks a commitment to consider the content of her obligations appropriately, but the source of that absence is important. The Childish Adult's lack of commitment comes about as part of her overall deliberation. She does not commit to taking the content of her obligations seriously in the same way that I might not commit to seeing clearly when I cross my eyes. I am doing something else, and the instability of my vision, the fact that the world will come in and out of focus, is the necessary consequence of my choice. The adolescent's lack of commitment stems from her impairment. Her vision comes in and out of focus because the mechanics of her eyes do not yet function appropriately. And so we might think that the Childish Adult has made a morally relevant judgment about the content and object of her moral obligation, while the adolescent has not. In circumstances such as these, the Childish Adult allows herself to be overcome by other considerations because she does not appropriately value her obligations and their source, and she deserves blame for that instability. Her instability is a result of moral judgment she has made, namely, that she either values something inappropriately over the maintenance of appropriate deliberative regulation (in which case, she has failed to populate the deliberative field appropriately *and* she has failed to weigh the considerations that are morally relevant correctly), or she does not have a genuine

commitment to regulate her ends appropriately at the most basic level. If the latter is true, then she has failed to acknowledge the reasons-generating force of her obligations adequately.

Whatever the particular shape of her failure, she deserves blame for it. The adolescent, on this point, has done something very different. To be precise, the adolescent has made no choice upon which we can assess.

We can see why, then, the condition of adolescence is mitigating. Moral obligations demand we regulate our action appropriately as moral acknowledgment, and that includes the regulation of our deliberative process. This regulation is constitutive of the respect due the source of the obligation. An obligation to not bully, for example, may not permissibly be overridden by a desire to gain social status. To allow this inclination to override one's obligation represents a lack of appropriate deliberative regulation, and this is, in a basic way, a display of disrespect towards the object of the obligation. It is implying that the object's worth qua source of obligation is vulnerable to being set aside. This lack of commitment to regulate one's deliberation appropriately justifies blame for the agent, since part of the demand of the obligation in question just is the appropriate regulation of one's deliberation. The adolescent is also not showing appropriate respect to the object of the obligation, but the absence of appropriate respect does not, on its own, establish grounds for blame. We can assess the Childish Adult's lack of commitment to regulate her deliberation appropriately because regulation, of a morally inappropriate sort, is still occurring. The Childish Adult is still deliberating as a normal agent, with no impairment or unusual circumstances asserting themselves upon her judgment. The population of her deliberative field is hers, in a deep way, as is the content of her deliberation. The Childish Adult certainly deserves less blame than the agent who bullies out of purely malicious intent, but the lack of respect is still a reflection of her agency, and her lack of

appropriate regulation is a moral failing. Unlike the Childish Adult, the adolescent I described above hasn't actually regulated, or failed to regulate, her deliberation. That process has been interrupted, much like a deliberative tic. And since there hasn't been any genuine exercise of agency at this point, for better or worse, there is nothing to assess. And so the adolescent has a claim to mitigation, a reduction in the amount of blame she deserves, because she has not actually failed to regulate herself appropriately. There is more to say about impairments such as adolescence, but first it will be helpful to add another case to the discussion. The similarities and differences will be illuminating. Let's turn to intoxication.

II.

While extreme intoxication may excuse a person from blame entirely,⁴⁷ many people have the intuition that mild to moderate intoxication might mitigate without excusing. The consumption of alcohol, for example, prompts change in the activity of the frontal lobe, both at the time of consumption and, with extended alcohol dependence, over time.⁴⁸ These brain changes impact deliberation in ways not unlike the adolescent's state impairs hers. These changes in brain state result in lowered inhibitions, reduced future-orientation, and so on. Within the deliberative process, this will result in instability in how the agent weighs her considerations as contenders for action. What it is to have "lowered inhibitions" is to be in a state where certain kinds of considerations are fragile and vulnerable to being overtaken by other sorts of considerations. Like the adolescent, the agent's deliberative process is both stemming from her agency and interrupted by an obstacle foreign to it. This impairment affects how she works with those considerations, and how she manages the considerations that she has. Mild-to-moderate

⁴⁷ Though this will be controversial given the origin of the intoxication. Many feel that the choice to drink to extreme drunkenness, for example, will perhaps open up the agent to moral assessment for the choices she makes while intoxicated. I will discuss this question further later in this chapter.

⁴⁸ Hamdy F. Moselhy, et al, "Frontal Lobe Changes in Alcoholism: A Review of the Literature."

intoxication will justify mitigating blame for the same reason as adolescence. The agent is no longer controlling the deliberation at the level of forming her intention. The obstacle she encounters – the effects of the intoxicant on her cognitive processing – makes it the case that she is no longer driving the deliberative process. The effects of the intoxication, in part, are driving her.

We can see this with an example. Let's say a person – call her The Intoxicated Agent – has had a few drinks at a party. She is not blacked-out drunk. She knows where she is and is partially, though unsteadily, in charge of her deliberative process. The Intoxicated Agent is in a committed monogamous relationship with her partner, and she has pledged vows of fidelity. During the course of the party, however, someone other than her partner sexually propositions her, and as a result of undervaluing the importance of keeping her promise and overvaluing the desire for excitement, she accepts the proposition and cheats on her partner. She has committed a moral wrong. Like the adolescent, her claim to mitigation depends on the particulars of her deliberative process. It is possible that the alcohol had no effect on her deliberation, and her action flowed uncomplicatedly from her agency to the expression of her action in the world. We can imagine a case where the intoxication didn't impact her deliberation in this particular instance, and if that is the case, just because she is biologically intoxicated doesn't mean she has a claim to mitigation. It will only be when the agent's intoxication actually presents impairment to deliberation that she will have mitigating circumstances.

Let's return then to the case where The Intoxicated Agent's intoxication does indeed corrupt a piece of her deliberation. Her deliberative process would look like this: on her deliberative field is the fact that she has made a promise to be faithful to her partner, along with an appreciation for the status of her partner as a moral peer and the knowledge that to break her

promise would be a moral wrong. There is also an attraction to the person who propositions her, and a desire to engage with them sexually. Under normal (or “cool”) circumstances for this agent, without the influence of intoxication, she would value the importance of her promise more highly than her desire, and she would not cheat. But in this case, The Intoxicated Agent’s deliberation is, in a sense, halted near the end. Intoxication reduced inhibitions, which is another way of saying that it makes the agent less responsive to considerations that require reflection, such as moral considerations, and more responsive to immediate desires. This is not to say that intoxication makes it the case that an agent will always choose immediate desires over moral (or other reflection-oriented) concerns. This is to say that intoxication alters the orientation of one stage of her deliberation. The Intoxicated Agent deserves blame for the parts of her deliberation that occurred uncorrupted, and does not deserve blame for the parts that the intoxication garbled. And thus, she deserves less blame than the agent who also cheated on her partner, but not under the influence of intoxication. That piece, the being “under the influence” of something that destabilizes deliberation, corrupts a stage of her deliberation and makes it ineligible for moral assessment, and thus, ineligible for the assignment of blame.

I would like to be clear about the point I am making. It would be easy to appeal to some variation of an “ought-implies-can” theory of responsibility and say that because these impaired agents can’t regulate their deliberation appropriately, it would be unfair (in the way Wallace uses the word) to blame them for the resulting deliberation. While some variation of this picture may be true – and indeed, though I offer no argumentation on this point, I think it probably is true – that is not the claim I am making. My argument about these impairments such as the ones I have described works like this: agency is not a single, unified state. It requires the capacity to regulate one’s behavior in light of moral reasons in many different ways. Impairments that prevent this

regulation altogether will be excusing, but impairments that prevent this regulation only at select points in the action-formation process will be mitigating, not because the agent can't do better, but because what is happening at the points of impairment are not genuinely pieces of deliberation. They are events that do not stem in any way from agency – hiccups interrupting speech – and morally assessing those events are like grammatically assessing hiccups. Agents with these impairments do not have mitigation for their failures because they couldn't have done better. They have mitigation because no choices are being made at the point of impairment, and so we have nothing to assess them on. We might think of the impaired agent as an archer taking aim at a target. She pulls her bow back, and right before she releases, someone pushes her. The arrow goes where it goes, but even though the archer took careful aim before releasing, it is not accurate to say that the arrow went where she meant it to go. If we try to evaluate her skill based on the destination of the arrow, we have made a fundamental mistake. The data is corrupted. We have no way to evaluate her based on her arrow's proximity to her target because an obstacle disrupted the relationship between her aim and her release.

This is not, however, to say we can't assess her preparation and aim before the disruption happened. This part of her performance was in every meaningful way hers, and if she aimed in the opposite direction of her goal, we can say a great deal about her skill. Blame is deserved with the performance of genuine moral failures. An absence of performance means that we have nothing to assess, and thus, it will be inappropriate to assess blame. With impairments such as adolescence and intoxication, the reduction in blame deserved for their performances does not stem from the agents having "good reasons" to do morally wrong things; cases like that will be discussed below. They deserve a reduction in blame because they did not perform a fully wrong

action at every stage of action. And this is because they did not genuinely perform every stage of action. They are vulnerable to moral assessment only for what is theirs.

III.

There is another lurking concern with intoxication, however, one that often influences intuitions about mitigation. And that is the question of whether the agent's choosing to become intoxicated doesn't contribute in some way to the amount of blame she deserves for wrong acts committed while under the influence. First, it is worth pointing out that intoxication can and does occur outside of a person choosing to become intoxicated. There are many ways that people become intoxicated through no direct or intentional action on their part. People can be involuntarily sedated after accidents, inhale toxic fumes, and so on. The deliberation-affecting features of intoxication must at times be regarded independently of any story about the agent *choosing* to become intoxicated. And when the circumstances surrounding intoxication look like this, the intoxicated agent should be morally assessed in the same manner as the adolescent; the obstacles affecting their deliberations are relevantly similar.

But those are not, to be sure, the most typical circumstances surrounding intoxication. The vast majority of people who commit moral wrongs while intoxicated intentionally chose to become intoxicated, and it is an open question how these circumstances affect mitigation. To answer this question, we must establish exactly what action such an agent is performing. Let's return to The Intoxicated Agent. This agent went to a party and consumed several alcoholic drinks, with knowledge that the consequence of her consumption would be intoxication. Later, when her deliberative process was corrupted by her moderate impairment, she broke her vow of monogamy with someone who propositioned her at the party. Our moral assessment will begin

by establishing which moral obligations bind her; in this case, she is bound by her promise of faithfulness to her partner. Next, we turn to her deliberation and the intention-formation process to establish the exact shape and nature of her moral violation, starting from the context of her promise. Here, we might see vastly different possible scenarios of deliberation relative to this obligation.

First, the decision to drink to intoxication might be directly connected to the end of breaking her vows. The Intoxicated Agent could have formed the intention to commit this moral wrong but judged it an easier task if she was drunk, perhaps, or she might be planning to use her intoxication as a cover for her guilt. In either of these cases, the decision to become intoxicated is no different than a murderer who decides to put a silencer on her gun. The choice to become intoxicated is a means to committing a morally wrong end, and it is part and parcel of the performance of the moral wrong. The murderer who uses a silencer does not have mitigating circumstances because her murder made less noise, and the agent who chooses to become intoxicated in service of performing the moral wrong does not have mitigating circumstances either. Here we see that the mere presence of intoxication is not sufficient for mitigation, even though the intoxication still corrupts the end of the deliberative process. The reason there is no mitigation, however, is not strictly because the agent chose to get intoxicated. The lack of mitigation is because the agent is deserving of full blame for the formation of the morally prohibited intention (in this case, deciding to break her promise), for that intention formation occurred independently of the corrupted deliberation. There is no mitigation because the intention to violate her obligation was formed in full independently of the later deliberation that becomes corrupted, and she showed maximal disrespect to the moral status of her partner in all parts of the deliberation vulnerable to moral assessment. Earlier in Chapter 3, I discussed cases

where an agent has mitigating circumstances because she fundamentally gets her obligation right at the level of the deliberative field and wrong at the level of deliberative content – she deliberates about correct considerations in incorrect ways. This version of the Intoxicated Agent performs her entire intention-formation process and gets it wrong from the beginning. She has merely taken action to eliminate the possibility of doing better in a separate deliberative process. This agent deserves maximal blame.

But this is not the only possible story for the Intoxicated Agent. We can imagine a case where her deliberation goes like this: she chooses to drink, knowing that it will lead to intoxication, but this agent does *not* do so as a means to break her vow. When she decides to drink, she has no impairments blocking her memory or understanding of her promise, and she knows that intoxication puts her at risk for making hasty sexual decisions that her sober self does not endorse. But she did not include that judgment on her deliberative field. She chose not to consider the possible impacts of her decision to drink on the obligations that bound her. And so she drank with the intention of become intoxicated, and then became impaired once the intoxication took effect. This agent has a different relationship to her intoxication than the agent who did not contribute in any way to her impairment (such as a person who accidentally inhales toxic fumes, for example), and with the agent who intentionally becomes impaired to facilitate the performance of a moral wrong. This version of the Intoxicated Agent case is one where the agent failed to properly consider the possible moral consequences of her actions, and as such, she deserves blame for that piece of deliberation. She did not populate her deliberative field appropriately. There was no impairment and no mitigating circumstances for the agent's choice to become intoxicated when that intoxication could reasonably lead to a risk of performing moral wrong. There is, however, something else to be said about this case. There is no necessary

connection between intoxication and promise-breaking. It is one thing to fail to consider something that in itself is a violation of one's moral obligation, a consideration concerning a direct and constitutive violation, and another thing to fail to consider a path that carries a *risk* of a violation. Built into the second is an additional piece of deliberation that is open to moral assessment. We only think that agents have obligations to include considerations involving risk-taking when we think that the risk-taking passes a certain threshold of acceptableness. When someone fails to adequately consider the risk of a moral harm, we call it a case of negligence.⁴⁹

Cases of negligence, at their core, are about taking morally unacceptable risks. We judge someone as morally negligent when we assess them as taking a risk that they should not reasonably have taken, one that is unacceptably high given the moral stakes. We say that not maintaining one's breaks is negligent (and worthy of blame) because part of responsible car ownership is taking into consideration both the possibility of brake failure and the harm that such a failure might cause. To drive without proper brake maintenance is to turn one's car into a weapon that could strike at any time, aimed at anyone, and that is morally impermissible.⁵⁰ The Intoxicated Agent who intentionally performs behavior that she knows will cause intoxication, for the sake of becoming intoxicated, is like the driver who does not maintain her breaks. She deserves full blame for her negligence, since the deliberative processes that lead to the intention-

⁴⁹ Establishing exactly how much blame this agent deserves will require sorting out what kind of harm this particular form of negligence is, and that will require determining what obligations agents have concerning the consideration of possible outcomes (and their likelihoods) of their actions. For more on how the role of what the agent knows plays into blame and culpability, see Holly Smith, "Culpable Ignorance."

⁵⁰ One of the fundamental difficulties of negligence is sorting out which risks are morally acceptable and which are not. Many of us agree that failing to maintain one's breaks is an act of negligence, but few people believe that driving at all – even though the mere act of getting behind the wheel puts self and others at risk – is negligent. Sorting out just what risks are morally reasonable and which ones are not is a project worthy of attention all of its own. I think it is contentious that all acts of intentional intoxication are negligent because they necessarily weaken one's reflective capacities and moral resolve, but perhaps such a case could be made. Regardless, in the case I am describing here, I have in mind a case where the agent has good reasons to think her choice to become intoxicated will lead to a significant increase in the likelihood that she would break her vows, significant enough that this concern should have been included in her deliberative process, and significant enough that weighing her moral obligations appropriately would have required her choosing not to become intoxicated in these particular circumstances.

formation relevant to the negligent choice was unimpaired and non-corrupted by any deliberative interruptions. What comes after the intoxication, corrupted by the results of that negligence, is still not genuine deliberation. The Intoxicated Agent did not form an intention to break her vows *because* of her intoxication. The intoxication undermined part of her deliberative process, preventing the piece of not-deliberation from being properly open to moral assessment. So we end with an agent who deserves full blame for her negligent action, since the entire process of her deliberation and intention-formation process that led to the negligence was hers, and because part of the appropriate regulation demanded by moral obligations is not to be negligent.

Nevertheless, negligence is a very different kind of wrong than full and explicit violations of obligation, and so will be worth different amount of blame. But the negligent agent will have the same claim to mitigation for the broken vows as the Intoxicated Agent whose intoxication was through no fault of her own. Blame does not compound like interest. Her choice to become intoxicated was not directly connected to an intention to break her vows. And thus, the part of deliberation relative to her negligence gets considered in moral assessment. But that negligence does not make the negligent agent deserve as much blame as someone who becomes intoxicated maliciously, as a means to the overt disrespecting of her promise. Being morally responsible for the existence of one's impairment does not undermine mitigation, if the actions that cause the impairment are not deliberately connected – part of the same intention-formation process – as the wrong performed. It only means that one is vulnerable to moral assessment for the intention to bring those circumstances about. A crater might be caused by an intentional bombing, or it might be caused by an asteroid. Moral assessment will require us to assign blame for the metaphorical bombing and not for the metaphorical asteroid, but when it comes to assigning blame for the results – the corrupted parts of deliberation -- there is still just a crater, regardless

of its origin, and it is not appropriate to assign blame for deliberation that doesn't genuinely belong to the agent.

IV.

But there are other kinds of circumstances that many of us intuitively believe are mitigating. Adolescence and intoxication mitigate because they undermine the performance of deliberation. By making a piece of the deliberative process unstable, they make it such that part of the deliberative process originates from and is genuinely the result of agency, and part of what appears to be deliberation is not genuine deliberation at all; it is instead the product of the agent losing control over the process. There are other ways, however, that deliberative obstacles can affect deliberation, and thus affect moral assessment. One circumstance that is often linked with mitigation is the presence of strong emotion. Statutes that distinguish second-degree murder as "crimes of passion" seem to be playing on this intuition, and so I will now turn to circumstances that are characterized by a "loud" subjective experience for the agent deliberating. I will call these circumstances passion-inducing.

Passion-inducing circumstances come in a broad variety of types. Their unifying feature is that they create desires or other emotional responses that inappropriately command attention from other deliberative considerations. There is a subjective intensity attached to these circumstances that is lacking in cases of normal deliberation. The extreme anger of the Betrayed Lover from Chapter 1 might fit here (depending on how we describe that case), as will cases of desire like Rapunzel's mother from the Grimm Brother's fairy tale, whose pregnancy craving for rampion led to the forfeiture of her daughter to an evil witch.⁵¹ Some argue that this intensity of

⁵¹ Jacob and Wilhelm Grimm, "Rapunzel" in *Household Tales*.

desire is one of the distinctive features of addiction, characterizing addition as being gripped by subjective experiences of intense and extraordinarily powerful craving.

One such proponent of this view of addiction is Stephen J. Morse.⁵² While Morse acknowledges that addiction may create other morally relevant circumstances due to brain restructuring and other biological changes (changes which would either be mitigating or excusing impairments, depending on their nature and severity),⁵³ he argues that the subjective experience of an addict's cravings also affects the moral assessment of their actions. Morse rejects the notion that addicts are compelled by their cravings, and that there is no genuine deliberation taking place when someone is in the grip of such a craving.⁵⁴ Addicts, he argues, are typically responsive to reasons, and this implies that addiction isn't like falling off of a cliff – the addict is still, in some sense, acting. Morse argues that the subjective experience of craving itself impacts, or corrupts, the agent's ability to deliberate well. He compares it to being near the point of orgasm, without the pleasurable sexual feelings. When the desire for their drug mounts and nears its crescendo, it becomes a kind of reflection-inhibiting distraction. He writes that for addicts, "Fundamental components of rationality – the capacities to think clearly and self-consciously to evaluate one's conduct – are compromised. The agent may not recognize the various options at all or may not be able coherently to weigh and assess those that are recognized."⁵⁵

For all of these passion-inducing circumstances, the unique deliberative obstacle isn't that a part of the deliberative capacity has gone awry; it is that the subjective experience of the

⁵² Stephen J. Morse, (2000). "Hooked on Hype: Addiction and Responsibility."

⁵³ The origin story for the addiction might matter as well, in the way that it might matter for the source of intoxication. We will assess the addict differently if her addiction is something she intentionally cultivated for morally wrong ends, or if she cultivated it through negligence, or if it came about through no intentional action on her part whatsoever.

⁵⁴ Morse, p. 13.

⁵⁵ Morse, p. 39.

passion is preventing the agent from using her deliberative capacity in the manner that genuinely reflects her orientation towards her considerations. We might imagine a driver who is navigating a curvy road while her radio is stuck at top volume. The radio isn't directly interfering with the driver's driving capacity itself; she can still see the road ahead, recognize that she must turn the wheel to keep pace with the curved road, and apply gas and brakes at appropriate times. But the presence of loud music perhaps interferes in her ability to do these things well. The distraction perhaps captures her attention periodically, shifting her focus from the road ahead to the music on the radio. Its volume inserts itself into her thoughts and captures her attention, and it distracts her from her task in a way that she did not choose. She values stopping at the appropriate time far more than she values listening to music, but the blaring radio distracting her makes her slower to break. If this driver crashed on the curvy road, we'd certainly tell some story about how her malfunctioning radio contributed to the situation.

Passion-inducing circumstances have a distinctive character to them, something that is harder to resist than a run-of-the-mill emotion or desire. There is something especially difficult about constant noise, for example, and its effect on deliberation is different than, say, an equally strong feeling of boredom. Being at the point of orgasm is one kind of distraction, and having a word on the tip of your tongue is another, and not just of degree. There is a sense in which certain kinds of distractions are just more viscerally powerful than others. They have the ability to command our attention in ways that other distractions – even of the same degree – simply do not. They are fundamentally disorienting.

Strong passions create irregularities within deliberation in a different way than impairments. Where impairments undermine the agent's ability to perform parts of deliberation, passions act as deliberation corrupters in a different way. Impairments block the ability of the

deliberative capacity in some way – they weaken or impair the muscle of deliberation. Passions affect deliberation by demanding attention, whether the agent judges that the passion merits attention or not. They are like trying to use our perfectly functioning muscles on a planet with a much greater gravitational force. Whether we try to walk, run, or jump, gravity will pin us to the ground regardless. Intense sensory experience is inherently disorienting because it severs the relationship between choice and focus. Part of the regulation demanded by moral obligation is appropriate focus; this is part of what it is to place something on the deliberative field and take it seriously as reasons-generating. It is to judge it, acknowledge it, as attention-worthy. Strong passions corrupt this process by demanding attention independently of any judgment from the agent. They do not corrupt the system of deliberation. They are corrupted data entered into the system.

Intentions formed in the presence of these passions include genuine deliberation. The agent does indeed populate her deliberative field and weigh her considerations. But the content of what she considers has been, in a manner of speaking, tampered with. And so the resulting product of the deliberation does not represent the genuine judgments of the agent. Rapunzel's mother does not genuinely endorse the fulfillment of her pregnancy craving over the lives of her husband and daughter. In commanding more consideration than she would otherwise give it, the passion makes it the case that the resulting decisions are not, in a genuine sense, fully hers. The passions create deliberative irregularity because they contain instances of genuine deliberation that stem authentically from the agent, and instances of deliberation that do not. Passions intrude in deliberation; it is not accurate to say that the agent deliberates upon them. The agent deserves blame for the violations and failures to regulate that are genuinely hers, which stem from her agency. She does not deserve blame for deliberative performances that happen *through* her,

without her judgment. Regulation is, by definition, an intentional organizing and constraining of behavior. But passions play on a quirk of human attention. They prevent the agent from fully dismissing them, even when she judges them worth dismissing, and so the resulting failures to dismiss are not genuinely hers. Blame is deserved by agents who fail to perform moral obligations, and agents do not deserve blame for what they did not do.

Another category that can corrupt the deliberative process is circumstances that require extreme urgency, or snap decisions. Extreme time constraints or rushed circumstances may also act as conditions that necessarily drive focus and attention away from where the agent judges they belong. Urgency requires a kind of deliberative regulation that is external to the requirements of one's obligations. Appropriate acknowledgment of a person's moral status as the grounding of an obligation exists independently of time constraints. No one deserves acknowledgment *because* time is short or conditions are urgent. Their status makes it the case that they deserve acknowledgement. Urgency just places constraints on the time the agent has to deliberate about that fact. And so now the agent faces the task of regulating deliberation and intention-formation according to two regulative principles: the demands of her obligation and the demands of her time constraints. There is nothing problematic about this on its face – part of the content of our duties is to regulate our deliberation and action-formation in appropriate ways in the real world, and sitting down and pondering whether to save a drowning child for a good hour – just to make really sure you get it right – is a way of avoiding the morally appropriate decision. The child will drown while you naval gaze. Part of acknowledging the status of the other person is coming to judgments about what will be a genuine respecting of their status.

But urgency is something different. The cases I have in mind require what I will call non-reasonable time-frames for deliberation and intention-formation. One of your friends has been

bitten by a poisonous snake, and she will die if she does not receive the antidote. You have one vial of the antidote, but it's rare and precious, and you were planning to sell it to the highest bidder. You have enough time to snap off the lid and inject her before it's too late, but just barely. I will stipulate that the morally correct option is to inject your friend, but I will argue that you have a claim to mitigation if you do not. Urgency requires the prioritization of a system of deliberation and intention-formation regulation that is not based on morally relevant features. Something must be done or the opportunity for the action will pass. Your friend will no longer require help if she dies. But the time-frame, an obligation-independent factor, has necessarily inserted itself into your deliberative process. Like a passion, urgency means that timing demands attention, even if the agent does not hold the judgment that time concerns are more attention-worthy than her obligation. In fact, with extreme enough urgency, the only possible deliberative regulation might be to abort: to not deliberate at all. These are cases where people essentially freeze. I am inclined to think of these as cases of full excuse, not mitigation. But in cases where some deliberation is possible, the product of that deliberation will be necessarily regulated by factors external to the agent's judgment. And so even if we think of this as genuine deliberation, it is not deliberation that stems in the correct way from the agent.

V.

But there are other circumstances that many people believe are mitigating that aren't yet captured by the conditions I've discussed. In particular, let's return to the Betrayed Lover case I introduced in Chapter 1. In this case, a woman returns home to find her lover in bed with someone else, and feeling jealousy, anger and resentment, she picks up a nearby vase and kills him with it. The Betrayed Lover might have a claim to mitigation under the category of passion-

inducing circumstances that I described above. It might be the case that her emotion is so strong, so visceral, that it in a sense overran her own deliberation. But I don't think that is the only way to understand this case. A strong emotion will not necessarily impact deliberation in the same way as the cravings of addiction. The distinctive feature of the passion-inducing circumstances I describe above is their quasi-sensory way of demanding attention and consideration independently of whether the agent judges them worthy of attention. These cases are like the driver and the stuck radio, who is desperately trying to focus on her driving *in spite of* the noise she is encountering. This is a very different kind of case than a driver who is distracted by a radio that she turned on, and that she can turn off. This is not to say that we can turn our emotions on and off at will, of course, but that agents might have a very different relation to their emotions than they do to their cravings. Emotions such as the anger and resentment in the Betrayed Lover case are the results of certain judgments of the agent – they result from and are part of the deliberative and intention-formation process, and part of what is demanded by moral obligations is the requirement to regulate those emotions appropriately.

It is contentious whether we are morally responsible and open to assessment for the emotions and attitudes we hold, and this is not a debate I currently stake a position in.⁵⁶ I am, however, making the claim that we are properly morally accountable for how we choose to treat our emotions within our deliberative processes. We choose whether to consider those emotions as reason-generating, and we choose how strongly to weigh them against other considerations. This is part of the process of intentional action, and it is part of what we must regulate appropriately in meeting our moral demands. Passion-inducing circumstances play on a trick in human cognition, making us, in a manner of speaking, unable to turn away. But this is not true

⁵⁶ For a nice discussion of this topic, see Angela Smith's "Responsibility for Attitudes: Activity and Passivity in Mental Life."

for most cases of strong emotion. Anger is a tricky emotion to categorize. We have expressions such as “blinding rage” in common usage, which certainly suggests that in its extremes, it should be categorized as passion-inducing. But if that’s the case, then interesting questions surrounding the origin and cultivation of such emotions become relevant, just as questions about the origins of intoxication become relevant. If strong emotions like rage are passion-inducing, then it might be a form of negligence to cultivate one’s character in a way such that one is disposed to rage. We might think that if someone makes the decision to pick heated fights for fun, or get involved in situations where aggression is celebrated, one is doing the equivalent of drinking to intoxication. The agent might have mitigating circumstances when the rage hijacks her deliberation, but she is fully blameworthy for the negligence that led to that hijacking.

I believe that intuitions such as these are behind many views about “out of character” actions.⁵⁷ It is not entirely clear what people mean when they say an action is “out of character.” In one sense, if the case we are discussing is a case of genuine action where the agent did indeed perform the action, then such a phrasing appears absurd. The action cannot be “out of character” for an agent, because she (and her character) genuinely performed the action. This cannot be literally true. And of course, there are other problems with taking this description at its face. The first morally horrific act someone performs will, in one sense, be “out of character,” as will the early bad acts people perform that shape their character accordingly. This cannot be what people are trying to intuitively capture.

Instead, I suspect people use this term to mean something like “the agent did not cultivate the disposition that would normally motivate this bad act.” They mean that the moral violation in

⁵⁷ For examples of this defense in legal contexts, see Modal Penal Code, § 2-3, on duress, excuse, and mitigation. The force of duress appealed to here seems to be that the agent acted in a way that does not reflect his genuine character. It might be that part of the force of many cases of duress comes from the involvement of another agent, and part comes from the fact that the action in question does not seem to cohere in some way with the character of the agent in question.

question is in conflict with the other judgments and commitments that the agent holds, and at the level of disposition, they mean that the agent has not courted episodes such as these through negligence (or intentional malice). There is no reason, however, to think that this in itself is mitigating. The moral assessment of action and the moral assessment of actors are different processes, and while the first contributes to the second in very direct ways (evaluating the shape of intentional actions that an agent performs tells us about what kinds of conditions, over time, she takes to be reasons-generating and motive-worthy, which illustrates deep judgments and commitments she holds independently of any given action), the relationship is not as direct in the other direction. An agent's character qua character only becomes relevant to the moral assessment of action when some element of that character plays a part in the action-formation process. If an agent has a disposition to be cruel, and this generates a desire not to help a drowning child because such an action will amuse her, we can say that she is deserving of blame for not populating the deliberative field appropriately, and for under-weighting the content of her obligation to aid. We should not use elements of the agent's character that are not relevant to particular action in morally assessing that action, because we use the assessment of actions to make judgments about character. It would not make sense to say "She is disposed towards kindness" if there was no evidence of her performing any intentional acts of kindness. To use character to make assessments about action would make the assessment processes circular, and thus meaningless.

While I don't take "out of character" actions to be a mitigating category on its own, I do think that people often consider cases that are genuinely mitigating under this rubric. People speak of acts impaired by passion-inducing circumstances as "out of character," and in a way they are correct. Those actions do not genuinely reflect the judgment of the agent. But people

often speak about cases like the Betrayed Lover case under this rubric as well, and I'd like to offer a way to understand the mitigating circumstances of that case in a different way than just involving passion. Another salient factor is that the Betrayed Lover did what she did for "good" reasons, or at least, reasons that don't stem from morally inappropriate considerations. To see what this kind of mitigation might look like, let's turn to a real-life case.

In February of 2008, Barbara Sheehan, a quiet school secretary, shot her husband while he stood at the sink shaving.⁵⁸ She shot him eleven times, with two guns. Police discovered him dead in a pool of blood, water from the faucet still running. *The New York Times* described Sheehan as "a churchgoing mother of two who wears sober gray suits." Defense attorneys argued that Sheehan was frequently abused and battered and a long-time victim of domestic violence at the hands of her husband. New York law holds that killing in the face of an immediate threat is justified and legally permissible, but Sheehan was not facing an immediate threat. Her husband was at the sink shaving. And lawyers on both sides debated the question of whether domestic violence could justify a homicide even when no immediate threat was imminent. Since the defense could not appeal to self-defense at the moment of the act, they instead argued something like this: Sheehan had good (meaning morally appropriate) reasons to do what she did. And they argued that this should serve as full justification for her action.

None of us know the internal deliberation and intention-formation process that was part of Sheehan's act of killing her husband. It may be the case that the fear of constant abuse gave her an excusing condition altogether, or that ongoing abuse gives us a reason to think that the "immediate" clause currently connected to self-defense justifications is in error. But that is not the part of the case I will focus on. For my purposes, the part of the case that is most interesting

⁵⁸ Quotations and details of this case from Dan Bilefsky, "Murder Trial Hinges on Questions of Domestic Abuse," *The New York Times*, September 18, 2011, p. A23 of the New York edition.

is the attempt of both sides to characterize the nature of Sheehan's character in order to draw conclusions about her reasons for acting. The defense characterized Sheehan as a kind, quiet woman, beaten down from years of physical and mental abuse. The prosecution argued that she simply despised her husband after a long and unpleasant marriage, and that she was a hateful and angry person. The idea behind both approaches was this: talking about someone's character will tell you something about the reasons they acted in a given case, and those reasons matter. In this case, it is Sheehan's status as victim that the defense argued was especially relevant. As I've argued above, I do not think a direct appeal to character is appropriate when it comes to the moral assessment of actions, but I do think it worth discussing the general method of this approach, constrained to the description of action. In particular, how are we to think about people who commit moral wrongs (at least in part) because they themselves have been wronged? I will argue that such cases might indeed contain deliberative irregularity, and such circumstances will be mitigating.

Being the victim of a moral wrong alters the scope of morally justified – and thus morally relevant – considerations. Specifically, being the victim of a moral wrong makes it the case that one is morally justified in feeling anger and resentment towards the perpetrator. This can affect the moral assessment process in the following way. If an agent is bound by a moral obligation to aid someone trapped in a burning building, normal circumstances will demand she take his moral status as reason-generating, place that consideration on her deliberative field, populate her deliberative field with only other morally relevant considerations, and then weigh those considerations appropriately. There are many considerations that would be inappropriate to place on the deliberative field in this deliberative process. If the victim is of a race that the agent is exceptionally bigoted against, and she places the fact of his race (and her ill will towards

members of his race) on her deliberative field, she has performed a blameworthy sin. Her lack of good will towards the victim is not morally justified, and it has no place in this deliberation.

But circumstances change if this particular victim had significantly⁵⁹ wronged her.⁶⁰ In this case, the agent now has good moral reasons to withhold good will from the current victim, and these good moral reasons are relevant considerations that are appropriate to place on the deliberative field. When an agent is wronged by another, certain moral responses are not only acceptable, they themselves are a display of respect for one's own status as an agent. If someone maliciously stabs me, for example, and I feel no indignation, or anger, or in Strawsonian terms, resentment, I am either not properly acknowledging my own status as the source of the perpetrator's moral obligations, or not properly acknowledging the perpetrator's status as an actor (or both). When I fail to get angry at the stabbing, I am either displaying a judgment that there is nothing to get angry about, that no obligation bound the agent because I am not the right kind of creature to ground such an obligation, or I am treating the agent as something not of the appropriate sort to be bound. Respecting both my own status and the status of the perpetrator requires I react appropriately, and in this case, the appropriate reaction will be some attitude that indicates a reduction in good will towards the now-victim.

Whereas the racist made an entirely inappropriate decision to populate her deliberative field with her emotional response/reduction in good will, it is entirely appropriate for the wronged agent to consider her emotional response/reduction in good will on her deliberative

⁵⁹ I use the word "significantly" because the scope of the wrong will determine whether the lack of good will she has towards the victim is justified. If you didn't return my wave hello this morning, I might feel slighted and resentful. But if I consider that resentment when deliberating about whether to save your life, I have made an enormous error in scope. The wrong in question will have to be of significant magnitude, just as coercive threats must be of significant magnitude to count as genuine coercion.

⁶⁰ I think it is open as to whether the moral wrong had to be committed towards her to justify emotions like resentment, or whether a wrong being performed towards a close loved one, for example, will count. The defining feature of this circumstance is that it will appropriately ground a withholding of uncomplicated good will towards the victim, either through anger, resentment, or a similar emotion.

field. The wronged agent's response is a moral response, something demanded by both her and the victim's statuses. It is of the right type to place on the deliberative field, the right kind of consideration to weigh against her moral obligation to aid. And so if the agent mis-weighs these relevant considerations and performs a morally wrong action, then her failure occurs only at the final stage of deliberation, and she still succeeds at the level of the deliberative field. This gives her an appeal to mitigation that looks similar to cases of coercion. The agent performs the wrong action, but her mistake happens with the handling of moral overlap.

And finally, we can start to see the full picture of why the Betrayed Lover has a claim to mitigation. Independently of the possibility of passion, strong emotion, or urgency, The Betrayed Love was betrayed. She was morally correct to feel anger towards her lover, and to rescind the good will she once had towards him. Behavior that would in other circumstances be morally problematic – like kicking him out of the house, for example – are now justified by her circumstances. And so on her deliberative field we might see something like this: she knows she has an obligation not to kill him, or any other moral agent. She takes that to be reason-generating, and she places it upon her deliberative field. But she also places her anger and her ill will for the wrong, along with a desire to make him understand how badly he has made her suffer. She deliberates about these conditions before acting on her anger. She has committed a moral wrong, but the consideration that led to her intention wasn't itself morally wrong. It was just mis-weighed against something that should not have been undervalued. She performed a moral failure at the level of her deliberative content, and she is to blame for that failure. But she is not to blame for considering her anger and lack of good will in the first place. Allowing those considerations on the deliberative field was a way of showing respect to both herself and the victim qua agents.

The Betrayed Lover case illustrates an interesting feature of mitigation. It can be messy. There can be multiple reasons for moral success or failure at any deliberative point, and there might be inner-deliberative impairments and agency-blocking conditions, all in the same deliberative performance. This makes it extraordinarily hard to tease out all the relevant features for genuine moral assessment, but this way of breaking down the deliberative process for assessment at least allows us to make sense of the project. In the next and final chapter, I will be discussing where this approach is situated in the larger landscape of moral assessment. I will argue that my approach to mitigation compliments existing theories of excuse and justification, filling explanatory gaps and giving more robust content to the notion of mitigation.

CHAPTER 5: EXCUSE, JUSTIFICATION, AND MITIGATION

In the past four chapters, I have argued for a picture of moral mitigation focusing on the moral assessment of deep deliberation and the fullness of action performance. In this, the final chapter of the work, I will situate this view within the larger body of work on complications in moral assessment. In particular, I will argue that theories of excuse, exception, and justification are not on their own sufficient to capture the force of moral mitigation, and that my picture helps fill this gap. Taken as a whole, these theories will get us closer to a complete and accurate picture of moral assessment. In the final part of this chapter, I will discuss what I take to be my successes and failures in this endeavor, where this leaves moral assessment, and where the project can go from here.

I.

Within philosophical literature, many have thought the nature of mitigation could be understood by an adequately sophisticated theory of excuses. That is, they've thought that the normative force of mitigating circumstances is of the same type (but to a lesser degree) than what gives excuses their normative force. To understand the work my theory is doing, it will be helpful to explore this claim. A good starting point of contemporary work on excuse is J.L. Austin's "A Plea for Excuses."⁶¹ Austin began with the assumption that having an excuse for moral wrongs was possible, and then he went to work arguing that a thorough investigation of excuses would yield insight helpful to both ethics and philosophy of language. Austin starts his discussion by making a distinction between excuse and justification. A justification, he claims, is to admit that someone actually committed a moral wrong, but to argue that in this case, it wasn't

⁶¹ Austin, J.L. "A Plea for Excuses."

actually morally wrong. That is, perhaps we might say that an agent, call her the Knife-Wielder, stabbed another person in the neck. It is true that she did it; all of her faculties were functioning normally, no one forced her hand, she was not being controlled by an evil scientist, etcetera. And it is also true that, given no special story, going around stabbing people in their necks is morally wrong. But in the Knife-Wielder's case, there *was* a special story. She is a trained paramedic, and she had stumbled upon someone miles from civilization, throat closed from a severe anaphylactic reaction. She stabbed the person in the neck to open his airway and perform a tracheotomy. What would normally be morally wrong, a violation of an obligation not to harm, was in this case no such violation. There were special circumstances that made this action permissible (and perhaps, even morally praiseworthy). Those circumstances served as a justification for her action, making it the case that though she was fully morally responsible for the action in question, she is not deserving of blame or censure.

Austin contrasts these cases of justification with cases of excuse. The issue of justification is whether the action in question is actually morally wrong. Excuses take the moral wrongness of the given action as granted, but these circumstances show that perhaps the agent didn't really and truly (in Austin's words, "baldly"⁶²) *do* the action. Austin writes

"We may say it isn't fair just to say X did it; perhaps he was nudged. Or it isn't fair to say baldly he did A; it may have been partly accidental, or an unintentional slip. Or, it isn't fair to say he did simply A – he was really doing something quite different and A was only incidental, or he was looking at the whole thing quite differently."⁶³

This notion of excuse will cover cases like the Innocent Knife-Wielder, another throat-stabber. In her case, she has no special story to tell that makes throat-stabbing morally permissible. She is not a paramedic, and the victim was not facing death without an emergency tracheotomy. The

⁶² Austin, p.2

⁶³ Ibid.

action was wrong, full-stop. But there is something unique about the circumstances in her case. Let's say it is true that she held the knife that stabbed the victim's neck. But her arm was being controlled by another, much stronger woman, and it was this interloper who propelled the knife held in the Innocent Knife-Wielder's hand into the victim's neck. So yes, the victim was stabbed, and yes, stabbing people is wrong. But it appears unfair to say that the Innocent Knife-Wielder actually *did* the stabbing.

That said, Austin's notion of excuse covers a lot of ground. Excuses will apply to cases like the Innocent Knife-Wielder, but also to cases of accidents, incidental actions, unintentional actions, cases of "looking at things differently" (however that gets cashed out), and other cases where the direct agency of the person in question is somehow compromised. This group is far from homogenous. Austin gives examples like dropping a tea tray because of a wasp as legitimate cases of excuse,⁶⁴ and this seems to be in keeping with his inclusion of accidents in this grouping. This presents an obvious difficulty: how exactly are accidents, involuntary moments, and so on, cases of action in the first place? To be more specific, the notion of an excuse presupposes that a wrongful act has been performed, and its distinctive feature is that *despite* the fact that a wrongful act has been performed, the agent is not morally culpable. If excuses function by indicating that the action in question wasn't really an action, then it seems like excuses, as I've described them, are not actually possible.

Wallace runs into this issue when describing Austin's view. He glosses Austin as claiming that the unifying force of excuse is a lack of intentionality.⁶⁵ This blunt view can't be quite what Austin had in mind, if for no other reason than he was taking it as given that excuses are indeed possible. But we have additional reasons to think Austin was working towards

⁶⁴ Austin, p. 2.

⁶⁵ Wallace, p. 120.

something more subtle, namely his use of qualifiers around the verbs in question. He did not claim that the excused agent did not *do* X. He claimed that the excused agent did not *baldly* do X. That is to say, he claimed that there is something about the story that removed moral culpability from the agent, even when a morally wrong action was actually performed. In other words, his story of excuse must be that the story of action is complicated somehow in these cases, and this complication has moral force which makes it inappropriate to hold the agent in question blameworthy. If that's Austin's claim, then we are left with the deeper question. What exactly are these complications, and from where do they derive their force? By saying the story is more complicated than a mere lack of intention, we regain the intelligibility (and possibility) of excuse. But we're left with an even more puzzling set of questions.

Wallace takes up this project in the fifth and sixth chapters of his book.⁶⁶ He begins with the broad distinction that Austin makes, between excuses and justifications, and then sets aside the question of justification. Once he turns to the notion of excuse – circumstances where we acknowledge that a given action is indeed morally wrong but have some reason to consider the agent's responsibility inhibited – he makes a further distinction. He identifies two types of circumstances that fit this rough notion.⁶⁷ First, there are circumstances which inhibit responsibility locally, and second, there are circumstances that inhibit responsibility globally. These also track the categories that Strawson identifies as inhibitors of reactive attitudes.⁶⁸ To understand why these notions don't, on their own, explain the force of mitigation, it will help to examine these conditions in more depth.

⁶⁶ Wallace, pp. 118-194.

⁶⁷ Wallace, p. 119.

⁶⁸ Strawson, pp. 64-65.

I will begin with Wallace's first distinction: circumstances that inhibit responsibility locally. Wallace refers to these (and only these) conditions as excuses. He gives several examples of what he has in mind with these cases: inadvertence, mistakes, accidents, unintentional body movements, physical constraint, coercion, necessity, and duress. Like Strawson, Wallace believes that morality requires members of the moral community to regard each other with certain kinds of attitudes, and blame attaches to the failure to meet these obligations. He writes, "We demand that people should display toward each other a degree of respect, consideration, kindness, and so on, and we tend to resent or blame people to the extent that they fail to satisfy these demands. Excusing conditions, then, show that a person who appeared to violate such expectations did not in fact violate them after all."⁶⁹ In other words, Wallace believes that these local responsibility-inhibitors (excuses) function by showing that underneath the basic action description, the agent did not actually fail to display the appropriate attitude towards her peer.

How this works with cases of inadvertences, mistakes, and accidents is quite straightforward. These, Wallace argues, are not actually performed intentionally at all. If I take someone's lunch from the communal refrigerator, it appears on first pass that I am violating my obligation to regard that person with the respect that she deserves. But if I take the lunch by mistake – her lunch bag looks just like mine, and I had forgotten just what I'd packed – then the assessment of the action changes. It's true that I performed an action, I removed a lunch bag from the refrigerator. But only knowingly removing *her* bag from the fridge would be showing her something less than the respect she deserves. This is because displays of attitude and respect require me to have certain beliefs about the content of my actions, and if I do not believe that what I am doing is stealing her lunch (in fact, I am unaware she has anything to do with the bag

⁶⁹ Wallace, p. 125.

in my hand), then it is not possible for that action to reflect anything of my attitudes towards her. Wallace argues that this is because our actions importantly reflect our choices, and when circumstances are such that we are not performing the action that we think we are, then our action does not actually reflect our choice. He captures this by writing “intentions are sensitive to beliefs. To do something of a certain kind intentionally, one must know that one is doing something of that kind. More precisely, to do something of the kind x, as the result of the choice to do something of the kind x, one must believe that what one is undertaking to do is of the kind x, at the time that one makes the choice to do it.”⁷⁰ In other words, if I don’t know that I’m doing something of kind x, then I didn’t *choose* to do something of kind x. And if I didn’t choose to do something of kind x, then I am not actually displaying the attitudes indicated by kind x. And if I am not displaying the attitudes indicated by kind x, then I am not failing to meet the obligation not to regard my moral peers in inappropriate ways. This is essentially treating excusing conditions as conditions that force action redescription.

It is clear how cases of inadvertence, mistakes, or accidents block appropriate assessments of blame in these cases. But it is not clear if we should actually regard these as cases of excuse. It’s not clear that all of these are cases of genuine action in the first place. It is true that when I take a lunch-bag from the refrigerator that I am performing an action. But I am not performing the action in question. The action for which I am supposedly excused from blame is stealing my colleague’s lunch, and it is unclear how we should consider me performing *that* action at all. I am doing something else entirely. But this becomes even more clear when we include cases such as tripping, falling, dropping things, involuntary body movements, and so on. Wallace seems to think that these cases legitimately fall under his category of excuse. Within this

⁷⁰ Wallace, p. 136.

category, he describes epileptic fits, sleepwalking, and so on.⁷¹ He argues that these aren't intentional actions, and thus don't reflect any kind of choice on the agent's part, and thus they excuse the agent from blame. It is unclear exactly why he thinks cases like this belong under the rubric of excuse at all, since it seems that there is no morally impermissible action that needs to be excused. In other words, it is not clear what these cases have in common with other cases of excuse, since these seem to disqualify the event from moral assessment earlier in our consideration. We don't need an excusing condition to tell us why blame doesn't attach. It fails to be the appropriate thing to consider by failing to be an action under any description. Thus, the inclusion of these cases is a bit mysterious, and perhaps, misleading. The force of Wallace's picture of excuse is much more persuasive when we turn to his other examples.

Cases of physical constraint, for example, are much more interesting on Wallace's picture.⁷² These are cases where an agent is unable to fulfill her moral obligation because she is physically limited in some way, either externally (say, by someone holding her hands down, or being tied to a chair) or internally (say, paralysis or muscle exhaustion). In this category, we can imagine cases where an agent makes the relevant choice to do, or not do, what she should. Let's say she hears someone begging to be rescued from an attacker. If the agent chose to ignore the cries, she would be failing to meet the appropriate expectations for her – she'd be violating a moral obligation to show the appropriate kindness and respect to her moral peer. In this case, she would be both fully responsible and fully deserving of blame for her action. But if we re-imagine the case slightly, the interesting features of Wallace's view come into sharper focus. Suppose the agent did indeed fail to come to her neighbor's aid. But in this case, say she formed the appropriate intention to help (indeed, in Wallace's terms, she made the appropriate choice), and

⁷¹ Wallace, p. 140.

⁷² Wallace, pp. 141-143.

as she stood up to help, she realized that her wrists and ankles were cuffed together. In this case, the agent fails to do what she should, to help, but that action does not reflect her choice, and thus does not reflect an inappropriate attitude towards her peer. Cases of excuse by physical constraint are one of the most straight-forwardly appealing cases on Wallace's picture, since the break between the attitude of the agent and the action that she performs comes about after the appropriate intention has been formed. In other words, the problem is not merely that the agent hasn't formed an attitude (or made a choice) about the situation in question. She has done so, but she has been thwarted in her attempt to turn that choice into action. And so these are the clearest possible cases where action does not accurately reflect the attitude of the agent.

On Wallace's view, the thing that unifies all of these cases – the whole body of excuses in general – is that in each case, the agent is not actually making the choice that it appears. He writes, "Whenever people are genuinely excused from responsibility for their actions, those actions will not have been morally wrong. Obligations regulate the choices that are expressed in action, but when a valid excuse obtains, it turns out that what an agent has done did not express a choice at odds with the moral obligations to which we hold that agent."⁷³ And so, it will be inappropriate to blame an agent for a choice she did not make. Thus, excuses serve as "special circumstances" that force us to take a second look at the action in question.

So far, we've been speaking only of circumstances that block blame locally, for a particular action. Austin, Strawson, and Wallace all consider cases where the blocking force is global. That is, there are circumstances that make it the case that a person ought not be considered an appropriate target of blame at all, regardless of what it appears she is and is not doing. While Austin still considers these cases of excuses, Wallace, adopting the terminology of

⁷³ Wallace, p. 147.

Gary Watson,⁷⁴ refers to these conditions as exemptions.⁷⁵ Conditions that succeed as exceptions will make it the case that the person in question doesn't meet the requirements for accountability, and so instead of discussing (or redescribing) a particular action, exemptions force us to look at the (possible) agent herself. Common examples of this include insanity, childhood, brain-damage, and so on. These are all conditions that block the person in question from meeting the requirements for accountability, and thus, even if such a person performed an action that would be paradigmatically blameworthy if performed by a normal agent, her circumstances will prevent us from considering her responsible at all. This is important to note, because exceptions will inhibit blameworthiness even for cases where an action was performed intentionally, with all the other normal hallmarks of paradigmatic cases of action. The special circumstances will not be tied to the action itself, it will be tied to something distinctive about the *person* performing the action.

There are several ways to think about cases like these. We might think that the force of exempting conditions comes from the fact that they prevent the person, globally, from making the kinds of moral choices that are open to assessment. And if, as Wallace argues, fairness requires us to only hold people accountable for the fruits of their moral choices – in other words, if it requires that they be able to have avoided blame – then it will not be fair to blame people who cannot make the appropriate kinds of choices. Wallace takes this approach, claiming that what unifies this class of exceptions is that they somehow corrupt or undermine the powers of reflective self-control. With some of these conditions, it may be that these powers simply aren't

⁷⁴ Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme."

⁷⁵ Wallace, chapter 6.

fully present at all.⁷⁶ Different exempting conditions will act on different elements of these powers. Some conditions will prevent good deliberation at all, preempting any possibility of good moral deliberation. Some will perhaps impair an agent's ability to consider moral reasons. And though it is a question whether this is entirely possible, some excusing conditions might somehow interfere with the success of willing. Regardless, there is something in each successful exception case that shows that an important element of these powers is missing in the person being assessed, and that status makes her an inappropriate target for blame.

The study of excuses (and exemptions, and justifications) is extraordinarily complex, but Wallace's picture seems to capture what is appealing about many of our intuitions about excusing conditions. With the notions of excuse and justification on the table, we can now return our attention to mitigation. In particular, it is an open question whether these theories of excuse and justification shed light on the appropriate way to understand mitigation.

II.

As I've discussed earlier, the distinctive feature of mitigation is that it involves a reduction, not elimination, of the amount of blame an agent deserves for a given action. This notion has received brief mention in many discussions of both moral responsibility and excuse, and both Austin and Wallace mention it in passing. I will consider each of their remarks in turn, considering the implications of their positions, before returning to my own picture. Austin begins by assuming that the picture of mitigation will look exactly like the picture of excuse (and justification), differing only in degree. In other words, he claims that whatever gives excusing circumstances their excusing power will also be what gives mitigating circumstances their

⁷⁶ In the case of children, for example, we might think that these powers are importantly immature and underdeveloped. And thus the circumstance of childhood doesn't block or undermine the powers, but it does indicate that the relevant powers cannot yet fully function.

mitigating power. They will be identical in type, different in amount. He writes that “‘mitigation’ hovers uneasily between partial justification and partial excuse. ... Few excuses get us out of it completely: the average excuse, in a poor situation, gets us only out of the fire into the frying pan – but still, of course, any frying pan in a fire.”⁷⁷ On this view, mitigating circumstances would be excusing circumstances that don’t get us all the way out of the heat. And so, to understand mitigation, one only needs to understand excuse and its degrees of success. This is a popular take on mitigation, and most likely why it has received so little consideration in its own right.

Wallace often says things that suggest a similar approach. During his discussion of powers of reflective self-control, he acknowledges that these powers seem to come in degrees.

He writes:

These powers are matters of general psychological competence or capacity, and it is in the nature of such general capacities that they admit of various levels of development. Just as one may distinguish between rudimentary and more developed abilities to speak a foreign language, play squash, or solve differential equations, so one might in principle distinguish different levels of development of the powers to grasp moral reasons and to control one’s behavior accordingly.⁷⁸

In other words, he suggests that we might understand impairments in the relevant powers as limiting an agent to greater or lesser degrees, and this gives a natural picture as to how the process of mitigation might work. If fairness requires that we only hold appropriate targets blameworthy, and possessing these powers makes one an appropriate target, then it makes sense to think that possessing more or less of these powers might make one more or less appropriate. Thus, talking about degrees of cognitive and affective powers would give us the tools to consider

⁷⁷ Austin, p. 3.

⁷⁸ Wallace, p. 159.

degrees of blame and responsibility, and his picture of mitigation would simply track the degree of powers remaining. I call this a graduated-degree theory of agency.

This approach is problematic for two reasons. First, it seems that this method of establishing mitigation does not track with many of our common intuitions. If we thought that blame attaches appropriately to an agent depending on the degree to which she possesses powers of reflective self-control, then it would seem that it would be appropriate to assess people with higher than typical degrees of power as deserving larger amounts of blame than average agents, all other things constant. In other words, if a graduated-degree theory of agency is correct, it grounds the reduction in blame for reduced-skill agents in their reduced skill, all other features of action remaining constant. If that's the case, it seems there is nothing preventing an agent of extra-high skill to require an assessment of extra-high levels of blame, all other features of action remaining constant. Picture an agent particularly skilled in self-reflection and self-control. Assume that she is very knowledgeable about moral principles, has a high degree of accuracy when assessing moral circumstances, and is exceptional in her deliberative process. She also seamlessly moves from the fruits of her deliberation to her will – she is perfect in forming her intentions. Contrast this woman with a man who is normal in his reflective powers of self-control. He is good at what he does, deliberates well, and forms intentions effectively. Assume that both perform a morally wrong action, violating an obligation to a particular agent. It is not clear at all to me that we would judge the woman deserving of more blame (or say she is more responsible) just because she has a higher degree of the relevant powers.

It seems strange to think that some sort of super-obligation is created towards others just because one has a higher degree of skill. This is also explicitly problematic for theories of obligation that depend on obligations stemming from (or acknowledging) relationships between

morally equal peers. A view that requires blame to be dependent on the skill of the agent could not ground obligations in the status of the object of the obligation, since the nature of the obligation would change depending on the agent's skills. We'd be forced to say something like "Sue has an obligation to be kind to Jill. But since Ann has more skill, she has an obligation to be even kinder to Jill than Sue." If blame attaches to the violation of an obligation, and a more skilled agent deserves more blame for a given act than some average agent, the only option we are left with is that she has a different – and more robust – obligation. And this is problematic if we want to source the grounding of obligation in the status of the object, not certain facts about the agent. Even if we want to claim that obligations stem from the relationship between two agents, it seems incorrect to say that given the exact relationship between a person and a normal agent and a person and a super-skilled agent, that there would be a difference between the two resulting obligations. There might be cases where an agent's extra-high level of cognitive and affective might alter the content of a particular relationship, but in those cases, it will be the altered relationship that grounds differences in obligations, not the extra-high level of skill of the agent. Assessing an agent as deserving more blame for an act because she has extra-high skill is making a foundational mistake in the kinds of features relevant to moral assessment.

But there is another reason to be concerned with this approach; it is too narrow. Understanding mitigation as reductions in cognitive and affective powers doesn't explain many intuitively mitigating circumstances. Wallace himself mentions such cases in his discussion of excuses. He writes, "The lack of a persisting malicious motive or character trait might mitigate the degree of blame that would be appropriate,"⁷⁹ and situations such as these, or cases of coercion, or being a victim of a moral wrong, and so on, cannot be explained through this theory. There is nothing about a coercion case, for example, to suggest that there is anything abnormal

⁷⁹ Wallace, p. 123.

about the *agent* in question, or her powers. This is instead a set of circumstances that concern the context of the agent's acts, not her powers of agency. These examples of mitigation suggest that at best, an appeal to reduced-capacities can't ground many intuitive cases of mitigation. It is clear that the story cannot stop here if we are to find a plausible picture.⁸⁰

Another way of expressing these concerns is to point out that theories about excuse and justification are fundamentally theories about the necessary and sufficient conditions for agency. This is importantly different than a theory about the necessary and sufficient conditions for blame. It will no doubt be true that agency will be a necessary condition for the appropriate assignment of blame, but a full theory of the conditions for agency will still not be sufficient to capture the conditions for blame. Wallace and Austin were not wrong to suggest that circumstances that impair agency will also rightly impact blame, but that does not mean that a simple graduated-degree theory of agency will capture mitigation, especially since there are many intuitive cases of mitigation that don't involve any impairments in agency. And so I do not argue that these theories are wrong so much as I argue that these theories mislocate the problem at hand. In order to effectively bridge the gap between these conditions for agency and conditions for blame, we will need a greater story about where and when blame is appropriate.

This is precisely the role that my theory of mitigation plays. By showing how moral obligation requires success at multiple points in the deliberative and intention-formation process, I show how a particular action can contain moral success and failure at different morally relevant points. I argue that blame is deserved not merely at the top level of action, to be adjusted higher or lower depending on facts about the agent, but instead at each level of regulation required by a moral obligation, with the amount of blame that an agent deserves for a particular action being additive. And so minor impairments in cognitive or reflective powers, like adolescence or

⁸⁰ It should be noted that Wallace himself is not satisfied with this picture of mitigation either.

intoxication, will be mitigating, not because they represent a graduated-degree theory of agency, but because they prevent a genuine moral success or failure from obtaining, leaving us without the appropriate sort of thing to morally assess. And because my picture depends on necessary and sufficient conditions for the appropriate assessment of blame – not the appropriate assessment of agency – it will capture mitigating circumstances that have to do with deliberative context without in any impairment in agency. We can understand coercion as a case of mitigation because it represents a certain kind of moral failure, one which contains the appropriate regulation of deliberation throughout some of the process, and inappropriate regulation in another elsewhere. Having a theory about the appropriate moral assessment of the inner-workings of action, deep deliberation, allows us to make more nuanced assessments of action, and this nuance is the way to properly understand the notion of mitigation. Before concluding, however, there is one other popular view of mitigation that is worth considering. And so I will now say a few words about moral effort.

III.

There is one other view that is often raised by philosophers looking to capture the force of mitigation, and both Wallace and Austin give us clues that they have something close to this picture in mind. On this picture, mitigation obtains when not violating a moral obligation is very difficult, and the degree of mitigation presumably increases with the degree of difficulty. When discussing why extreme stress might be mitigating, Wallace argues that the mitigation might depend on the easiness (or difficulty) of not violating a moral obligation. Presumably, when it is easier to respect moral obligations, one gets less “slack” if she violates one. And if it is

extraordinarily difficult to respect moral obligations, perhaps that provides the grounds for a reduction in blame.

There is much that is appealing about this story, and it captures an interesting feature of many cases of mitigation. The woman I describe in Chapter 1, the Betrayed Lover, certainly would have found it more difficult to respect her moral obligations than the woman who cold-heartedly killed her lover out of boredom and financial interest. Being in extreme emotional states makes it harder to be susceptible to moral reasons, and it often sways people away from their better (read: more moral) judgments. Such a degree-of-difficulty picture also effectively captures a feature of coercion cases that feels morally relevant. When someone is forced into a pick-x-or-y situation, where x is a moral wrong and y is harm to the agent, choosing to respect the moral obligation is now attached to a harm that is difficult for the agent to swallow. What might have been an easy choice – respect one’s moral obligations – now has quite a cost attached. And that difficulty seems to be at least part of the reason that we feel sympathy for coerced wrong-doers. In fact, I believe that this degree-of-difficulty picture is lurking underneath the great majority of theories of mitigation, including perhaps the graduated-degree theory of agency I considered above. After all, it seems reasonable to think that fairness might require us to cut some slack (so to speak) to agents who must overcome great difficulty to respect their moral obligations. This picture seems to imply that fairness demands we respect moral effort. And while this seems intuitive on its face, I will argue that once we dig under the surface, this notion cannot be what is grounding our moral intuitions.

The first problem with this view is not dissimilar to the first problem with the graduated-degree theory of agency. Specifically, it gives us very strange results when we apply the view. If degree-of-difficulty carries with it some justification for reduction in blame, then we are, in a

sense, rewarding people who have a more difficult time respecting their moral obligations and punishing those who find it easier. This seems, on its face, ridiculous. If someone is faced with a decision that would agonize a normal person – commit a moral wrong or be killed, for example – and a morally gifted person makes the decision with ease because she’s just that good, it seems strange to think that she should receive less moral “credit” than someone who made the same decision with much greater difficulty. Yes, the person who had a harder time with the decision had to struggle with it; she perhaps exerted more effort. But that is because she is less good as a person. The same results seem to manifest even more dramatically when we look at cases of people who are exceptionally morally corrupt. Imagine a person who is quite a rotten human being, full of malicious inclinations. Should we give her more moral credit when she respects her moral obligations, because she had to extend so much more effort to do so? Does a malicious person deserve praise for being only a little malicious? It is possible that these strange results are particular to cases of praise, and perhaps they will not appear with the mitigation of blame. But if the degree-of-difficulty picture is mitigating because it is somehow connected to the fairness of the assessment of praise for the agent, it’s not clear why the same demands of fairness would not be relevant in the assessment of blame for the agent.

Regardless, the degree-of-difficulty approach will not actually capture the relevant features in many cases of mitigation, and a final case will help us see why. Imagine two people walking along the edge of a lake, coming from separate directions. Each spots a man in the deepest part of the lake, flailing and sputtering, obviously in distress and begging for help. Both of our bystanders are excellent swimmers, and the water is warm. They would both be able to swim to the drowning man and save him with relative ease, and both are normal agents, fully

developed in their reflective powers of self-control. Both recognize that they have a moral obligation to save the drowning man.

Our first bystander approaches the edge of the lake and considers diving in, but he is paused by his fear of water. He's had an irrational fear of water since his childhood. The man is not phobic; he is not overwhelmed by his fear, and he has overcome it many times before. In fact, he has developed very strong swimming skills to compensate for his anxiety. Nevertheless, getting into the water would be deeply unpleasant for him, making him feel large amounts of fear for his own safety, even though he knows that his fear is irrational. He deliberates, considering both the emotional distress it will cause him to dive in and his obligation to come to the drowning man's aid on his deliberative field, and he finally decides not to save the man. He has committed a moral wrong.

The second bystander, on the other side of the lake, has no such fear. This bystander is a famous director, on his way to the premiere of his new movie. He's been looking forward to this moment for years, the celebration and fanfare. This movie represents his first taste of enormous financial and critical success, and he wants to bask in the media attention, not to mention having his photograph appear all over the entertainment media. He knows that such attention will make his supporters proud, and his critics jealous. He spots the drowning man and realizes that if he jumps in to save him, he will no doubt miss his premiere, and thus his opportunity to be photographed. He deliberates, considering both the strength of his desire to have this moment in the limelight and his obligation to come to the man's aid on his deliberative field, and he finally decides not to save the man. He has also committed a moral wrong.

When we consider these two cases, I believe there is a clear intuition that the Fearful Man deserves less blame than the Director. So what explains this difference? It can't be degree of

difficulty. The Fearful Man had a very difficult decision to make, weighing his significant fear and anxiety against respecting his obligation to aid the drowning man. The fear was visceral and intense, and ultimately, that fear is what motivated his inaction. But the Director had an enormously difficult decision to make as well. This premiere would be the celebration of his fame and accomplishment. It would also be an opportunity to make his enemies bitterly jealous, something he desperately desired. To him, missing this event also felt emotionally painful, and this is what motivated his inaction. Both men faced a very difficult choice, and I will stipulate that the degree of difficulty was the same in each case. And yet, the degree of mitigation seems very different. If intuitions about these cases are correct, then it can't be a degree-of-difficulty theory that is doing the work in all mitigation cases.

My picture of mitigation can, however, explain the morally relevant differences between these two cases. There are certain features that remain constant between the cases of these two men. Both experienced the same degree of difficulty in their deliberative processes, both considered the fact that they had a moral obligation to aid the drowning man, and both undervalued the force of that obligation. The reason that each of them undervalued their obligation to aid is because they overvalued something else: anxiety and emotional distress in the case of the Fearful Man, and social recognition in the case of the Director. Thus, we might say that that part of their deliberation was a moral failure in exactly the same way. The difference lies in the content of what they overvalued. That is to say, the Fearful Man overvalued the emotional distress that he would encounter if he rescued the man, and this kind of fear is an appropriate consideration to weigh in the moral decision-making process.⁸¹ In fact, there are ways to adjust the case such that this consideration would become an adequate justification for

⁸¹ In the way I'm describing the case, this fear in question isn't a passion-inducing condition, but it is all too easy to reimagine the case where the fear is distracting him accordingly, and that would be a different and independent qualifying condition for mitigation.

the action of letting the man drown – certainly if he knew that jumping in would cause him to have an emotional breakdown and become permanently catatonic, it would no longer be wrong to not save the drowning man. The Fearful Man has a moral obligation to his own well-being as well as an obligation to aid the stranger, and his circumstances demand he navigate this moral overlap. And so the threat of severe emotional distress is a morally relevant consideration that he simply mis-weighed. We can stipulate that some emotional distress is not enough of a harm to act as a justification for his inaction – it doesn't adequately compare to the harm that the drowning man suffers when the fearful man fails to save him. But it is a consideration of the right type.

This is not the case with the Director. Missing a premiere, or making one's enemies jealous, are simply not the right types of things that ought to be on the moral deliberative field when considering whether to save a drowning man. It's true that the director overvalued this inclination of his and undervalued the obligation, but the content of the inclination matters. If someone stops before saving a drowning man and thinks "Hm, I could save this man... but I could also really go for some ice cream," something has gone very wrong. That inclination has no place in the deliberative process, and it does not belong on the deliberative field. And so the problem with the Director's violation is not merely that he undervalued his obligation. It's that he undervalued his obligation by placing it underneath something he should not have been considering in the first place. The Fearful Man has succeeded at acknowledging what the relevant moral reasons are and placing the life of the drowning man among other legitimate considerations. He constrains his considerations appropriately but then weighs those relevant considerations inappropriately. The Director's deliberative process fails in two places. That he even chooses to deliberate about whether he will attend the premiere or save the drowning man is

an enormous insult to the man. He does not acknowledge or respect that there's something special about his obligation to the man that puts a limit on the kind of considerations it is appropriate to weigh it against. Not constraining his deliberation as appropriate reveals a more foundational disrespect.

Moreover, the picture I offer captures what I think is most attractive about degree-of-difficulty views. Views that connect mitigation to difficulty, or moral effort, are views that express the notion that the context of an action matters to the moral assessment of the agent's performance. They play on the – true! – belief that the raw description of a motive-intention pair is not enough to shed light on that action's formative story, deep deliberation, or its relationship to the agent. The underlying problem with degree-of-difficulty views, other than the fact that they can't explain the force of some intuitively mitigating circumstances, is that they don't give us an adequate story about why we should give agents more "slack" when performing a morally required action is more difficult. The main virtue of my view is that it helps ground the intuitions that support theories like these, without the problematic commitments. It captures the reasons we think difficult circumstances are often mitigating circumstances, without claiming that the mitigation stems from the difficulty.

CONCLUSION

I believe that the picture I have offered goes a long way towards filling an explanatory gap between excuse and justifications on one side, and agents who deserve blame in full on the other. The picture I offered functions by understanding the moral assessment of blame as connecting to particular failures and successes at the level of deliberative regulation required by moral obligation, treating the intention formation process as a process that can succeed or fail at more than one place, in more than one way. One consequence of a view like mine is that cases of mitigation will turn out to be far more common than we might otherwise have thought. Many, many wrong acts are performed not out of malice, but out of fear, or laziness, or distraction, and opening up the deliberative particulars of those actions might well mean that far few people are fully to blame for their wrong acts than we perhaps like to believe. Mitigation might not be best viewed as a rare and exotic exception to normal maximally blame-triggering performances, but instead as an indispensable way to capture to nuance of everyday moral assessment accurately.

That said, the view does have some limitations. My picture is predicated on a desert-based theory of blame, and if one does not subscribe to such a notion, then the discussion I have offered about the points of deliberation relevant to desert of blame will not be helpful. If someone were to view blame as instead depending only on the public perceptions of the act, then these deliberative backstories of an action simply won't be able to do as much work. There are also theories of blame that require necessary and sufficient conditions for the provisioning of blame independently of desert. These views might include considerations such as mercy or remorse as mitigating, even though they do not concern desert. My view does not speak to such conditions, nor does it offer the tools to ground those intuitions. My view also depends on a

particular picture of the demands and content of moral obligation, and the regulative demands that stem from them. This is woefully under-argued for this this piece.

It is also not clear whether such a picture will work in parallel with other moral responses such as praise. I hinted at the notion of praise above, but there is much work to be done to establish whether the necessary and sufficient conditions for the deserving of praise will be parallel to the conditions for the deserving of blame. I suspect that the picture of praise will not turn out to be a mirror image of the picture of blame, and mitigation of praise thus requires a theory of appropriate assessment all its own.

But it is my hope that the approach of looking deeper at the deliberative part of action during the moral assessment process helps us gets closer to accurate assessments. The method of assessing the constructive parts of intention-formation in isolation, looking at the successes and failures during the process, helps us understand action not as a single brute decision on the part of the agent, but instead as a series of subtle and morally relevant sub-decisions, all of which are worth evaluating.

BIBLIOGRAPHY

- Albritton, Rogers. "Freedom of Will and Freedom of Action," *Proceedings of the American Philosophical Association* 59 (1985-1986).
- Arenella, P. "The Diminished Capacity and Diminished Responsibility Defenses: Two Children of a Doomed Marriage." *Columbia Law Review*, 77(6), 827-865 (1977).
- Arizona Capital Sentence Guide A.R.S. §13-751(G)(1). 2012. Web.
<<http://www.azcourts.gov/Default.aspx?alias=www.azcourts.gov/ccsguide> >
- Atkins v. Virginia, 536 U.S. 304. 2002. Web.
<<http://supreme.justia.com/cases/federal/us/536/304/case.html>>
- Austin, J.L., "A Plea for Excuses." *Proceedings of the Aristotelian Society*, vol. 57, (1956-1957): pp 1-30.
- Baker, B. M. "Theorizing about Responsibility and Criminal Liability." *Law and Philosophy*, 11(4), (1992): 403-430.
- Bilefsky, Dan. "Murder Trial Hinges on Questions of Domestic Abuse," *The New York Times*, September 18, 2011, p. A23 of the New York edition, (2011).
- Brandt, R. B. "The Insanity Defense and the Theory of Motivation." *Law and Philosophy*, 7(2), (1988): 123-146.
- Buss, S. "Justified Wrongdoing." *Nous*, 31(3), (1997): 337-369.
- Carr C. L. "Coercion and Freedom." *American Philosophical Quarterly*, 25, (1988): 59-67.
- Carr, C. L. "Duress and Criminal Responsibility." *Law and Philosophy*, 10(2), . (1991): 161-188.
- Cauffman, E. & Steinberg, L. "(Im)maturity of Judgment in Adolescence: Why Adolescents May Be Less Culpable Than Adults." *Behavioral Sciences and the Law*, 18, (2000): 741-760.
- Commonwealth v. DeMarco, 809 A.2d 256. 2002. Web.
<<http://caselaw.findlaw.com/pa-supreme-court/1377368.html>>
- Darwall, Stephen. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press. 2009.
- Feinberg, J. *Doing and Deserving*. Princeton: Princeton University Press, 1979.
- Feld, B. C. "Abolish the Juvenile Court: Youthfulness, Criminal Responsibility, and Sentencing Policy." *The Journal of Criminal Law and Criminology*, 88(1), (1997): 68-136.

- Feld, B. C. "Juvenile and Criminal Justice Systems' Responses to Youth Violence." *Crime and Justice*, 24, (1998): 189-261.
- Fingarette, H. "Disabilities of Mind and Criminal Responsibility – A Unitary Doctrine." *Columbia Law Review*, 76(2), (1976): 236-266.
- Fischer, J. M. "Recent Work on Moral Responsibility." *Ethics*, 110(1), (1999): 93-139.
- Fischhoff, B. "Risk Taking: A Developmental Perspective." In J.F. Yates (Ed.), *Risk-Taking Behavior*, (1992): 133-162. Chichester: Wiley.
- Florida Statutes Section, No. 921.0026, 2012. Web.
<http://www.flsenate.gov/Laws/Statutes/2010/921.0026>
- Frankfurt, H. "Coercion and Moral Responsibility," *Essays on Freedom of Action*, Ted Honderich (ed.), London: Routledge & Kegan Paul, (1973): 65–86.
- Frankfurt, H. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy*, 68(1), (1971): 5-20.
- Frankfurt, H. "Identification and Wholeheartedness." In F. Schoeman (Ed.), *Responsibility, Character, and the Emotions*. Cambridge: Cambridge University Press, (1987): 27-45.
- Frankfurt, H. *The Importance of What We Care About*, Cambridge: Cambridge University Press. 1988.
- Freeh, Sporkin & Sullivan, LLP (2012-07-12). "Report of the Special Investigative Counsel Regarding the Actions of The Pennsylvania State University Related to the Child Sexual Abuse Committed by Gerald A. Sandusky." 2012. Web.
 <http://i2.cdn.turner.com/cnn/2012/images/07/12/freeh.report.pdf?hpt=hp_c1>
- Glover, Jonathon. *Responsibility*, New York, Humanities Press, 1970.
- Goodin, R. E. "Apportioning Responsibilities." *Law and Philosophy*, 6(2), (1987): 167-185.
- Grimm, Jacob and Wilhelm. *Household Tales* (English translation by Margaret Hunt), "Rapunzel", (1884).
- Hieronymi, Pamela. "The Force and Fairness of Blame." *Philosophical Perspectives*, 18, (2004): 115-148.
- Hieronymi, Pamela. "The Will as Reason," *Philosophical Perspectives* 23, no.1 (2009): 201–220
- Herman, Barbara. *Moral Literacy*, Harvard University Press, 2008.
- Herman, Barbara. *The Practice of Moral Judgment*, Harvard University Press, 1996.

- Hookway, C. & Wilkes, K.V. "Conscious Belief and Deliberation." *Proceedings of the Aristotelian Society*, Supplementary Volumes 55, (1981): 75-89 and 91-107.
- Husak, D. N. "Partial Defenses." *Canadian Journal of Law and Jurisprudence*, 11(1), (1998): 167-192.
- Kahan, D.M., & Nussbaum, M.C. "Two Conceptions of Emotion in Criminal Law." *Columbia Law Review*, 96(2), (1996): 269-374.
- Kant, Immanuel. *Groundwork of the Metaphysics of Morals*, Mary Gregor (trans.), New York: Cambridge University Press, (1997 [1785]).
- Kenny, A. "Duress per Minas as a Defense to Crime: II," *Law and Philosophy*, 1(2), (1982): 197-205.
- Korsgaard, C.M. "Self-Constitution in the Ethics of Plato and Kant." *The Journal of Ethics*, 3(1), (1999): 1-29.
- Lavie, N. "Distracted and Confused?: Selective Attention Under Load." *Trends in Cognitive Sciences*, 9(2), (2005): 75-82.
- Locke, L. A. "Personhood and Moral Responsibility." *Law and Philosophy*, 9(1), (1990): 39-66.
- Model Penal Code and Commentaries (official draft and revised commentaries): Part I, General Provisions* (reporter, with others). 3 vols. (American Law Institute, 1985).
- Morse, S.J. "Hooked on Hype: Addiction and Responsibility." *Law and Philosophy*, 19(1), (2000): 3-49.
- Moselhy, Hamdy F., et al. "Frontal Lobe Changes in Alcoholism: A Review of the Literature," *Alcohol and Alcoholism*, Vol. 36, No. 5, (2001): 357-368.
- Nozick, Robert. "Coercion," in *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, Sidney Morgenbesser, Patrick Suppes, and Morton White (eds.), New York: St. Martin's Press, (1969): 440-472.
- Nozick, Robert. *Philosophical Explanations*. Cambridge, MA: Harvard University Press, 1981.
- O'Neill, Onora. "Which are the Offers You, Can't Refuse?" Chapter 7 in *Violence, Terrorism, and Justice*, R. G. Frey and Christopher Morris (eds.), Cambridge: Cambridge University Press, (1991): 170-195.
- Paus, T. "Mapping Brain Maturation and Cognitive Development During Adolescence." *Trends in Cognitive Sciences*, 9(2), (2005): 60-68.
- Penry v. Lynaugh, 492 U.S. 302. Web. 1989.
<<http://supreme.justia.com/cases/federal/us/492/302/case.html>>

- Pettit, P. "Responsibility Incorporated." *Ethics*, (2007): 117, 171-201.
- Richards, N. "Criminal Children." *Law and Philosophy*, 16(1), (1997): 63-89.
- Ripstein, Arthur. "Authority and Coercion," *Philosophy and Public Affairs*, (2004): 2–35.
- Sagen, Carl. *Cosmos*. New York: Random House, 1980.
- Scanlon, T.M. *Moral Dimensions*, Harvard University Press, 2008.
- Scanlon, T.M. "The Significance of Choice" in Sterling M. McMurrin, ed., *The Tanner Lectures on Human Values*, vol 8. Salt Lake City: University of Utah Press, 1988.
- Scanlon, T.M. *What We Owe To Each Other*, Harvard University Press, 1988.
- Schleifer, M. "Instrumental Conditioning and the Concept of the Voluntary." *Ethics*, 82(2), (1972): 163-170.
- Schrag, F. "The Child in the Moral Order." *Philosophy*, 52(200), (1977): 167-177.
- Shultz, T. R., Wright, K., & Schleifer, M. "Assignment of Moral Responsibility and Punishment." *Child Development*, 57(1), (1986): 177-184.
- Sinnott-Armstrong, Walter. *Moral Dilemmas*, Oxford: Basil Blackwell, 1988.
- Smith, Angela M. "On Being Responsible and Holding Responsible." *The Journal of Ethics*, 11, (2007): 465-484.
- Smith, Angela M. "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115 (2): (2005): 236-271.
- Smith, Holly. "Culpable Ignorance." *The Philosophical Review* , Vol. 92, No. 4, (Oct. 1983).
- Snyder, O.C. "Criminal Responsibility." *Duke Law Journal*, 1962(2), (1962): 204-218.
- Sparks, R.F. "'Diminished Responsibility' in Theory and Practice." *The Modern Law Review*, 27(1), (1964): 9-34.
- Steinberg, L. "Cognitive and Affective Development in Adolescence." *Trends in Cognitive Sciences*, 9(2), (2005): 69-74.
- Steinberg, L. & Scott, E. S. "Less Guilty by Reason of Adolescence: Developmental Immaturity, Diminished Responsibility, and the Juvenile Death Penalty." *American Psychologist*, 58(12), (2003): 1009-1018.
- State v. Dyke, 361 N.J. Super. 403, 825 A.2d 1163. 2003. Web.
<https://njcourts.judiciary.state.nj.us/web0/mcs/case_law/a2658-01.pdf>

- Strawson, P.F. "Freedom and Resentment," in Gary Watson, ed., *Free Will*. Oxford: Oxford University Press, 1982.
- Thomson, Judith Jarvis. "Self-Defense" *Philosophy and Public Affairs*, Vol. 20, No 4 (Autumn, 1991).
- U.S. v Atencio, 586 F.2d 744. 1978. Web.
<<http://law.justia.com/cases/federal/appellate-courts/F2/942/795/282172/>>
- U.S. v Jennell, 749 F.2d 1302. 1985. Web.
<<http://openjurist.org/749/f2d/1302/united-states-v-jennell>>
- Viera, Mark and Jo Becker (November 14, 2011). "Ex-Coach Denies Charges Amid New Accusations". The New York Times. Report: Former coach Jerry Sandusky used charity to molest kids (Harrisburg Patriot-News, November 6, 2011).
- Virginia Code. Section 19.2-264.4.B. 2011. Web.
<<http://law.justia.com/codes/virginia/2011/title19-2/chapter15/19-2-264-4/>>
- Vuoso, G. "Background, Responsibility, and Excuse." *The Yale Law Journal*, 96(7), (1987): 1661-1686.
- Watson, G. *Agency and Answerability*. Oxford: Clarendon Press, 2004.
- Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In Ferdinand Schoeman, ed., *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, Cambridge: Cambridge University Press, (1987): 256-286.
- Wallace, R.J. "Addiction as Defect of the Will: Some Philosophical Reflections." *Law and Philosophy*, 18(6), (1999): 621-654.
- Wallace, R. Jay. *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press, 1994.
- Watson, Gary. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," in Ferdinand Schoeman, ed., *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge, Mass: Cambridge University Press, 1987.
- Yaffe, G. "Recent Work on Addiction and Responsible Agency." *Philosophy and Public Affairs*, 30(2), (2001): 178-221.
- Zimmerman, Michael J. "Another Plea for Excuses." *American Philosophical Quarterly*, Vol 41, Num 3, 2004.