

City University of New York (CUNY)

CUNY Academic Works

All Dissertations, Theses, and Capstone
Projects

Dissertations, Theses, and Capstone Projects

5-2019

Who Needs Blame?: Answerability Without Expressed Blame

Sarah Gokhale

Graduate Center, City University of New York

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/gc_etds/3087

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

**WHO NEEDS BLAME?:
ANSWERABILITY WITHOUT EXPRESSED BLAME**

by

SARAH GOKHALE

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2019

© 2019
SARAH GOKHALE
All Rights Reserved

**WHO NEEDS BLAME?:
ANSWERABILITY WITHOUT EXPRESSED BLAME**

by

Sarah Gokhale

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Hagop Sarkissian, PhD
Chair of Examining Committee

Date

Nickolas Pappas, PhD
Executive Officer

Supervisory Committee:

Miranda Fricker, PhD
Jeffrey Blustein, PhD
Hagop Sarkissian, PhD

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Advisor: Miranda Fricker

This dissertation argues that we can hold other agents morally responsible without expressing blame and, more strongly, that doing so is preferable. I first argue that blame is fundamentally retributive, and that blame's retributive foundation is incipiently present even in civilized guises. As such, even though some forms of expressed blame are quite civilized, expressed blame always involves a risk of emotional damage, entrenchment, and escalation. To make things worse, I argue that anger is an exacerbating feature of blame's retributive foundation. I then argue that, generally speaking, cases of public blame involve higher stakes than cases of private judgments of blame. This difference in stakes informs the warrant we have to make private judgments of blame, as compared to the heightened warrant we need to make public expressions of blame. Throughout the dissertation, I make repeated use of the idea that even if expressing blame to a wrongdoer is fitting for a given case of wrongdoing, it might not be appropriate to express that blame, due to the practical, epistemic, and ethical risks that expressions of blame invoke. Finally, I present three alternate modes of response to moral wrongdoing that allow us to hold wrongdoers responsible (construed as answerability) without expressed blame. I further argue for their context-dependent superiority to expressed blame. My hope with this project, more generally, is to expand the range of useful responses to moral wrongdoing one can take. Put simply, I aim to shift our default impersonal moral strategies of response to wrongdoing away from expressed blame and towards alternative practices.

Acknowledgements

Thank you to my advisor, Miranda Fricker, for taking such a genuine interest in my project from day one and for helping me bring each of my nascent ideas to fruition. While I had been planning to work in aesthetics, taking your course on Blame & Forgiveness in 2016 sparked my lasting interest in this topic. I owe so much of this project to not only your wealth of knowledge, but to your patience, kindness, and thoughtfulness in our work together. Thank you for helping me become a better writer and clearer thinker. Thank you for your mentorship, both personally and professionally, over these past three years. I am profoundly grateful.

Thank you to my committee members, Jeffrey Blustein and Hagop Sarkissian, for pushing me to explain myself more precisely, and for bringing so much depth and expertise to my research and writing.

Thank you to Elise Springer, for introducing me to virtue ethics and for first helping me conceive of the possibility and joy of being a strong female philosopher. Thank you to Linda Alcoff and Carol Gould, for helping me normalize that possibility.

Thank you to my mom, step-dad, and grandmother for instilling in me the importance of education and generosity from an early age. Thank you for leading the way and reminding me, time and time again, what matters most.

And thank you to my friend John, for believing in me.

For AER, for teaching me grace

Table of Contents

Introduction | What is Blame?

I. Statement of Thesis.....	1
II. What is Blame?	8
III. Blameworthiness Versus Blame & Appropriateness Conditions.....	12
IV. The Best Case for Blame.....	15
V. Conclusion.....	17

Chapter One | Blame and Retribution

I. Introduction.....	19
II. Blame is Retributive.....	19
III. Civilizing Blame.....	25
IV. Domesticated Blame.....	28
V. Objections to McGeer.....	31
A. Building on Top of a Retributive Core.....	31
B. Bad is Stronger than Good.....	36
VI. Fittingness Versus Appropriateness.....	39
VII. Vargas.....	42
VIII. Conclusion.....	46

Chapter Two | Blame, Retribution, and Anger

I. Introduction.....	48
II. Nussbaum.....	50
A. Anger and Retribution.....	50
B. Transition-Anger.....	54
C. Nonviolent Communication.....	57
D. Is Nussbaum's Case Overstated.....	59
E. Fittingness Versus Appropriateness.....	61
III. Wolf.....	62
A. Wolf's Account.....	62
B. Objections to Wolf.....	67
IV. Flanagan.....	69
A. Flanagan Against Anger.....	69
B. Eliminativist Views About Anger.....	71
C. Modified Eliminativism	73
V. Bommarito.....	76
A. Bommarito's Account.....	76
B. First Objection: Restraint.....	78
C. Second Objection: Anger Has Too Much Heat.....	80
VI. Conclusion.....	83

Chapter Three | A Further Argument against Expressed Blame: Epistemic Skepticism

I. Rosen's Epistemic Skepticism.....	86
A. Rosen's Account.....	86
B. Is Rosen's Account Overstated?	91

i. Against Wholesale Skepticism.....	91
ii. Stakes.....	92
II. Harman’s Reply to Rosen.....	98
III. Sher on the Searchlight View.....	104
IV. Revisiting the Doctor Case: A Moderated Skepticism.....	107
V. Confucian Support for Epistemic Skepticism.....	110
VI. Conclusion.....	114
Chapter Four Answerability Without Expressed Blame	
I. Introduction.....	117
II. Responsibility as Answerability.....	120
III. Responsibility Without (Affective) Blame.....	126
A. Answerability Without (Affective) Blame.....	131
B. Inverting the Paradigmatic Response: Foregrounding Compassion.....	133
i. Compassion.....	133
ii. Curiosity.....	135
iii. Answerability.....	137
IV. Transforming Hostile Emotions: The Emotional Conversion View.....	139
A. Emotionally Cooler Modes of Response.....	140
B. Empirical Research on Anger and Sadness.....	142
C. Answerability Through the Emotional Conversion View.....	145
V. Confucian Answerability.....	147
A. Sarkissian’s Commentary.....	148
B. Confucian Answerability.....	155
C. When Immediate Response is Required	157
VI. The Moral Failure to Not Express Blame.....	160
A. The Moral Duty to Express Blame.....	161
B. Response to the Moral Duty to Express Blame Objection.....	163
i. Manifestations of Respect.....	163
ii. Dignity.....	165
VII. Conclusion.....	166
Bibliography.....	168

Introduction. What is Blame?

I. Statement of Thesis

There has been a recent surge of interest in the philosophical study of blame. For example, some philosophers discuss the connection between blame and moral responsibility. (Gideon Rosen, Elizabeth Harman) Others discuss the connection between blame and free will. (Derek Pereboom, J.J.C. Smart) Others try to pin down just what blame *is*. (Angela Smith, T.M. Scanlon, George Sher) Others defend the connection between blame and retributive punishment. (Christopher Bennett, Antony Duff) Others discuss blame from a clinical therapeutic context. (Nicola Lacey, Hanna Pickard)

Whatever blame is, we surely do it a lot. Judging and then expressing blame to wrongdoers is a central feature of our daily lives. Expressing blame is a moral practice with wide ranging expressions. Sometimes expressed blame is angry or indignant. Consider expressions like *How could you do that!! I cannot believe you. What were you thinking?!?! Other times expressions of blame are somber or disappointed. Consider expressions like You let me down. This really hurts me. I'm so disappointed in you. We infer the emotional valence of expressions of blame through the blamer's tone, body language, volume, facial expression, etc. We express blame to a variety of agents—ourselves, our friends and family members, acquaintances, strangers, celebrities, social institutions, politicians, etc. We blame these agents for morally culpable (or what we *perceive* to be morally culpable) actions. When someone does something that we find morally objectionable, our natural response is to judge them blameworthy and then express blame towards them.*

This practice of expressing blame to wrongdoers does not always last for a long period of time. Sometimes we express blame towards someone and then forgive them mere moments later. For example, when my friend misremembers the time that we arranged to meet at a local coffee shop and mistakenly shows up a half an hour late, I might express blame towards her for a few seconds, and then we quickly move on. Her tardiness is not a big deal to me, so I do not hold a grudge or even really think twice about it. It is not typical behavior from her. It was an innocent mistake. I forgive her completely. Sometimes the process of expressing blame is brief and gets resolved quickly. Other times, the blame we express towards others is much more hostile, and lasts for months, if not years. We might angrily express blame towards someone and then continue to blame them and stew in those resentful and indignant feelings for years to come. The feeling might weaken in intensity, but not fully dissolve, even years later. Major cases of moral wrongdoing often involve a more drawn out process of expressing blame and do not get resolved quickly. Sometimes the wrongdoers in these cases are never forgiven. Either way, expressing and receiving blame with others in our shared moral communities is deeply embedded into the fabric of our interpersonal lives.

Indeed, blame is such a fundamental part of our daily lives that we often express it publicly without even realizing that we are doing so. When a close friend forgets to call us on our birthday, someone cuts in front of us in line at the grocery store, our partner betrays us, or a colleague takes all the credit for a project we jointly worked on, our first response is likely to express blame (to varying degrees) to each of these agents.

Blame has become our default response to being morally wronged. As Jeffrey Blustein explains, the standard philosophical view is “Why wouldn’t one want to blame the offender if one is convinced of his wrongdoing? After all, if one wants the offender to change, to stop doing

what he has done, isn't blame the most efficacious way of bringing this about?"¹ Rather than think through all of the possible reasons why the individual acted in a morally bad way, most of us jump immediately from feeling wronged by someone to expressing blame to them. There is rarely a pause for reflection about how certain we are about the wrongdoer's culpability. Worse still, many times this expression of blame is highly affective, meaning that the blame is expressed with some set of *negative* emotional attitudes, such as resentment, indignation, anger or a desire for revenge.

This heated, emotional expression of blame is considered justified by most of us because when others wrong us, we take their bad actions to mean that they deserve to feel badly in response. When others hurt us with their wrong actions, we not only want them to know that they hurt us, but, more strongly, we want them to *feel badly* about it. We can achieve this goal—wanting the wrongdoer to feel badly—through expressing blame to them. One of the major functions of expressing blame is to inspire remorse in wrongdoers. Blamers want wrongdoers to feel badly about their actions, in order to motivate them to change their behaviors. On this view, wrongdoers need to feel badly about their actions to be motivated to act better next time. By feeling remorse, wrongdoers gain the motivation to make morally better choices. If wrongdoers lack remorse, the traditional view argues, then wrongdoers might not be motivated to change their behavior. So, it is typically thought that expressing blame is not only warranted, but, more strongly, that it is a *required* response to wrongdoing because it serves the positive function of inspiring remorse in wrongdoers with the aim of helping them gain moral motivation.

Is this the picture of our daily moral lives we want? Do we want to be so critical of each other's actions, so constantly looking for and then expressing the faults in others? Without

¹ Blustein, Jeffrey (2018). "Should We Get Rid of Blame?" Unpublished manuscript, p. 2.

proper pause for reflection, we risk doling out hostile expressions of blame to innocent agents. Given blame's centrality in our daily lives, it is important to ask whether it is, in fact, morally and socially valuable to us. As D. Justin Coates and Neal A. Tognazzini ask, "If blame turns out to have little or no value, then shouldn't we try to excise it, and its associated pain and suffering, from our moral lives?"² Traditionally, expressing blame to moral wrongdoers has been deemed valuable on the grounds that its evaluative force, while not always pleasant to endure, is necessary for wrongdoers to grasp the significance of their wrongs and to be motivated to change. At the same time, all of us have been on the receiving end of blame and can viscerally understand how emotionally (and sometimes physically) unpleasant it can be. Michelle Mason aptly describes this philosophical tension:

Being viewed with reproach, at least by those we care about, is disconcerting and something we properly aim to avoid. Yet, blame has a more appealing face: Holding, and being held, accountable is, I take it, important to us and such accountability arguably requires that we remain within blame's purview. Blame thus challenges the philosopher to ground its special importance without denying its reproachful quality and associated justificatory burdens.³

My dissertation takes up this philosophical tension in blame—its simultaneously pleasant and unpleasant faces. The scope of my project is quite narrow. The type of blaming practices I am interested in are specifically *interpersonal* and *public*. I discuss public communications of blame between two people, or at most a small group in interpersonal moral settings. This dissertation does not address blame that is expressed in larger group contexts, such as blaming institutions, corporations, or world leaders. Further, I will not discuss the role of blame in larger structural contexts, such as blaming the stock market, Capitalism, or the Walmart Corporation. Finally, I

² Coates, Justin D. and Tognazzi, Neal A. (2012). "The Contours of Blame" in *Blame: Its Nature and Norms*, eds., D. Justin Coates and Neal A. Tognazzini. Oxford University Press, p. 23.

³ Mason, Michelle (2011). "Blame: Taking it Seriously". *Philosophy and Phenomenological Research* 83:2, p. 473.

will not discuss the role of non-human blame, such as Hurricane Sandy being to blame for a loss of power and significant, sustained structural damage in New York City.

I am interested in how expressed blame manifests in our daily lives interpersonally among and between those closest to us. Of course, I recognize that interpersonal expressions of blame cannot help but intersect, to some degree, with institutional or political forms of blame. Nevertheless, the two are distinct, and for the purposes of this dissertation, will remain distinct. In this dissertation, I argue that expressing blame in interpersonal contexts comes with a host of troubling social risks such as escalation and entrenchment. I further argue that we can hold others responsible in our interpersonal lives without expressed blame, and that we may be *best served* doing so. However, this does not carry over to a claim about whether blame is necessary in institutional or political contexts. The two realms—interpersonal and institutional—might need to be attended to entirely differently.

As another logistical note, this dissertation focuses exclusively on dyadic cases of blame. I have chosen to limit myself to this dyadic, interpersonal realm of blame so that I have a manageable topic to work with given the constraints of this dissertation. I leave open whether expressed blame should continue to be the default response in political or legal spheres. Though I anticipate that my worries about interpersonal blame certainly bleed, to some degree, into political and legal realms, I am not making an argument against expressed blame in settings outside of interpersonal cases. I have now mentioned the term *expressed blame* in this chapter several times. In what follows, I will refer to the type of blame we communicate interpersonally in response to moral wrongdoing as *expressed blame*. I define the term as follows.

Expressed Blame: Blame that is publicly communicated to a wrongdoer (or perceived wrongdoer) in response to a wrongdoing (or perceived wrongdoing) in interpersonal settings. Such communication is most often expressed affectively, in emotional registers that can include anger, indignation, resentment, and retributive impulses.

We express blame daily to our mothers, our partners, friends, coworkers, bus drivers, and so on. The blaming practices I will home in on are those that are *communicated*. This communication usually happens verbally, but sometimes silence itself is a way to express blame. Further, we communicate our blame through our body language, facial expressions, gestures, volume, and tone, too. Expressed blame is a social practice that mutates over time as our social conventions change.

Expressing blame in interpersonal contexts is a *practice* we each, to varying degrees, participate in. The practice of expressing blame is a wholly separate issue from the issue of determining (or judging) whether or not one is blameworthy. I am interested specifically in the social harm done by publicly expressing blame, and not in the cognitive judgment of determining an agent's culpability. The difference is obvious. Judgments of blameworthiness are not always expressed. I can have the private thought 'My brother oversleeping is the reason we are now late to the concert' and in doing so internally judge my brother as responsible for our lateness. But, that private judgment does not mean I will actually *express* blame to him verbally. Judging one blameworthy is to judge that one has acted wrongly. Oftentimes, when we make this cognitive judgment, we also go the additional step of expressing blame on the basis of deeming one blameworthy. However, going this extra step is not required. This extra step of actually expressing the blame publicly—and the social risks involved in doing so—is the sole focus of this project.

Thus blame, as I refer to it throughout this dissertation, is fundamentally a public communication between a blamer (or blamers) and a wrongdoer (or wrongdoers). My argument about expressed blame works in two stages. First, I will argue that expressed blame, even in its most civilized forms, involves an underlying retributive spirit. I will argue that this retributive

spirit is *morally bad*, in that it often leads to escalation such that expressions of blame can quickly spiral out of control. Drawing upon work by Victoria McGeer and Martha Nussbaum (among others), who each discuss the connection between retribution, anger, and blame, I claim that blame's retributive nature gives us good reason to be wary of expressing it interpersonally in response to moral wrongdoing. I will further claim that anger is an exacerbating feature of blame's retributive spirit, rendering angry forms of blame the worst kind.

I shall then strengthen this argument by offering an epistemic argument about reasons to be skeptical about judgements of blameworthiness. In doing so, I discuss Gideon Rosen's epistemic skepticism and use his view as support for my own. Rosen argues that we rarely have the proper warrant to make positive judgments of culpability. I will expand Rosen's discussion regarding warrants about judgments of culpability into the realm of public communications of blame. I argue that if we cannot be sure whether or not someone is culpable, it follows to exert caution before making private judgments about culpability about them, in addition to exerting extreme caution in publicly expressing blame to them. As I will argue, the stakes involved in public expressions of blame are much higher than the stakes involved in private judgments of blame, and our practices of expressing blame should reflect this difference.

Finally, chapter four provides practical, alternative methods of responding to moral wrongdoing that do not require or rely upon the use of expressed blame. Contra the standard view that expressed blame is required to serve the functions of 1. Motivating and 2. Holding others morally responsible, I will argue that we do not need expressed blame to achieve either function, and that we have other viable options. If I am right that it is possible to cultivate habits of moral response such that we can hold others morally responsible and motivate them to change without expressed blame, and granted that blame often has at a problematically hostile

expression, then there is a strong prima facie case for thinking that holding each other responsible without expressed blame is a preferable option. Though many philosophers raise skepticism about blame's positive value, few end up arguing, as radically as I will, that we can hold wrongdoers responsible without expressed blame, and that doing so is preferable.

I will use the remainder of this introductory chapter to briefly discuss the four predominant accounts about what blame is in the literature. I will be drawing upon them in the chapters to come.

II. What is Blame?

There are four established views about what blame is: cognitive, conative, affective, and functional. In this section, I will spend more time discussing affective and functional accounts of blame, because I argue that blame is a combination of affective and functional. I will not be developing my own definition of blame in this dissertation. Instead, I will make the case that expressed blame often plays a functional role and takes an affective expression.

First, cognitive accounts suggest that blame is a *judgment* or *evaluation* someone makes towards another person due to her attitudes or actions (See Hieronymi (2004), Zimmerman (1998), Scanlon (1986, 2008) and Watson (1996, 2004)). If blame is primarily a judgment, then it need not be expressed. Instead, the judgment could simply remain privately held in the blamer's mind. For example, Zimmerman writes that when we blame someone, there is a "discredit or debit in his ledger...(we have decided) that his 'moral standing' has been

diminished.”⁴ Zimmerman’s account suggests that blame merely involves the cognitive judgment that someone’s moral standing has been diminished.

Second, conative accounts of blame argue that *desires* and *intentions*, rather than judgments, are essential to blame (See Sher (2004)). According to George Sher, to blame is to have a belief-desire pair. The belief held by the blamer is that the wrongdoer acted wrongly. The desire held by the blamer (and maybe the wrongdoer, too) is that the wrongdoer *not* have acted badly in that way. The desire is a backward-looking desire about the wrongdoing, which can be quite similar to *regret* for the wrongdoer. The wrongdoer desires that she not have done that wrong action. Sher claims that we can understand blame based solely on this belief-desire pair.

Third, affective accounts take our emotions to be central to what blame is (see Wolf (2011), Bell (2013)). These blaming emotions are largely, if not entirely, negative. Common blaming emotions include: anger, indignation, resentment, or disappointment. In the interpersonal domain, expressed blame is highly affective. We can make sense of cooler forms of blame when blame is expressed to a stranger for some minor moral wrongdoing or to the leader of a country far removed from one’s own. These expressions of blame are likely cooler in intensity due to the distant nature of the relationship between the blamer and wrongdoer. It might be hard to get all that worked up over situations of wrongdoing that do not directly apply to our lives or to the lives of those close to us. Differently, with those close to us, our blaming emotions tend to arise more strongly and are more personal. For example, when my sister says something racist about our mutual friend, I feel much more betrayed than if a stranger makes the same racist statement. I expect *better* of my sister, and I have an expectation of her that is not being met. Though I might also express blame to the stranger who expresses the same racist sentiment, it is

⁴ Zimmerman, M. (1988). *An Essay on Moral Responsibility*. Totowa, NJ: Rowman & Littlefield, p. 38.

less personalized, because we have no prior relationship or expectation of one another that has now been broken.

P. F. Strawson's *Freedom and Resentment* (1962) is widely regarded as the catalyst for emotional theories of blame. Strawson argued that to blame someone (or oneself) is simply to have an emotional response, or set of what he termed *reactive attitudes*. According to Strawson, the reactive attitudes are "attitudes belonging to involvement or participation with others in interpersonal human relationships,"⁵ such as resentment, gratitude, forgiveness, and anger. This list of reactive attitudes is quite diverse. Gratitude is clearly a positively-valenced emotion, and forgiveness also seems positively-valenced, whereas resentment and anger are clearly negatively-valenced. Gratitude is unlikely to crop up in our expressions of blame, whereas the negative emotions on the list are ordinary blaming emotions.

Although anger, resentment, and indignation are not identical emotions, and it might seem too hasty to lump them all together, they each have a negative charge for the one who emotes them, and they are received negatively. It hurts to be on the receiving end of anger, resentment or indignation. The subtle differences between these terms in the particular context of expressing blame, while interesting and important, do not concern me. Which of these negative emotions we habitually gravitate towards in our blaming practices depends on the particulars of one's personality traits, lived experience, upbringing, cultural customs, etc. The crucial point is how negative and hostile most of the typical blaming emotions are.

Fourth, functional accounts of blame pick out blame according to its *functional role* (See Smith (2013), Fricker (2016)). For example, Miranda Fricker argues that the most explanatorily basic form of blame is communicative blame. She claims that the function of communicative

⁵ Strawson, P.F. (1962). "Freedom and Resentment". *Proceedings of the British Academy* (48), pp. 1-25.

blame is to inspire remorse (understood cognitively) in wrongdoers, thereby achieving shared moral understanding between the wrongdoer and the wronged party. Fricker argues that communicative blame “aims to bring the wrongdoer to see or fully acknowledge the moral significance of what they have done or failed to do.”⁶ The process of bringing the wrongdoer to acknowledge the moral significance of her actions aims to increase the alignment of moral understanding between the wronged and the wrongdoer.

Remorse, in this context, refers to “a moral perception that delivers a pained understanding of the moral wrong we have done.”⁷ Remorse, while painful, has a well-intentioned aim: to get wrongdoers to see things differently and gain the desire to make positive moral changes in themselves. The psychological pain remorse brings to the wrongdoer is considered a necessary motivational tool to get the wrongdoer to want to change her behavior.

I agree with Fricker that a primary function of blame is to get wrongdoers to see the moral significance of their actions and be motivated to change their behavior. The blamer aims to promote the moral development of the wrongdoer by forcefully getting her to see reasons to act differently next time. This aim is useful and important. We of course want wrongdoers to acknowledge the moral significance of their actions and feel motivated to change their behavior. However, as I will argue, because of its forcefulness or hostility, expressed blame might not be the best way to achieve this aim.

I will assume, in what follows, that expressed blame plays a combination of *affective* (negatively-valenced) and *functional* roles in the aftermath of wrongdoing. This leaves room for expressed blame to take many forms. Affectively, blame might take expression as anger, indignation, resentment, or even disappointment. Functionally, expressed blame serves to make

⁶ Fricker, Miranda. (2016). “What’s the Point of Blame? A Paradigm Based Explanation”. *Noûs*, 50, p. 173.

⁷ Fricker, p. 173.

wrongdoers feel badly for what they did and to motivate them to change. But, how this process of getting wrongdoers to feel badly unfolds depends on the particular blamer as well as the particular dynamic between the blamer and the wrongdoer. We each have different emotional and psychological dispositions. Some of us are quick to anger, some of us are quick to depression. Some of us are quick to take the blame, some of us are extremely defensive. Some of us work well with others, some of us are better off working alone. Blame can surely be other things—a judgment, a belief, an attitude, etc. But, whatever else it is, I argue that it is primarily affective (negatively-valenced), and functions to communicate to the wrongdoer the significance of her wrongdoing and inspire remorse.

III. Blameworthiness Versus Blame & Appropriateness Conditions

In addition to trying to pin down what blame is, many philosophers discuss if, and under what circumstances, blame is *appropriate*. Determining whether one is blameworthy involves determining 1. Who is culpable for a given wrong, and 2. Why. Recall that I am focused on the public practice of expressing blame, not on private judgments of culpability. However, the two are closely related, as many authors take expressed blame to be appropriate *only if* one is culpable for a given wrong. Specifically, many argue that one can only be considered blameworthy, and thus a candidate for blame, for actions over which one had the capacity to control. If one could not but have done a given wrong action, it seems obvious that one is not to blame for the wrong action. Matthew Talbert summarizes this standard view:

Many theorists assume that wrongdoers are open to serious moral blame only if they had the psychological, emotional, and motivational resources to respond to the moral considerations that counted against their behavior. This assumption imposes a moral competence requirement on blame...(according to which) only competent wrongdoers can reasonably be

exposed to moral blame because only these agents are reasonably held to the expectation that they respond to moral considerations.⁸

For example, sometimes blame is a fitting response to wrongdoing, but it is not appropriate to express it.⁹ The two can come apart. For example, sometimes a situation or statement is very funny, but it would be rude to laugh about it given the circumstances (at a funeral, during a breakup, and so on). In this same way, so too is it not always the proper moment to express blame, even if blame is fitting for the situation. We rightly do not always express blame to those we deem blameworthy, due to a variety of contextual factors. I will return to this distinction between fittingness and appropriateness throughout the dissertation.

Additionally, blame comes in degrees. The severity of a case of expressed blame depends on the severity of the wrongdoing. For example, it is not appropriate—in fact, it is nonsensical—to blame to such a degree that I end a relationship with a close friend for her being five minutes late to meet me for lunch. However, it makes more sense to express blame harshly, perhaps to the point of rupturing or even ending the relationship, when I learn that my assistant has been stealing money from me for ten years. The stakes of the case matter a great deal in determining the appropriateness and severity of expressions of blame.

A related question is who has the standing to blame. For example, many argue that if one has engaged in or is currently engaging in the very behavior of the wrongdoer, one's standing to blame the wrongdoer is weakened, if not totally negated. This condition is called the 'hypocrisy condition'. Put simply, the pot cannot call the kettle black. For example, if Rita recently had an affair with a married man, she likely does not have the standing to judge her friend, Sanaz for also having an affair with a different married man. In this case, Rita's judgment of Sanaz is

⁸ Talbert, Matthew (2012). Moral Competence, Moral Blame, and Protest. *Journal of Ethics* 16 (1), p. 89.

⁹ See D'Arms and Jacobsen (2000) for more on the fitting versus appropriate distinction.

hypocritical. The hypocrisy condition does not prevent blame from being appropriate. It may well be that Sanaz's actions are worthy of blame, and she should be blamed for them. But, the hypocrisy condition limits who has the standing to express that blame. Rita might not have the standing to blame Sanaz, but someone else who has not broken the hypocrisy condition regarding this issue might have the standing to blame Sanaz for her actions.

In addition to the hypocrisy condition, there may be agents we simply are not in an epistemic position to appropriately express blame to, such as strangers or acquaintances. In these cases, we might not know enough about the wrongdoer to blame them with full confidence about their culpability. Again, blame might be appropriate in these instances if expressed by those who have the standing to blame. However, strangers may not have the standing to blame each other, because they lack sufficient knowledge of the case. Angela Smith (2007) argues that one's standing to blame depends both on one's relationship to the wrongdoer and on whether one has a "relevant interest or stake in the matter."¹⁰ By relevant interest or stake in the matter, she means that the blamer should likely either know the wrongdoer personally, or should be affected by the wrongdoer's actions in some way, broadly construed. For example, if a stranger starts a giant fire in my apartment building, even though I do not know this stranger, I have a relevant stake in the matter, since my property is at stake because of her actions.

As a second example, it may be inappropriate for me to express blame towards a stranger on the subway who is speaking condescendingly towards his friend. In this case, I may feel that it is not my place to chime in. I lack the standing to express blame since I do not know these people and will, in a matter of minutes, likely never see them again. I do not have any background context for their specific situation, and I do not know what their relationship is like.

¹⁰ Smith, Angela (2007). On Being Responsible and Holding Responsible. *The Journal of Ethics* 11, p. 478.

However, I do have the standing to express blame when my roommate suddenly decides to stop paying her portion of our rent. This is because, following Smith's conditions, I have a close relationship with the wrongdoing agent and a relevant interest in the matter. So, my expressed blame is likely to be appropriate in this case. Assuming I have not committed the same wrong towards her, i.e. suddenly decided to not pay my portion of our rent, my expressed blame is not hypocritical, either.

IV. The Best Case for Blame

Finally, in this introductory chapter, I will briefly explore the underlying features of cases of expressed blame that work well. I will consider the best-case scenario for expressed blame. Surely blame is *sometimes* useful to express. When, specifically, does expressing blame work well. Put another way, when does expressed blame *help* a given interpersonal conflict? By 'work well' or 'help', I argue that the blamer's expression of blame needs to have *practical uptake*. Practical uptake means that the wrongdoer understands the content of the blame and is motivated to make positive moral changes as a result of the blamer's blame. The moral point comes across. In addition, a part of blame *working well* or having *practical uptake* means that *there was no major emotional damage done* by the blaming interaction. Not only does the moral point come across, but it does so without creating lasting emotional damage between blamer and wrongdoer. On my view, such emotional damage (shouting, giving the other person the cold shoulder, nasty comments, vengeful acts, etc.) is not a hallmark of effective blaming practices.

Before I begin my arguments against expressed blame in the chapters to come, I would like to give expressed blame the most charitable interpretation I can. As such, below I briefly

sketch some characteristics of cases in which expressed blame may usefully serve an interpersonal situation. After doing so, I will spend the rest of this dissertation demonstrating that a given interpersonal situation must be so *idealized* for expressed blame to be useful, that on the whole, it is not worth putting any hope into expressed blame usefully serving our interpersonal moral conflicts. More often than not, expressed blame makes things worse for the blamer, the wrongdoer, or both.

Still, I will imagine the characteristics involved in cases in which expressed blame works well. I argue that there are three general elements underlying these sort of cases:

1. The blamer does not express the blame angrily
2. The wrongdoer does not get defensive
3. The blame has practical uptake with the wrongdoer (she learns the lesson and does not do it again)

These three elements are intertwined. Expressed blame's best shot in interpersonal contexts is when the wrongdoer does not get defensive. When the wrongdoer gets defensive, it makes intuitive sense that she is less likely to listen to the blamer's attempts to motivate her to change her behavior. Her emotions forestall positive moral change as she is focused only on defending herself, and not on hearing the other person's perspective. This is especially true when the blamer gets overtaken by anger. In chapter two, I will argue that anger is an exacerbating feature of blame. The blamer's anger, coupled with the wrongdoer's defensiveness, prevents actual, productive dialogue from occurring. Lost in the heat of emotional turmoil, it is hard to understand the moral message that another agent is trying to express.

Indeed, the first two conditions (the blamer does not get angry and the blamer does not get defensive) are *process oriented*. They are conditions involving the process of expressing blame. The third condition is *outcome based*. It is a condition regarding the ways in which the blamer's moral point comes across. The best case for practical uptake is when the wrongdoer is

not defensive, and when neither agent lets their anger overtake the communication. If both agents (blamer and wrongdoer) are able to calmly and rationally respond to a case of wrongdoing, there is certainly a chance that expressed blame can effectively lead to positive moral change.

Expressed blame does sometimes work well. However, so many things have to go correctly in order for blame to work well. First, the blamer has to not overreact, and not let her anger spiral. Second, the wrongdoer has to not get defensive when receiving the blame. Third, the blame has to have practical uptake. The point of this initial discussion was to prompt reflection on when blame has practical uptake, in order to contrast it to the myriad cases I will present in the chapters to come in which blame is counterproductive. As I will argue, these three features (defensiveness, anger, and lack of understanding) are pervasive elements of many of our interpersonal blaming practices, and so much the worse for the positive case for expressed blame.

V. Conclusion

As a preview of the arguments to come, my overarching view is that despite any good intentions that blamers have when they express blame, to carve up the world through the lens of expressed blame in response to wrongdoing is, on the whole, worse for our interpersonal moral lives than not expressing blame is. Indeed, I am arguing something quite radical: *We can, and often should, hold wrongdoers responsible and incite proper motivation in them without expressing blame.* I am not suggesting that no one is ever responsible for a given moral wrong or that we should never respond to or discourage wrongdoing. Rather, I argue that we have good reason to think that expressed blame is not required to hold others responsible. Even though

expressed blame is currently the default response to being wronged in our interpersonal moral lives, we may well be better off without it.

Chapter One. Blame and Retribution

I. Introduction

This dissertation argues that expressing blame in interpersonal moral settings is, generally speaking, *unhelpful* and *unnecessary*, even in its most civilized guises. This is an argument about social expressions of blame, not about private judgments of blame. I propose that alternatives to expressing blame can equally serve the goals of motivating wrongdoers to recognize the significance of their actions and change their ways, as well as holding them responsible for their actions. Not only can we hold each other responsible without expressed blame, but, more strongly, we have reason to think that our interpersonal moral lives will *go better* without expressed blame.

I will argue for this conclusion by first offering a critical discussion of recent literature on the reactive attitudes, retribution, and blame. In particular, Victoria McGeer's recent work provides a defense of a certain form of blame and offers a strong foil to my view. I will mount my case against expressed blame by arguing that blame is fundamentally retributive and that blame's retributive root is incipiently present even in civilized guises.

II. Blame is Retributive

There is good reason to suspect that all forms of blame are at least incipiently retributive. This holds not only for overtly angry forms of blame but also for more civilized, less emotionally-heated forms of blame. This does not imply that all forms of blame will involve

patently retributive elements. Rather, I argue that blame's primitive foundations involve a retributive spirit that remains at least incipiently at play in even our modern-day interpersonal expressions of blame.

There is support in the philosophical literature for the view that blame is, at root, retributive. McGeer discusses the connection between blame and retribution in her two papers "Co-reactive Attitudes and the Making of Moral Community" (2010) and "Civilizing Blame" (2012). In the first half of this chapter, I will provide a critical survey of McGeer's work in order to, in the second half of the chapter, build my case for blame's retributive nature.

In "Co-reactive Attitudes and the Making of Moral Community", McGeer develops and builds upon P.F. Strawson's account of the "reactive attitudes" in his paper "Freedom and Resentment" (1974). As mentioned in my introductory chapter, Strawson first coined the term "reactive attitudes". He describes the reactive attitudes as special affective responses that are deeply ingrained in our interactions with one another. These reactive attitudes include: gratitude, resentment, hurt feelings, indignation, shame, guilt, remorse, and so on. According to Strawson, the reactive attitudes constitute the ways in which we relate to one another in our practices of holding each other morally responsible.

Strawson's account has had wide-reaching influence since it was first published. For example, in *Responsibility and Moral Sentiments*, R. Jay Wallace builds on Strawson's account. Wallace argues that we can use the reactive attitudes as a criterion of responsibility. Specifically, his account "treats the stance of holding people responsible essentially in terms of attitudinal conditions."¹¹ According to Wallace, to hold an agent responsible is essentially to subject them to the blamer's reactive emotions. For example, when you betray my trust and tell my secret to a

¹¹ Wallace, R. J. (1998) "Responsibility and the Moral Sentiments". Harvard University Press, p. 66.

room full of people, I hold you responsible through making you a target of my hostile reactive attitudes, such as anger or indignation. According to Wallace's view, wrongdoers are held to account through being on the receiving end of these attitudes and then responding by making positive changes to their moral behavior.

However, the reactive attitudes do not cover the full range of emotional responses one might have in the aftermath of wrongdoing; they are not exhaustive. McGeer notes that an agent might well feel pity or compassion for a wrongdoer, neither of which are included in Strawson's list of reactive attitudes.¹² It seems unlikely that expressing pity towards a wrongdoer could serve as a practical way to hold a wrongdoer responsible, even though pity is a perfectly sensible response to some, likely minor, forms of wrongdoing. Still, generally speaking, when an agent wrongs us, and we feel that the agent is responsible for the crime, we naturally respond with one (or a few) of the specific negative reactive attitudes.¹³ While being on the receiving end of these negative reactive attitudes is largely unpleasant, McGeer claims that the reactive attitudes deliver a positive and important message.

McGeer explains

Just what are we communicating (to wrongdoers) through our reactive attitudes then? (The reactive attitudes say that) we hold them accountable to an ideal of moral agency because we think them capable of living up to that ideal. So reactive attitudes communicate a positive message even in their most negative guise – even in the guise of anger, resentment, indignation. The fact that we express them says to the recipients that we see them as individuals who are capable of understanding and living up to the norms that make for moral community.¹⁴

¹² McGeer, Victoria (2012). "Co-Reactive Attitudes and the Making of Moral Community". In *Emotions, Imagination and Moral Reasoning*, eds., C. MacKenzie & R. Langdon. Macquarie Monographs in Cognitive Science, Psychology Press, p. 2.

¹³ It should be noted that not all individuals are appropriate targets of our reactive attitudes. The standard view is that individuals with cognitive-abnormalities and young children might not be in the position to have known better when they acted wrongly. Only agents who are participants in our shared moral norms are proper targets of the reactive attitudes.

¹⁴ McGeer, p. 7.

In this passage, McGeer defends the positive message conveyed through the harsh affective expression of the reactive attitudes. When we communicate blame towards a wrongdoer through our hostile emotions like anger or resentment, McGeer claims that this anger positively and helpfully communicates that we see the wrongdoing agent as someone capable of knowing better and acting more in line with our shared moral norms. We are communicating, through our harsh reactive attitudes, that we demand that the wrongdoer acts better next time.¹⁵ This communication shows wrongdoers that we respect them as moral agents who should know better than to act as they did. Further, it encourages them to act better next time.

On McGeer's view, then, there is nothing *wrong* with communicating our negative reactive attitudes to wrongdoers as a way to show respect towards them and to get them to act better next time. She calls the responsiveness a wrongdoer brings to being the target of a blamer's reactive attitudes *co-reactivity*. McGeer defines co-reactivity as "a sensitivity to the scaffolding structure of reactive emotions that is displayed by most human beings most of the time."¹⁶ Responsible agency requires just this sort of co-reactive sensitivity. Indeed, this was one of Strawson's own points. Strawson argued that the reactive attitudes are a way to hold other agents accountable in an essentially *interpersonal* and *interactive* way, and that this process involves a response from the wrongdoer.

Similarly, in "Respect and the Second Person Standpoint" (2004), Darwall expands on Strawson's point. Darwall argues that a second-personal reason "is one whose validity depends on presupposed authority (hence accountability) relations between persons and, therefore, on the possibility of the reason's being *addressed* person-to-person."¹⁷ McGeer's proposal regarding the

¹⁵ See also Darwall (2006) for a similar argument.

¹⁶ McGeer, p. 1.

¹⁷ Darwall, Stephen (2004). "Respect and the Second Person Standpoint". *Proceedings and Addresses of the American Philosophical Association*, Vol. 78, No. 2, p. 46.

co-reactive structure of blame involves reasons of just this second-personal variety. Darwall further argues that second-personal reasons “always derive from agents’ relations to one another, (so) they are invariably agent-relative in some way or other; they apply to us from within the network of these relations.”¹⁸ Darwall highlights the context-sensitive nature of our interpersonal relationships. How we respond to wrongdoers and how they respond to us, in our interpersonal lives, depends on the dynamics involved in that specific relationship and other context-sensitive factors about the given situation.

McGeer accepts Strawson’s picture of how the reactive attitudes arise in our interactions with others, particularly in our practices of holding each other responsible. Further, she combines this Strawsonian picture with recent cognitive science research about the reactive attitudes to strengthen her defense of blaming. For example, she discusses Greene and Cohen’s paper “For the Law, Neuroscience Changes Nothing and Everything” (2004), in which the authors argue that our reactive affective system is part of our evolved biological heritage, making it “very unlikely to be cognitively penetrable.”¹⁹ Greene and Cohen agree that although “we may come to regard our reactive attitudes (e.g. resentment, indignation, retributive anger) as embedding a hopeless error in the way we regard human agents, we may be stuck with such attitudes in the hurly-burly of everyday life.”²⁰ By this, Greene and Cohen mean that no matter how evolved we have become, to some degree, we are *stuck* with our primitive-based negative reactive attitudes.

Put more strongly, Greene and Cohen suggest that these reactive attitudes, including retributive feelings, are likely to be an *ineradicable feature of the human psyche*²¹ that are “driven

¹⁸ Darwall, p. 46.

¹⁹ Greene, J. & Cohen, J. (2004). “For the Law, Neuroscience Changes Nothing and Everything”. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*: 359, p. 1784.

²⁰ McGeer, p. 16.

²¹ McGeer’s phrase not mine (McGeer, p. 23).

by phylogenetically old mechanisms in the brain.”²² Try as we might, we cannot simply get rid of or evolve past our negative reactive attitudes. This is a substantial and, on first pass, seemingly hopeless conclusion. The conclusion is that while we may be able to train and tame these retributive feelings through habitual effort, we simply cannot eliminate them from our brains. This gives us reason to think that those feelings will, to some degree, continually crop up in our interpersonal moral lives, despite our best efforts to tame them. This finding mirrors Adam Smith’s view in *A Theory of Moral Sentiments*, in which he argues that “revenge, the excess of resentment, appears to be the most detestable of all the passions, and is the object of the horror and indignation of everybody.”²³ Even though revenge is a natural habit and one we likely cannot eliminate, Smith finds it the worst of all our emotional responses, deeming it a “savage disposition”²⁴ to embody.

Optimistically, however, Greene and Cohen do not think that the fact that retributive feelings are ineradicable features of our human psyches should force us to make legal and policy decisions that involve retributive forms of punishment. We should not feel resigned to a retributive-based legal system. Instead, Greene and Cohen suggest that we might be able to *bracket-out* our reactive attitudes when we make policy decisions about criminal justice, as well as when we make moral decisions in our everyday interpersonal lives. By *bracket-out*, Greene and Cohen mean that we might be able to design institutional forms of justice that intentionally do not cater to these negative impulses. Just because these negative attitudes and impulses will likely continue to arise in our minds regardless of how much we train to eliminate them does not mean we will necessarily be compelled to always act on them. Instead, we can train ourselves to

²² Greene & Cohen, p. 1784.

²³ Smith, Adam (1759). *The Theory of Moral Sentiments*. Gutenberg Publishers, p. 68.

²⁴ Smith, p. 160.

be aware of those reactive attitudes arising and take care to not let them solely dictate our legal policy decisions.²⁵ In the same way that many of us would prefer to eat cookies and ice cream every day, but are able to train ourselves to moderate our sugar intake because we value our health and longevity, perhaps we can also train ourselves to be aware of and moderate the expression of our reactive attitudes with consistent effort.

Greene and Cohen's finding applies not only to the legal realm, but also to our daily interpersonal moral lives, too. It is important to consider whether the fact that retributive feelings are ineradicable features of our human psyches forces us into a retributive-based interpersonal moral system. Greene and Cohen do not elaborate on whether the ability to bracket-out our retributive impulses in our daily moral lives is practically available to most of us. Perhaps the moral saints among us are able to express blame in a highly civilized manner, void of any of these primitive retaliatory impulses. But, this may not hold true for the rest of us. It is likely unrealistic to expect the average moral agent to so diligently and constantly bracket-out their retributive impulses. In the next section, I will consider the strategy of civilizing blame, which will help shed light on the plausibility of our ordinary abilities to bracket-out our retributive impulses.

III. Civilizing Blame

The project of civilizing blame aims to show that blame can be a “morally acceptable, even constructive feature of human life.”²⁶ An extreme version of *Civilized Blame* is what

²⁵ For further, related discussion, see Martha Nussbaum's *Upheavals of Thought* (2001) where she argues that we must not let emotions dictate the law.

²⁶ McGeer, Victoria (2013). “Civilizing Blame”. in *Blame: Its Nature and Norms*, eds., D. Justin Coates and Neal A. Tognazzini. Oxford University Press, p. 1.

philosophers refer to as *Sanitized Blame*, a view that McGeer finds psychologically unrealistic. Advocates of sanitized versions of blame argue that even if blame involves underlying retaliatory elements, perhaps these baser instincts can be sanitized, if not wholly eradicated. According to McGeer, it is simply not possible to purge blame of all its unsavory psychological features. The Sanitized Blame strategy is unrealistic because part of what it is for humans to respond to wrongdoing naturally is for us to express our reactive attitudes to one another, including our retributive impulses. Perhaps the moral saints among us are capable of such a drastic eradication of their baser instincts, but this achievement is not available to the rest of us.

A less extreme version of this is Civilized Blame, which McGeer explains as aiming to “preserve (blame’s) normative power while purging its unsavory psychological features.”²⁷ McGeer endorses a version of Civilized Blame under the title *Domesticated Blame*, which I will survey in the next section. Assuming it is at least psychologically possible to dispense with the negative emotional elements of blame, McGeer examines whether this would be a *normatively preferable* state of affairs. She recognizes and agrees with me that blame’s affective elements

Can be highly destructive, leading to cruel and excessively punitive behavior. They can also be self-destructive, leading to unpleasant ruminations, loss of trust, loss of more general goodwill...They can generate downward spirals, where one person’s anger begets what may seem to others justified anger and resentment in turn, refueling the anger originally felt and encouraging more anger in return.²⁸

Despite this recognition, McGeer argues that as dangerous as the blaming emotions are, “the state of affairs in which such emotions are missing or absent from blame should not be recommended on normative grounds.”²⁹ McGeer defends this claim with the premise that emotions constitute a *uniquely powerful* form of communication. McGeer argues that we need

²⁷ McGeer, p. 1.

²⁸ McGeer, p. 26.

²⁹ McGeer, p. 27.

the forcefulness of anger, indignation, and resentment in order to make salient the contents of our expressions of blame. Without the forcefulness of these negative emotions, the wrongdoer might not clearly get the message that what they did was wrong. McGeer explains that biologically speaking, we are

Involuntarily disposed to react to—that is, both detect and *care about*—the good or ill will that others manifest toward us... (reactive attitudes) are attention-grabbers for us, providing a form of triangulation that draws our attention both to the person manifesting the attitude and to the behavior—whatever it was—that prompted the attitude.³⁰

According to this view, we are involuntarily disposed to reacting negatively when we are on the receiving end of emotionally-laden forms of blame. The forceful charge of affective expressions of blame, then, incites a reaction in wrongdoers that non-emotional expressions of blame may not. McGeer, then, rejects the Sanitized Blame strategy on the grounds that it is unrealistic and that the reactive attitudes successfully grab wrongdoers' attention and force them to respond. To strip blame of these forceful components would not yield as successful a moral outcome. McGeer argues that blaming emotions “convey to transgressors both how seriously blamers regard the offense and their sincerity in pressing their normative demands.”³¹

Thus, McGeer concludes that “if we do not wear the costs of angry blame, then we cannot reap its considerable rewards... (we should not try) to take the emotional guts out of blame.”³² She is aware of the unsavory aspects of blame, i.e. the forcefulness and retributive nature of our blaming emotions. However, McGeer concludes that those negative emotions are necessary means to the important end of getting wrongdoers to forcefully recognize their wrongdoing and to commit to acting in morally better ways.

³⁰ McGeer, p. 27.

³¹ McGeer, p. 28.

³² McGeer, p. 29

While McGeer and I are in full agreement both that blame fundamentally involves retributive elements and that the Sanitized Blame strategy is ineffective, our arguments diverge beyond these points of agreement. Like McGeer, I also find the Sanitized Blame strategy psychologically unrealistic, but I disagree with her that we need to express the reactive attitudes in order to most successfully get wrongdoers to recognize their wrongdoing and commit to changing. That we disagree on this point is very important, because I conclude, in the chapters to come, that the social, ethical, and practical risks involved in expressing blame make it an ill-advised interpersonal moral strategy, whereas McGeer's conclusion is in favor of continuing our blaming practices. Before I argue for why I have come to this conclusion, I will give McGeer's account its due, and will discuss her defense of a form of blame, which she terms Domesticated Blame. Given that McGeer finds the Sanitized Blame strategy unsatisfying, she instead proposes an account of blame, Domesticated Blame, that aims to tame (but not shy away from) our retributive impulses. I will first summarize her account, and then provide objections to her view, of a kind that will reveal in what respects my position is much more radical than hers.

IV. Domesticated Blame

McGeer argues that her alternative to civilizing blame, Domesticated Blame, is a happy medium between blame's primitive, retributive origins and overly-sanitized versions of blame. McGeer's strategy is a form of blame, and one that she argues more accurately depicts the psychology of blame as a "complex, emotionally charged reaction to wrongdoing, with inseparable savory and unsavory features."³³ Rather than transform the psychology of blame

³³ McGeer, p. 1.

(Sanitized Blame’s futile goal), McGeer’s domestication method aims to reform our very *institutions* of blame, while leaving the psychological origins of our blaming practices intact. McGeer’s strategy does not attempt to sanitize or strip blame of its retributive origins, but rather aims to get at the foundational, structural underpinnings of our blaming practices.

McGeer urges her readers to recognize that the ‘affective complex’ associated with blame, while unsavory—which she terms “warts-and-all”—still provides valuable normative work “so long as it is constrained by social and institutional practices that support its more constructive features.”³⁴ In this way, Domesticated Blame retains all of the negative emotional aspects of blame, but constrains them within a structured social and institution model that encourages morally productive dialogical exchanges between blamer and wrongdoer.

At the heart of McGeer’s strategy is a dialogical exchange between blamer and wrongdoer. Recall my earlier discussion of the co-reactivity involved between blamer and wrongdoer in response to moral wrongdoing. That co-reactivity is central to the dialogical exchange. For McGeer, blame grounded in a dialogical exchange is normatively acceptable because the dialogical dimension is what makes blame “no longer just a primitive punitive response but one that builds onto this crude regulative structure a cognitively sophisticated dialogical dimension.”³⁵ According to McGeer, Domesticated Blame is not merely a hostile and retributive exchange because it involves a non-hostile, co-reactive back and forth dialogue between the agents.

This co-reactive dialogical exchange involves two elements. First, the blamer *responds to* something a wrongdoer has done. Second, the blamer *calls for* the wrongdoer to appropriately respond to the blamer’s sense of being wronged.³⁶ The two agents co-react to the situation and to

³⁴ McGeer, p. 4.

³⁵ McGeer, p. 30.

³⁶ McGeer, p. 30.

what one another is saying in the moment. The wrongdoer is allowed, even encouraged, to defend herself against the blamer's charges if she feels that they are unfair. The dialogue encourages each agent to explain their own perspective. McGeer further argues that the point of this dialogical exchange is to

Draw wrongdoers into a kind of exchange where they are perforce challenged to exercise their capacities as responsible agents, to reflect on what they have done, whether or not it is legitimate, and if it is not, to take responsibility for what they have done and for what they will do in the future. Thus despite the fact that such a process aims to change a person's actions and attitudes, it is not by means of the blamers' exercising engineering control.³⁷

In this passage, McGeer explains that the dialogical force of blame is, just as Darwall argued above, essentially *interpersonal*. It involves the two (or more) agents jointly reflecting on the wrong moral actions that occurred, and then jointly coming to a consensus about how to move forward. Importantly, McGeer admits that the dialogical process aims to change the wrongdoer's actions and attitudes, yet McGeer maintains that this is not equivalent to the blamer's attempting to gain control. The dialogical exchange is not an attempt on behalf of the blamer to gain control because the wrongdoer has the chance to respond to the charges of blame. It is not a one-sided verbal exchange, but rather is an active, ongoing, collaborative process between the agents.

McGeer further explains that

Drawing putative wrongdoers into such an exchange, and thereby fully crediting them with the capacities of responsible agency, opens blamers themselves to being challenged on the legitimacy of their blame.³⁸

Indeed, the dialogical exchange between blamer and wrongdoer that McGeer envisions aims to create space for wrongdoers to put the burden of justification back onto the blamer, if the wrongdoer feels that the blamer has unfairly charged them with the wrongdoing. Pushback from the wrongdoer is an expected and inevitable part of the dialogical exchange.

³⁷ McGeer, p. 25.

³⁸ McGeer, p. 25.

For instance, the dialogical exchange might involve a candid discussion of fault between the two agents. The wrongdoer might offer excuses or explanations of her actions. She might apologize. She might commit to reforming her actions. She might defend her behavior as being non-culpable. At this point, the blamer then has another chance to respond to the wrongdoer's own response. The back and forth dialogue might consist of several rounds of exchange. McGeer's account does not limit the number of back and forth exchanges the agents can or should have. Whichever form the conversation takes, dialogical exchange aims to reinforce both of the agents' commitments to their shared moral norms. This dialogical exchange builds on the fact that we are reactively sensitive creatures, and are emotionally primed to respond to one another through this sort of dialogue.

McGeer finds this dialogical dimension of our blaming practices to be normatively valuable "both from the point of view of repairing the variety of wrongs done by any transgression (to the victim, to the stability of shared norms, to a sense of security, and so on) and from the point of view of regulating behavior going forward."³⁹ McGeer's dialogue-based Domesticated Blame strategy thus aims to create a structure for fruitful dialogue *within a system of blame* to occur between the blamer and wrongdoer with an eye towards lasting moral change.

V. Objections to McGeer

A. Building on Top of a Retributive Core

While I am quite sympathetic to McGeer's dialogical strategy and her focus on constructive communication between the agents involved, I have two objections to her view.

³⁹ McGeer, p. 30.

These objections support my view that because expressed blame comes with a host of social risks, we would be best served avoiding expressed blame in our interpersonal moral lives.

My first objection is that McGeer's account is overly optimistic about the practicality of building a domesticated moral system *on top of* a fundamentally retributive core. McGeer writes that this dialogical dimension is what makes blame "no longer just a primitive punitive response but one that builds onto this crude regulative structure a cognitively sophisticated dialogical dimension."⁴⁰ The dialogical dimension might *superficially* mask the primitive foundations of blame, but it does not make those primitive foundations disappear.

Building a more domesticated form of blame on top of an undomesticated foundation seems both strange and risky. Why not build a new system of responding to wrongdoing, one that is not built on top of an existent, primitive structure? Blame's retributive core is much more powerful and enduring than McGeer acknowledges. In the same way that living in an earthquake fault zone always comes with the active risk of earthquake-induced damage to one's home, so too does building a system of responding to wrongdoing on top of a dangerous platform—retribution—always come with risks. We see frequent, seemingly glorifying depictions of retribution in our media, for example in the recent films *The Revenant* and *Inglorious Bastards*. Exerting retributive aims when others wrong us seems to be an ordinary feature of our daily moral lives both on screen and off. Relatedly, recall Greene and Cohen's persuasive argument that our retributive feelings are likely to be an *ineradicable feature of the human psyche*⁴¹ "driven by phylogenetically old mechanisms in the brain."⁴² Because of this, our retributive impulses can, in mere seconds, devolve back into purely retributive (without McGeer's domesticated dialogical

⁴⁰ McGeer, p. 30.

⁴¹ McGeer's phrase not mine (McGeer, p. 23).

⁴² Greene & Cohen, p. 1784.

structure) forms of blame, depending on the situation and dynamic between the agents involved. A retributive foundation is not something we can so easily build a sophisticated moral system on top of.

In light of this, McGeer is overly optimistic to think we are able to tame our retributive impulses, generally speaking. Although we might be able to tame our retributive impulses *on occasion*, and we certainly can attempt to train ourselves to do so, we cannot, so readily, domesticate blame's retributive core. To some degree, many of us have domesticated our retributive impulses fairly successfully in our daily moral lives. In our country on the whole, we peacefully exist in our shared moral communities much of the time without much trouble. However, the process of domestication is rigorous and ongoing. Except for perhaps the moral saints among us, we have not (nor can we) domesticate our primitive retributive impulses, full stop.

Even if McGeer is right to think that some forms of blame are more domesticated, and many steps evolved from this primitive foundation, *still*, these domesticated forms of blame remain descendants of more primitive forms of blame. Even if one's expression of blame *appears* quite civilized, we must remember that our evolutionary impulse is to retaliate, and that we are not so far evolved from that root impulse.⁴³ Therefore, because the dialogical exchange at the heart of McGeer's domesticated blame strategy is built on top of a retributive foundation, there is always a chance of it devolving (and devolving quickly) into its retributive core.

In response to my objection, McGeer might claim that there is no obvious harm to building a more domesticated version of blame on top of its primitive version. However, in response to this line of thought, I still argue that building a new version of blame on top of a

⁴³ See Haidt (2001) for further details

primitive foundation is not worth the escalated social risk. Further, I argue that there are additional harms to McGeer's domesticating blame strategy. Expressions of blame, even in their most domesticated or civilized guises, are rarely affectively neutral. Indeed, expressions of blame often induce a visceral, emotional reaction from both the blamer and the wrongdoer. As I argued in my introductory chapter, part of what it is to express blame is to express negative emotions (anger, indignation, resentment) at a wrongdoer with the aim of inciting moral change in them. When expressions of blame are especially hostile and heated, there is a specifically *social* harm involved. Sometimes, the rebuke involved in expressed blame is not merely a normatively symbolic retributive mechanism, but it is also a psychological sting explicitly aimed to incite remorse, guilt, or shame in wrongdoers. There is substantial empirical evidence for the view that moral judgments are emotional in nature (see Prinz (2006); Moll, de Oliveira-Souza, and Eslinger (2003); Greene et al. (2001)).

We aim to psychologically get revenge on each other in response to wrongdoing in all sorts of affective ways: passive aggressive verbal censure, aggressive body language, emotionally (or physically) distancing oneself from the wrongdoer, and so on. Each of these strategies involve a social harm in and of itself. It is hard to create space for the type of authentic dialogue that McGeer's strategy requires when one (or both) agents is predominately focused on psychologically wounding the other person by discharging their negative emotions onto that person.

In response, it may be remarked, what is so wrong with retributive expressions of blame in our moral lives? Someone might agree with me that blame is retributive, but also think there is nothing wrong with this. In response, I argue that all expressions of blame (however civilized)

remain socially risky. That social risk may be manifestly present, or may just be incipient, but the risk is there in either case. One obvious social harm is escalation.⁴⁴

For example, when I express blame to my friend for repeatedly running late and cancelling plans we made at the last minute, I might give her the cold shoulder or behave in a passive aggressive way when she finally shows up. I am not only expressing to her that I blame her, but I am further punishing her in a psychologically painful way. I want her to feel remorse. I want her to feel badly. In that moment, I am not trying to communicate authentically with her to resolve the conflict. Instead, I want her to feel badly because she has hurt me. I felt disrespected by her not valuing my time, and I want her to psychologically suffer in kind. This intention is a form of psychological retributivism, one that creates social harm.

The social harm is twofold. First, it is a social harm because the blamer's psychological retributivism aims to hurt the wrongdoer. Second, it is a social harm because this retributive aim is not only harmful for the recipient, but makes it harder for the agents to authentically communicate about how to resolve the wrongdoing. The productive and authentic dialogical exchange that McGeer envisions becomes hard to create in a psychologically hostile atmosphere. It is harder to resolve interpersonal moral conflicts when retributive aims are at play. In the case described above, I have deemed my friend's tardiness and flakiness as an expression of her unfairly treating my time as less valuable than hers. The harm I want to come to her is due to her unfairly harming me. This is psychologically harmful to her, and it is harmful to our attempts at resolving the conflict because I am too focused on trying to make her feel bad. That focus on

⁴⁴ Because my dissertation focuses on expressed, interpersonal blame, rather than private or self-blame, I am most worried about the social effects of blame. If I were focused on private blame, the social effects might be lessened. But even private forms of blame might manifest publicly in subtle ways, such as passive-aggressiveness, social withdrawal, etc. If private blame gets communicated at some level (even non-verbally), then it's incipiently expressed.

making her feel bad restricts my ability to engage in inquisitive moral dialogue with her. This further social harm provides an additional reason for why McGeer's Domesticated Blame strategy is overly optimistic about our ability to build a sophisticated and affectively-tame structure on top of a primitive retributive foundation.

B. Bad is Stronger Than Good: A Second Social Harm Argument

My second objection to McGeer's view involves an additional social harm invoked by the Domesticated Blame strategy. The additional social harm is that expressed blame creates a dynamic of resistance from both blamer and wrongdoer, a dynamic which produces a socially and morally unhelpful outcome. This is because each agent is wired to be resistant to change (indeed, they are likely to dig in their heels), after the blame has been expressed.

Empirical research in Baumeister, Bratslavsky, Finkenauer, and Vohs' paper "Bad is Stronger Than Good" (2001) sheds additional light on this resistance. The authors argue for the general psychological principle that bad is stronger than good.⁴⁵ This applies to the greater power of: bad events over good ones, bad feedback over good feedback, and bad stereotypes over positive stereotypes. In general, bad impressions, emotions, and information form more quickly in our minds and are more resistant to disconfirmation than good impressions, emotions, and information. Applying this insight to the case of interpersonal blame, the thought is that events that are negatively valenced will have a greater impact on agents than positively valenced events of the same type.⁴⁶ So, if I simultaneously blame and praise you, it is likely that my blame (because it is negatively valenced) will have a greater impact on you than my praise (because it

⁴⁵ Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). "Bad is stronger than good". *Review of General Psychology*, 5, pp. 323.

⁴⁶ Baumeister, Bratslavsky, Finkenauer, & Vohs, p. 323.

is positively valenced). Similarly, you will be more likely to remember my blame than my praise, because it is negatively valenced. My expressed blame likely has a more lasting negative impact on you than any positive impact my praise has.

Other researchers have confirmed this finding, calling it the *positive-negative asymmetry effect*.⁴⁷ Generally speaking, Baumeister, Bratslavsky, Finkenauer, and Vohs argue that “learning something bad about a new acquaintance carries more weight than learning something good about them, by and large.”⁴⁸ This makes intuitive sense. None of us likes to receive criticism, blame, or censure. It can be very hard for us to shake off the negative impressions of others, even in light of witnessing their positive attributes. Given that the content of blame is negative, and given that negative events will likely have a greater impact on us than positive events, there is a privileging of the negative events over the positive events in instances of expressed blame. The bad actions that an agent takes may carry too much weight in blamers’ minds, and the good actions that an agent takes may carry too little weight in the blamers’ minds. This bias impacts how we weight wrongdoer’s bad moral actions.

Baumeister, Bratslavsky, Finkenauer, and Vohs further explain that “Close relationships are more deeply and conclusively affected by destructive actions than by constructive ones, by negative communications than positive ones, and by conflict than harmony.”⁴⁹ This asymmetry cannot help but affect the interpersonal dynamics between blamer and wrongdoer. In addition, “bad moods and negative emotions have stronger effects than good ones on cognitive processing.”⁵⁰ Simply put, bad emotions have more power and lasting impact on us than positive emotions do.

⁴⁷ See Anderson (1965); Peeters & Czapinski (1990); Skowronski & Carlston (1989).

⁴⁸ Baumeister, Bratslavsky, Finkenauer, & Vohs, p. 324.

⁴⁹ Baumeister, Bratslavsky, Finkenauer, & Vohs, p. 355.

⁵⁰ Baumeister, Bratslavsky, Finkenauer, & Vohs, p. 355.

Although Baumeister, Bratslavsky, Finkenauer, and Vohs do not address expressed blame directly, they do conclude their article with a caution about punishment. They write that “punishment may not be optimal for education, even if it does produce optimal learning, because the side effects of punishment can be damaging.”⁵¹ By this, they mean that while punishment might create a positive environment for *learning*, the *harmful emotional effects* of punishment may outweigh the educational benefits. That is, even if we learn best (morally speaking) through expressed blame, it must be remarked how emotionally wounding receiving expressed blame can be. The learning outcome should not be prioritized above all else. Indeed, the sacrifice of creating a successful learning environment is that we create a hostile emotional and psychological environment, due in part to the positive-negative asymmetry effect. Other agents’ expressions of blame might motivate us to change, but it often comes at the great cost of emotional and social harm. A practice that is emotionally damaging, even if it produces good learning outcomes, is not a best practice.

In our daily moral lives, most of us do not exhibit much, if any, caution about expressing blame; we are largely blind to the positive-negative asymmetry effect. Additionally, our tendency to give negative emotions more power than positive emotions means that we might *exaggerate* the appropriate emotional reaction to a case of moral wrongdoing. Because we tend to give our negative emotions more attention than our positive ones, we might tend to give our angry or resentful blaming emotions undue credence. When analyzing the facts and deciding whether to express blame to someone, we may be giving unfair weight to our negative emotions about the wrongdoer.

⁵¹ Baumeister, Bratslavsky, Finkenauer, & Vohs, p. 362.

These observations, then, strongly encourage the idea that all expressions of blame, even the highly civilized versions among them, belong to a family of response that are, at root, retributive, and emotionally damaging. Though our retributive impulses will not always arise, there is always a risk that they will arise when expressing blame, and the results of these impulses arising can be quite socially risky, due to the volatile and heated nature of our retributive impulses. For this reason, we are better off, on the whole, moving away from expressed blame as the default mode of response to wrongdoing. This, again, is a much more radical claim that McGeer makes. McGeer’s account aims to domesticate blame. Differently (and more radically) I aim to move our interpersonal habits of moral response away from blame altogether.

VI. Fittingness Versus Appropriateness

There is an important distinction at play in cases of retribution and expressed blame, and one that I will continue to utilize throughout this dissertation. This distinction is motivated by Justin D’Arms and Daniel Jacobson’s paper “The Moralistic Fallacy: On the ‘Appropriateness’ of Emotions”. D’Arms and Jacobson argue that an emotion “can be fitting despite being wrong to feel”.⁵² D’Arms and Jacobson offer the following example. Imagine that you are at a funeral and your Uncle slips and falls while walking to his seat. It might be fitting to laugh at him—the situation is undoubtedly humorous—yet it might also be inappropriate to laugh, given the somber setting. Other guests might not appreciate the humor in the situation, given the somber setting. This example demonstrates the general claim that certain emotions can very well *fit* an event

⁵² D’Arms, Justin and Daniel Jacobson (2000). “The Moralistic Fallacy: On the ‘Appropriateness’ of Emotions”. *Philosophy and Phenomenological Research*, Vol. 61:1, p. 65.

(laughing at someone falling), without being *appropriate* to feel in the specific context (a funeral).

The authors claim that it is a moralistic fallacy to infer “from the claim that it would be wrong or vicious to feel an emotion, that it is therefore unfitting.”⁵³ Instead, the authors argue that an emotion can be fitting despite being wrong (or inexpedient) to feel. As such, “the wrongness of feeling an emotion never, in itself, constitutes a reason that the emotion fails to be fitting.”⁵⁴

D’Arms and Jacobson divide their considerations of fittingness into two kinds: shape and size. The authors offer the following example:

Suppose you envy Susan the cabin in the woods where you believe she spends her weekends. Your envy would be unfitting on the grounds of shape if the cabin is primitive and you’re keen on comfort, or if she really has no cabin and spends the weekends caring for a sick relative.⁵⁵

In this example, my envy’s shape would not match the situation in either of these cases, because I have no real grounds for envy over Susan’s situation. If the cabin is primitive, then there is no need for me to be envious of Susan, since I have better accommodations in my current apartment. Likewise, if Susan goes to her cabin in order to care for a sick relative, again my envy makes no sense (it does not *fit* the situation), because caring for a sick relative is painstakingly emotional and involves arduous physical labor, a circumstance far removed from an idyllic vacation.

My envy can also fail to fit with regard to its size. Perhaps Susan’s cabin in the woods is luxurious and I am quite right to be envious of her (in this case, my emotion has the right shape). However, if I am *exceedingly envious*, my emotion is an overreaction; it is the wrong size. If I am envious of every aspect of Susan’s life merely on the grounds of her having a luxurious

⁵³ D’Arms and Jacobson, p. 69.

⁵⁴ D’Arms and Jacobson, p. 69.

⁵⁵ D’Arms and Jacobson, p. 73.

cabin, I am overreacting. My envy is too large. I do not have the proper information about Susan in order to be envious of her in all areas of her life. The size of my excessive envy does not fit the situation.

Beyond the fittingness of size and shape of our emotions, we must attend to the moral reasons at stake in our expressions of our emotions. D'Arms and Jacobson offer another helpful example:

If you are widowed with young children, you must bring them up as best you can. Too much grief risks further harm to them, so it is incumbent upon you not to fall apart. Since the children need to go on with their lives, with as much security and as little trauma as possible, it would be wrong to indulge in the fitting amount of sorrow—the amount that accurately reflects the sadness of the situation. But this is not to suggest that the loss of a spouse isn't all that sad. Instead, it demonstrates that some moral reasons not to feel grief are irrelevant...in particular, those trading on the consequences of feeling some way, for yourself or others.⁵⁶

This example demonstrates that although widow's grief is very *fitting* for her situation, there are other considerations—the wellbeing of her children—that impact how the widow should act and impact the *appropriateness* of her grief. Further, the example reveals that a period of grief in which one cannot function in an everyday way might be fitting, and yet, if one has responsibilities—i.e. parenting—it cannot be morally appropriate to feel the grief, all things considered. Emotions may fit their object, despite being wrong to feel. Beyond grief, D'Arms and Jacobson claim that this holds for all basic emotions, such as anger, guilt, amusement, and jealousy.⁵⁷

I will now apply their argument to the specific context of expressed blame. Taking D'Arms and Jacobson's point, one's blaming emotions can be fitting, despite being wrong to feel or inappropriate in a given situation. The blaming emotions are often fitting to feel when one has been wronged, and they may indeed serve a useful role, but generally speaking I argue that they

⁵⁶ D'Arms and Jacobson, p. 77.

⁵⁷ D'Arms and Jacobson, p. 82.

remain *wrong to express* because of the harms that expressed blame can so easily bring in its wake.

In light of this, I propose that even when expressed blame fits a situation, it is often inappropriate, even wrong, to express, due to its retributive root and the social risks involved. I wholly agree that many of the things philosophers say are blameworthy are, indeed, morally bad and blameworthy, but we should not express blame in response to these actions. We can privately judge something blameworthy (and find the expression of blame fitting) without expressing blame. I will make use of this fitting versus appropriate distinction as applied to the case of expressed blame in the chapters to come, as well. For now, my goal in introducing this distinction has been to highlight the ways in which expressing blame and blameworthiness can come apart, as well as the ways in which expressed blame can be inappropriate to express, even if it fits the situation.

VII. Vargas

Finally, I will introduce Manuel Vargas's account of blame to explore how we can achieve blame's instructive and motivating aims without expressing blame. The goal of this critical overview of Vargas's work is to make the idea that we can understand our moral lives without expressed blame seem more plausible. I will extend this line of argumentation—that we can, practically speaking, make sense of our interpersonal moral lives without blame—at length in chapter four of this dissertation.

Vargas argues that blame serves an instrumental role in providing agents the opportunity to grow their moral agency and learn to act in morally better ways in the future. That is, blame can

be instrumentally useful insofar as it helps agents grow in morally positive ways. When the norms of blame are internalized by a wrongdoer, those norms can motivate that agent to act in line with the correct moral considerations in the future. This point is very much in line with McGeer's discussion of blame. We saw that McGeer argues that even though blame has an unsavory retributive root, still it can serve an instrumental role in a modified form. McGeer writes that "The fact that we express (the reactive attitudes) says to the recipients that we see them as individuals who are capable of understanding and living up to the norms that make for moral community."⁵⁸

However, in line with what I have argued up to this point, Vargas does not think that responsible agency always tracks being blameworthy. For example, he distinguishes between responsible agency and being responsible for some bit of behavior. He explains that

One can be a responsible agent at some time—alive to relevant moral considerations—without it following that one is blameworthy for some act. And, in cases where there are no moral considerations in play, where the considered act is devoid of moral significance, the issue of being morally responsible for some behavior is simply not live.⁵⁹

Moreover, Vargas thinks an agent can be responsible in one context, yet not responsible in a different context. There is a deep context sensitivity involved in our determinations of responsibility. A main utility of our moral responsibility system, again, is that the norms of blame help motivate agents to act better. This is a purely instrumental case for blame. It is not the argument that blame is always necessary to express in every case of wrongdoing. Rather, it is the argument that there are context-sensitive considerations at play in determining one's responsibility and, further, that when it makes sense to blame a given agent, doing so is an instrumental good. In this way, blame *can be* justified on the basis of the good effects it

⁵⁸ McGeer, p. 7.

⁵⁹ Vargas, Manuel (2013). *Building Better Beings: A Theory of Moral Responsibility*. Oxford University Press, p. 239.

produces, such as wrongdoers' motivation towards pursuing moral growth. However, blame is not automatically justified for a given case of wrongdoing.

Relatedly, at the end of *Building Better Beings*, in a chapter titled "Blame and Desert", Vargas discusses the ways in which blame can also be a liability in our social relationships. As important as holding others responsible is, he suggests that "it does not obviously trump all other considerations or concerns that can matter to a human life."⁶⁰ He argues that

Sometimes blaming will be (locally) very costly, and perhaps even costly enough that it makes sense to maintain that one should not blame those who deserve it. In such cases forgiveness or even ignoring the transgression may, all things considered, be the better route.⁶¹

Vargas rightly argues that sometimes blame is too costly from a social vantage point, and thus not worth the risks. This sentiment mirrors my argument in the previous section that even if blame is fitting, it might not be appropriate to express, due to blame's retributive core and the heightened social risks involved.

Vargas goes on to explain that forgiveness itself can be an alternative to blame. He explains that forgiveness can be administered without "turning a blind eye towards the moral significance of the transgression."⁶² This means that forgiveness is not incompatible with holding others responsible and taking seriously the moral significance of their wrongdoing. In fact, he argues that when done properly, forgiveness

Does not undercut the aim of the responsibility system so much as it presumes it. On an interpersonal level, it seems to me that there is much to be said in favor of forgiveness and the way it relieves the burdens of blame. Moreover, it permits a restoration of the social order. A degree of readiness to forgive might be salubrious for both individual and community.⁶³

⁶⁰ Vargas, p. 243.

⁶¹ Vargas, p. 242.

⁶² Vargas, p. 242.

⁶³ Vargas, p. 242.

As his chapter continues, however, Vargas does not readily endorse this position (forgiveness as a viable alternative to blaming). Throughout the book, he argues that the willingness of individuals to blame is plausibly a prerequisite on collective action⁶⁴, meaning that blame in human life is, to some degree, inevitable. Still, if we focus on his claim that forgiveness “does not undercut the aim of the responsibility system so much as it presumes it”, we can see that forgiveness does not contradict or get in the way of the structures of moral responsibility embedded into our blaming practices. Both blaming and forgiving can be ordinary parts of our moral repertoires in response to wrongdoing. As I will argue in the chapters to come, expressing blame is not an essential aspect of holding others responsible in shared moral community, and Vargas’s view offers strong support for this argument.

Due to blame’s retributive core, which is liable to come out (and forcefully at that) during any blaming episode, however civilized the expressed blame may appear, we have good reason to reduce our daily expressions of blame. Vargas’s view is consistent with that goal insofar as it acknowledges that we can still make sense of our moral responsibility system without expressed blame, and that we need not always publicly blame others even when they are morally responsible, due to any number of context-dependent factors. In this way, I am offering a generalized version of Vargas’s localized worry about blame sometimes being too costly. My view differs from Vargas’s view in that I argue that blame’s retributive root means that it is a liability in our daily moral lives. Further, because expressed blame is a liability in our daily moral lives, we are better off breaking the habit of blame and finding alternative practices through which we can hold each other responsible.

⁶⁴ Vargas, p. 243.

VIII. Conclusion

In sum, I have argued that expressions of blame belong to a family of responses that have retributive origins, and thus expressed blame involves a risk of emotional damage, entrenchment, and escalation. Even though some forms of expressed blame are quite civilized, and even though many of us train with great effort to tame our retributive impulses in our daily moral lives, it remains the case that expressed blame always has the capacity to devolve into its psychologically primitive form. On the one hand, just because retribution is evolutionarily hardwired in us does not mean that it is all that likely to surface in our interpersonal moral lives. On the other hand, when our immediate responses to wrongdoing are affectively hostile and retributive, those harsh emotions strip away the domesticated veneer of civilization in our everyday blaming practices.

In this chapter, I have suggested we have reason to believe we are not as many iterations past the original, primitive forms of blame as we may like to think we are. It still requires a lot of daily effort for many of us to not let our blaming responses devolve into their primitive origins. Suppressing our retributive impulses in the face of wrongdoing is important. However, directing that energy into a civilized guise in the form of McGeer's Domesticated Blame strategy fails to weed out the possibility of those primitive foundations manifesting.

For this reason—deep reasons about the retributive nature of expressed blame—I am skeptical of the utility of civilized forms of blame, because they are not so far removed from the primitive origins they aim to avoid. In arguing this, I recognize what a radical claim it is. I am not only suggesting that there are blaming alternatives, and that we would be well-advised to consider those alternatives to expressed blame. More strongly, I argue that due to the social risks that blame's retributive core invokes, alternatives to blame are preferable to expressed blame. I

have not (yet) laid out specific alternatives to blame—those will come in chapter four. In the next two chapters, I will present further arguments against expressed blame. This will prepare the way for me to propose concrete and practical alternatives to expressed blame in chapter four—alternative modes through which we may hold each other responsible.

Chapter Two. Blame, Retribution, and Anger

I. Introduction

In chapter one, I argued that blame has an essentially retributive core. In this chapter, I build upon that argument and take a closer look at one specific blaming emotion: anger. I suspect that anger is a common feature of many cases of expressed blame, particularly affectively-charged cases of blame. In fact, anger and blame have been called a natural pairing.⁶⁵ For this reason, a critical analysis of anger will help elucidate what is so off-putting and socially harmful about the affective component of expressed blame.

While there are many different forms of anger with varying emotional intensities, the form of anger that is most strongly associated with our expressed blaming practices, according to the Stoic philosopher Seneca, can be characterized as heated and blinding anger. Seneca (among others) discussed this form of anger at length. Generalizing Seneca's point, Owen Flanagan writes "anger almost always overreaches and overreacts; it is incontinent...As soon as anger gets its grip, I can't see things—you, our relationship—as they are. I overreact, and things are spinning out of control, especially if you react angrily to my unjust behavior."⁶⁶

I will focus on heated anger in this chapter, though I recognize that some forms of anger are less heated. For example, siblings (or any other agents in close relationship) can have a moral conflict which results in sustained, decade(s)-long angry feelings towards each other.

⁶⁵ For example, psychologists Malle, Guglielmo, and Monroe (2014) argue that expressions of blame are frequently accompanied by anger, and that the two share many properties. See also Wranik & Scherer (2010); Averill (1983) for further discussion of the connection between anger and blame.

⁶⁶ Flanagan, Owen (2016). *The Geography of Morals: Varieties of Moral Possibility*. Oxford University Press, p. 169-70.

Expressions of anger that involve such a long time-period are inherently less heated, since it seems implausible for anger to remain heated to the same intensity for the entire duration of a decades-long conflict. Over time, our emotions (positive and negative) naturally lessen, or weaken, in intensity. For example, while we may feel ecstatic the day we land our dream job, that positive emotion naturally weakens over time. It would be exhausting to feel *constant* joy or constant anger. Sometimes our expressions of anger are very intense, and sometimes they are more settled. Sometime our angry feelings are less on the surface of our emotional life, and instead are wedged deeper into our psyches, or even entirely repressed. In some ways, deep-seated anger seems *more* emotionally and socially dangerous than immediate, intense forms of anger, because it has a greater chance for lasting collateral damage. We can think of one of the siblings in their decade-long dispute finally exploding after all of the repressed anger they have felt for ten years, and we can easily imagine this explosion causing irreparable harm to their relationship.

Many philosophers have written on anger and the ways it is (or is not) retributive, including recent works by: Flanagan, Martha Nussbaum, Susan Wolf, and Nic Bommarito. This chapter offers a critical survey of each of these four philosophers' works to investigate the ways in which anger (especially heated, Senecan forms of anger), retribution, and expressed blame intersect, and problematically so. In my critical survey, the four authors will be split into two contrast pairs. Nussbaum and Wolf discuss angry blame in particular; Flanagan and Bommarito discuss anger more generally. Nussbaum and Flanagan are strongly skeptical about the helpfulness of anger in our daily moral lives; Wolf and Bommarito propose more moderate approaches in which they tentatively endorse the role of anger in our daily moral lives.

In my discussion of these four authors' works, I will investigate the close connection between anger, blame, and retribution. My investigation into each of their arguments will helpfully reveal this close connection between anger and blame. Further, I will argue that their arguments, even though not all explicitly about angry blame, are highly applicable to the specific context of interpersonal angry blame. Further, I will expand on my argument from chapter one that expressed blame is, at root, retributive by adding the additional argument that anger is an exacerbating feature of blame's retributive nature.

II. Nussbaum

A. Anger and Retribution

According to Nussbaum, 1. Anger is retributive, and 2. Retribution is bad because it is premised either on magical thinking, or on a status-demotion, neither of which is a helpful response in our daily moral lives. Nussbaum strongly argues that we *ought to avoid anger wherever possible*. Nussbaum's focus is on not just manifestations of anger in our moral lives, but on manifestations of anger in our lives more broadly. In her article "Transition Anger", Nussbaum argues that "a close philosophical analysis of the emotion of anger will show that it is normatively irrational: in some cases, based on futile magical thinking, in others, based on defective values."⁶⁷ Indeed, she is pessimistic about anger's capacity to serve any useful function.

Nussbaum largely follows Aristotle's definition of anger, which is that anger is "a desire accompanied by pain for an imagined retribution on account of an imagined slighting inflicted by

⁶⁷ Nussbaum, Martha (2015). "Transitional Anger." *Journal of the American Philosophical Association*, p. 41.

people who have no legitimate reason to slight oneself or one's own."⁶⁸ Breaking this definition down into sub-parts, Aristotle argues that anger involves five elements:

1. Slighting or down-ranking
2. Of the self or people close to the self⁶⁹
3. Wrongfully or inappropriately done
4. Accompanied by pain
5. And linked to a desire for retribution

Concerning the fifth aspect of Aristotle's definition, the desire for retribution, Nussbaum expands Aristotle's account. She suggests that this desire for retribution does not necessarily involve a violent wish for revenge. Anger's retributive arm is not always about wanting the other agent to physically suffer. Rather, retribution can take many forms. Nussbaum writes, "I may not want to get involved in revenge myself; I may want someone else, or the law, or life itself, to do it for me. I just want the (wrong)doer to suffer."⁷⁰ The suffering we wish upon wrongdoers can be psychological or emotional, not just physical or material.

For example, I might hope that a colleague who has wronged me will embarrass herself in front of our boss or that she will be psychologically unhappy. I might wish that her marriage goes badly or that she has a hard time on the academic job market this year. Even if I do not want to directly get involved in causing her suffering myself, I can still wish revenge on her indirectly. Desires for revenge come in a variety of types and degrees. The common thread behind our desires for revenge is the idea that because you, the wrongdoer, harmed me, you deserve to suffer as compensation for that specific harm you caused me.

⁶⁸ Nussbaum, p. 42.

⁶⁹ Nussbaum agrees with Aristotle that anger involves a slighting or down-ranking, but disagrees that it must be of the self or those close to us. She objects that we often get angry even when the slighting happens to a stranger, not just to ourselves or our loved ones. Nussbaum suggests that we get angry about instances of wrongdoing that touch on or threaten our own core values, regardless of whether the instance happens to someone we know or a stranger. These cases often directly affect ourselves or our loved ones, but they can also arise in cases of wrongdoing done by strangers, so long as the wrongs touch on our core values.

⁷⁰ Nussbaum, p. 46.

On my view, retribution is not purely or merely revenge. Rather, it is revenge that also has a *moralized rationale*. If I have retributive impulses towards you, I not only want you to get you back, but I want to get you back *because you harmed me and doing so is morally wrong*. With retributive aims, there is a clear moralizing intent behind my payback. I want to get back at you because you did x to me, and x is morally wrong. I might take revenge on someone for a non-moral wrong they committed, whereas retribution has a narrower scope in its moralized intent.

As I mentioned, Nussbaum finds any form of payback to be ill-conceived. The desire for revenge is ill-conceived because either it is premised on a form of magical thinking, or it involves a hyper-focus on status and an attempt to down-rank the wrongdoer. Neither of these paths serve a positive moral function. By magical thinking, Nussbaum argues that punishing offenders rarely repairs the damage that they did to us. Many of us may seek revenge in the hopes of something like closure or relief, but revenge rarely brings closure. It is simply magical thinking to believe that retribution will lead to anything like closure or relief.

As an example, Nussbaum discusses a TV interview of legendary athlete Michael Jordan after his father's murder. During the interview, Jordan is asked whether, if they ever caught his father's murderer, he would want the murderer executed. Jordan responds, 'Why? That wouldn't bring him back'.⁷¹ Jordan's comments in the interview reflect the type of magical thinking that Nussbaum argues is at the heart of anger's retributive edge. Though retribution may seem to compensate for the harm done to oneself, in reality, the suffering retribution causes will not bring back what was lost. As his statements in the interview make clear, Jordan recognizes this and has

⁷¹ Nussbaum, p. 47.

no desire to get revenge on his father's murderer. He knows it will not bring his father back, and that to think it will bring his father back is a form of magical thinking.

In addition to magical thinking, retribution can also take the form of a status down-ranking. Nussbaum explains

Many societies do encourage people to think of all injuries as essentially about them and their own ranking. Life involves perpetual status-anxiety, and more or less everything that happens to one either arises one's rank or lowers it.⁷²

Nussbaum finds this tendency towards status-anxiety fundamentally narcissistic, explaining

The tendency to see everything that happens as about oneself and one's own rank seems very narcissistic, and ill-suited to a society in which many aspects of human welfare have intrinsic value...The path of status...converts all injuries into problems of relative rank, thus making the world revolve around the desire of vulnerable selves for domination and control. Because this wish is at the heart of infantile narcissism, I think of this as a *narcissistic error*.⁷³

Nussbaum argues that this tendency towards an excessive focus on oneself and one's status prevents a focus on problem-solving or considering the situation as a whole.

Relatedly, as Jean Hampton notes, "if people are secure in their dignity, they won't see an injury as a diminishment; but people are rarely this secure. They secretly fear that the offense has revealed a real lowness or lack of value in themselves, and that putting the offender down will prove that the offender has made a mistake."⁷⁴ Nussbaum expands upon Hampton's claim to argue the more general assertion that because people care a lot about their own public standing, sometimes they express anger as a way to reaffirm their public standing in the face of being hurt. Broadly speaking, the claim is that people who are more secure in their public standing do not see harm done to them as an attack on their own worth. Nussbaum explains that anger may well induce a reversal of positions, "but only because the values involved are distorted: relative status

⁷² Nussbaum, p. 50.

⁷³ Nussbaum, p. 50-51.

⁷⁴ Nussbaum, p. 48.

should not be so important.”⁷⁵ In this way, focusing on status down ranking is not a good reason for retributive intent. Just like magical thinking, it too is based on a fundamentally distorted understanding.

B. Transition-Anger

Given the inadequacy of anger’s two retributive manifestations—as magical thinking or as a status down ranking—Nussbaum concludes that anger is not a helpful response to being wronged. Instead, she proposes a “borderline” species of anger. This form of anger, which she terms *Transition-Anger*, serves as a transition away from retributive anger and towards future-oriented, constructive thinking about how to be useful in the face of wrongdoing.⁷⁶ She urges victims of wrongdoing to focus on promoting future welfare, rather than on pursuing retribution. She calls Transition-Anger a “healthy segue into forward-looking thoughts of welfare and, accordingly, from anger into compassionate hope.”⁷⁷ The Transition mentality, Nussbaum argues, wants justice and brotherhood whereas the payback mentality of anger wants groveling.⁷⁸

According to Nussbaum, “the way anger goes away in the Transition seems much more promising (than retribution): one stops thinking about one’s own inner states and starts thinking about how to do something useful, and perhaps even generous, for others.”⁷⁹ While adopting Transition-Anger might seem challenging for many of us, Nussbaum is optimistic about each of our abilities to adopt Transition-Anger, instead of ordinary anger, in the face of being wronged. She thinks that being sympathetic about the motives of other people helps make Transition-

⁷⁵ Nussbaum, p. 42.

⁷⁶ Nussbaum, p. 42.

⁷⁷ Nussbaum, p. 53.

⁷⁸ Nussbaum, p. 53.

⁷⁹ Nussbaum, Martha (2016). *Anger and Forgiveness*. Oxford University Press, p. 136.

Anger more plausible. Rather than getting angry, Nussbaum argues we ought to aim to be moral problem solvers. Our aim should be to help wrongdoers morally develop. With this aim in mind, Transition-Anger is a useful and plausible strategy. Summarizing her view, Nussbaum explains,

I am saying something very radical: that in a sane and not excessively anxious and status-focused person, anger's idea of retribution or payback is a brief dream or cloud, soon dispelled by saner thoughts of personal and social welfare.... when anger makes sense, it is normatively problematic (focused narrowly on status); when it is normatively reasonable (focused on the injury), it doesn't make good sense, and is normatively problematic in that different way. In a rational person, anger, realizing that, soon laughs at itself and goes away.⁸⁰

Unlike traditional forms of anger, Transition-Anger does not seek revenge. If punishment is required, it is enacted in a helpful and kind spirit. For Nussbaum, it is pointless to dwell in the past. Instead, the future is the only thing we have the possibility of changing, so our focus should be on restoration, communication, forgiveness, and moral understanding.⁸¹

In a related article, "Beyond Anger", Nussbaum summarizes her view with the statement that payback "is a deeply human, but fatally flawed way of making sense of the world."⁸²

Nussbaum is hopeful about our abilities to engage in moral problem-solving without anger or retribution. Instead, with Transition-Anger, the motive to avoid morally bad behavior ought to be connected to "morality's positive goals."⁸³

Likening this motivation to avoid morally bad behaviors to one's relationship with others, she writes "a relationship of trust that depends on fear of anger is not a healthy relationship."⁸⁴ If one avoids bad moral actions purely out of fear of receiving anger or feeling guilt, this motive is

⁸⁰ Nussbaum, p. 30-31.

⁸¹ There is further support for Nussbaum's view in the Buddhist and Stoic traditions. For example, Seneca writes, "Let us be free from this evil, let us clear our minds of it, and extirpate root and branch a passion which grows again wherever the smallest particle of it finds a resting-place. Let us not moderate anger, but get rid of it altogether: what can moderation have to do with an evil habit? (Seneca, *De Ira*, III. 42).

⁸² Nussbaum, Martha C (2016). "Beyond Anger." *Aeon*, 26 July 2016, aeon.co/essays/there-s-no-emotion-we-ought-to-think-harder-about-than-anger.

⁸³ Nussbaum, p. 133.

⁸⁴ Nussbaum, p. 133.

not useful. Instead, one should avoid bad moral behaviors because one is committed to abiding by our moral system and wants to pursue morality's positive goals. Rather than pursue the correct moral actions to avoid a guilty conscience, one should pursue the correct moral actions out of a deep and authentic concern for the welfare of others.

In response, one might worry that while Nussbaum's account is appealing in theory, we humans are hard-wired for both retribution and anger (as evinced by Greene and Cohen's work), and so magical thinking and status-downranking are not elements we can so easily do away with. According to the worry, it might be overly optimistic to think that we can simply overcome our angry tendencies. Nussbaum predicts this sort of objection. In response, she explains that even though anger may have evolutionary roots, this does not imply that anger is an inevitable mode of behavior. She argues

We work hard to correct many tendencies or propensities that are hardwired in human nature, from myopia to memory lapses. As with diet and exercise: we do not have to believe that we will ultimately free ourselves from all illicit cravings in order to embark on a program of self-cultivation. Who knows? Maybe non-anger will make our lives go so much better that we will not even miss the strife-torn days of our past, any more than we always retain an acute craving for French fries and donuts. And even if we continue to experience anger, we need not make public policy based on its misleading normative promptings.⁸⁵

In this passage, Nussbaum claims that we have a choice over whether or not to give into our angry impulses. When the impulse to express anger arises, choosing to express it is optional. This mirrors Greene and Cohen's line of argumentation that even though retributive impulses are likely hardwired into our psyches, we still have choice (to some degree) over how we act on those primitive impulses. Like our inclination towards eating junk food, we have a natural disposition towards getting angry. However, this does not imply that we are *bound* to express

⁸⁵ Nussbaum, p. 248.

anger. As Greene and Cohen explained, we are able, to some degree, to bracket out those retributive impulses in our lives, even though they are hardwired into our psyches. With great effort, we might be able to train ourselves to bracket-out our retributive impulses, or to not habitually act on them when the impulses arise. Just as many of us are able to curb our junk food intake because we value our health, we might be able to tame (to some degree) our retributive impulses because we value non-hostile dialogue. Nussbaum is ultimately hopeful about our ability to decrease our expressions of anger from our daily lives.

C. Nonviolent Communication

Marshall B. Rosenberg, author of *Nonviolent Communication: A Language of Life*, offers further support for Nussbaum's view. Rosenberg agrees with Nussbaum about anger's pernicious effects in our daily interpersonal lives. He argues that anger "can be valuable if we use it as an alarm clock to wake us up—to realize we have a need that is not being met and that we are thinking in a way that makes it unlikely to be met."⁸⁶ However, Rosenberg argues that at its core, anger reveals a need that is not being fulfilled. For example, people often feel angry when they are being (or perceive that they are being) mistreated. People sometimes feel much *better* after the catharsis of expressing anger to the person who mistreated them, as that expression of anger has alerted the mistreater to the angry person's unfulfilled need. For example, when I angrily tell my partner that I resent doing most of the household chores myself, my expression of anger reveals my unfulfilled need to be respected. I feel disrespected by the unequal distribution of chores between us. I might also feel angry at him for other reasons relating to the case, but part of my anger stems from an unfulfilled need in me.

⁸⁶ Rosenberg, Marshall B. (2015). *Nonviolent Communication: A Language of Life*. PuddleDancer Press, p. 144.

Rosenberg suggest that although expressing anger may be a natural way to express our unmet needs to others, expressing anger can also blind us to relevant circumstantial factors, such as potential ways in which the conflict is partially our own fault. Accordingly, it is well worth asking, in cases of interpersonal moral conflict: Is this partially *my* fault? Sometimes we project our anger onto others, and doing so blinds us to important considerations. In an attempt to be right, sometimes we discount our own partial culpability. Though this is not always the case, it is worth considering whether we (the blamer) have partially contributed to the wrong.

One problem with expressed blame, according to Rosenberg, is that it is a *superficial* expression of anger. It is not a direct expression of anger. Rather, it delays the process of expressing one's anger by directing blame onto another person and making our anger their responsibility. In the case with my partner, it might be the case that my expressed blame aims largely to make him feel bad, rather than address the root of the issue, which is my unmet need for respect. This view mirrors Nussbaum's idea that the payback element of anger does not solve the problem. It does not bring back what was lost in the wrongdoing. Sometimes anger successfully makes a lot of noise, without solving a problem. Similarly, Rosenberg argues, expressing blame makes a lot of noise and certainly can direct the wrongdoer to the problem, but it tends to aim primarily to make the wrongdoer feel bad. This aim to make the wrongdoer feel bad can delay the goal of moral problem solving.

Rosenberg's non-violent approach to interpersonal conflict might be too extreme for many readers. It is not always the case that anger largely represents our unfulfilled needs. Surely our anger is *sometimes* very rightly about the other person and their wrongdoing. Nevertheless, I have included his view in this section to demonstrate precedent in the psychological literature for nonviolent alternatives to interpersonal conflict. Further, Rosenberg's account directly supports

Nussbaum's account of Transition-Anger in its renunciation of anger as a useful moral tool. Generally speaking, Nussbaum's view is part of an increasing range of views (mine included) that question the utility of anger, blame, and retribution in our daily moral lives.

D. Is Nussbaum's Case Overstated?

As mentioned above, Nussbaum defines anger in an Aristotelian sense. According to Aristotle, anger involves retribution of some sort. However, this might not be the case when anger arises in response to non-moral cases. It might be the case that Nussbaum's account is overstated insofar as she generalizes all anger as retributive, rather than specifically moral cases of anger. For example, when I am angry that it is snowing again today, in April in New York City, this does not involve a hyper-focus on status or on payback. Additionally, when I accidentally stub my toe on the door, I have no desire to exact revenge on the door, or on myself, for failing to step around it. I am briefly angry because of the pain, but there is nothing moralizing about my pain-induced anger. In each of these non-moral cases, there are no retributive aims.

Perhaps Nussbaum would say her account only applies to *human-originated* causes of anger. Even still, many daily cases of anger arise from human causes that do not lead to status injury or a focus on payback. Accordingly, I propose that Nussbaum's account is too general about the ways in which anger is always retributive. Rather than all anger being retributive, it is the case that *moralizing cases of anger are retributive*. For example, when I am waiting in line at the drug store and the customer ahead of me is taking a very long time to count out her exact change to pay for her items, anger sometimes arises in me. I get impatient. I might feel especially angry if I am running late or am very busy that day. I might feel as if the slow customer is not

valuing the time of the rest of us in line. However, I have never felt the need to get revenge on this customer (or those like her). Likewise, I have never felt that her slowness in counting her change threatens or diminishes my moral status in any way. The slow customer is not doing anything morally wrong. It is completely her prerogative to count out exact change for her items.

As this example demonstrates, Nussbaum is exaggerating the scope of anger's retributive element. Retribution is simply not omnipresent in all cases of moral and non-moral anger. Nussbaum fails to make this distinction between moral and non-moral anger clear. Yet, when anger is mixed with expressed blame, Nussbaum's account is spot on about the problematic ways in which retributive impulses arise. I argue that Nussbaum is right to think that we have a tendency towards retribution in response to being morally wronged. Thus, we can think of Nussbaum's account of anger, in a narrower sense, as a compelling and correct account of *angry blame*.

Nussbaum explains what is both incoherent and morally distasteful about retribution: that it is grounded in either magical thinking or a status demotion. Her persuasive argument about retribution helps make more salient why I find the risk of blame devolving into its primitive, retributive origins so problematic. But there is a deeper worry: when we add anger to the mix, it is an exacerbating feature of blame's retributive nature. Because anger involved with wrongdoing tends to be heated, and because anger in moral contexts involves retributive aims, anger cannot help but exacerbate the underlying retributive core of expressed blame. While Nussbaum might be wrong to think that all anger is retributive, she is quite right to think that angry blame is, at its core, retributive.

I have argued that moral cases of anger are essentially and conceptually retributive, but not all cases of non-moral anger are essentially and conceptually retributive. Similarly, some

forms of expressed blame overtly involve retribution, while others simply contain an incipient risk of devolving into retribution. When combining anger (especially Senecan anger) and expressed blame, it seems obvious that the risk of manifesting the retributive impulse will increase. Moreover, anger tends to restrict our ability to see things clearly. The raging bull is blind with anger, and sometimes in our interpersonal moral lives, we too become blind with anger. Anger can quickly become insufficiently focused, overdone, and excessive. Because anger tends to unravel, or to grow larger, it can make cases of expressed blame more heated and affectively nastier than they need to be. Other negative emotions associated with expressed blame (resentment, disappointment, and sadness) certainly run this risk, too. These emotions seem at least contingently hostile, if not retributive, depending on how they are deployed.

Senecan anger feels unique in its close connection to retribution. The retributive spirit underlying expressed blame, coupled with an outburst of anger, is inherently a social risk in our daily moral lives. This is not to say that angry forms of blame will always devolve into retributive charges. But, I posit that retribution is more likely to occur in instances of angry blame than it is in affectively cooler forms of blame. Nussbaum's account of anger's retributive foundation helps make the social risks involved in expressed blame more vivid. Given that expressed blame is always incipiently retributive, blame is highly likely to be expressed retributively when the forceful heat of Senecan anger is brought into the mix.

E. Fittingness Versus Appropriateness

The fitting versus appropriate distinction I introduced in chapter one is highly relevant to my current discussion of Nussbaum's view. On my view, both angry blame and retribution may well be *fitting* responses to a given case of moral wrongdoing. When others morally wrong us, it

quite rightly fits the situation to feel angry and want to express blame in response. However, as I have been arguing, more often than not, it is not *appropriate* to express blame (even when fitting) in our daily moral lives because the blaming emotions are such unpleasant modes of expression to receive, and they risk creating irreparable social harms. As discussed above, Nussbaum's definition of anger, in which anger is essentially retributive, fails to consider that there are many non-moral cases of anger that do not involve retribution. However, she has made a very plausible argument for angry blame, in particular. For this reason, the fitting versus appropriate distinction applies even more strongly to cases of angry blame than it does to cases of non-angry blame because angry blame is fundamentally retributive, and retribution amplifies the social risks involved in expressed blame.

Recall that Nussbaum thinks that anger is premised on either a form of magical thinking, or on a focus on status. A focus on status (and the anger that arises when one feels that one's status is downranked) certainly has an important place in society. There are times (war, for example) in which anger arising because of one's status being diminished serves the situation in useful ways. To say that retribution is bad in all contexts would be too extreme and is not my view. However, I am arguing that even if retribution may be a *fitting* response to a given moral situation in our daily lives, it is likely not *appropriate*. Barring extreme cases, in our interpersonal moral lives, we should be able to rise above our retributive and angry impulses as often as possible.

III. Wolf

A. Wolf's Account

Unlike Nussbaum's account, Wolf argues that, generally speaking, angry blame works well as a response to moral wrongdoing. In her paper "Blame, Italian Style", Wolf's focus is not simply anger, but angry blame in particular. In the section above, I argued that angry blame is Nussbaum's area of focus as well, even though Nussbaum herself presents cases she describes simply as cases of anger, not angry blame. However, I argued that while not all cases of anger are retributive, moral cases of anger, i.e. cases of *expressed blame*, involve retributive impulses. In this section, I will assume that Nussbaum and Wolf have largely the same focus—angry blame—in their respective accounts, even though they use slightly different terminologies.

Wolf explains that angry blame is a central feature of her family's daily interactions, and one that she cannot imagine them doing without. For her, angry blame is not the hostile, negative force that Nussbaum and I have made it out to be. Instead, Wolf finds angry blame to be a positive and indispensable daily moral tool. Wolf agrees with me that sometimes our expressions of angry blame can become excessive. But, on the whole, she argues that angry blame functions fairly well in our interpersonal lives and is best conceived as an important moral tool.

One reason Wolf claims that angry blame plays such a positive moral force is that it provides a way for us to express our attachment, concern, and care for those close to us. Part of being close to others and creating deep, fulfilling bonds with them involves moral conflict from time to time. Furthermore, Wolf argues that getting angry and blaming one another is a normal and healthy part of our relationships with our friends and family members. Despite its negative affective expression, Wolf suggests that blame projects to our loved ones that we are committed to them, and that we want them to make morally better decisions, because we care about them and their well-being. For example, Wolf argues

Although the angry emotions and attitudes do seem to me to be conceptually tied to a disposition to punish, and therefore with a willingness to make the object of blame suffer

in a particular way, it would be a serious mistake to identify this with a general withdrawal of good will. Even in the midst of my fury at my daughter's repeated raids of my closet, there was never a moment when I wanted harm to come to her...I may have wanted her to experience the painful feelings of guilt and remorse. I never wanted her to break her leg, or even scratch her knee.⁸⁷

The example Wolf references in the quote above is an example of a daily moral conflict in her household: her teenage daughter repeatedly steals her clothes without asking. Wolf responds to her daughter with angry blame, but she does not see this harsh response as harmful or problematic. However, Wolf does agree with me (and Nussbaum) that angry blame fundamentally involve the desire to punish wrongdoers.

While Wolf does not use the term 'retribution' directly in the passage above, she does remark that the "angry emotions and attitudes do seem to me to be conceptually tied to a disposition to punish, and therefore with a willingness to make the object of blame suffer in a particular way."⁸⁸ While a 'disposition to punish' is not equivalent to 'retribution', the two are close cousins. Recall that I defined retribution as *revenge that also has a moralized rationale*. I not only want you to get you back, but I want to get you back because you harmed me and doing so is morally wrong. With retributive aims, there is a clear moralizing intent behind my payback. I want to get back at you because you did x to me, and x is morally wrong. Wolf explains her 'disposition to punish' her daughter to be targeted, writing 'I may have wanted her to experience the painful feelings of guilt and remorse'. This is a form of psychological retributivism because Wolf wants her daughter to suffer in a psychologically damaging way (through guilt or remorse) because her daughter stole her clothes, and Wolf finds this action morally wrong.

⁸⁷ Wolf, Susan (2011). "Blame Italian Style". in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, Wallace, R. Jay, Rahul Kumar and Samuel Freeman (eds.). Oxford: Oxford University Press, pp. 12-13.

⁸⁸ Wolf, p. 12.

Generalizing the point, anger often has an *object*. Anger is usually about something in particular. When the angry emotions involve punishment, that punishment is directed towards a specific agent for a particular reason. I am angry at you and now I want to punish you because you wronged me. However, Wolf does not think that angry blame is conceptually retributive. She merely thinks that angry blame is punitive. Further, Wolf does not perceive blame's punitive nature as problematic. Wolf does not think this desire for wrongdoers to suffer is excessive. On Wolf's view, then, expressing blame to those close to us in our daily moral lives is not as problematic from a social risk standpoint as I have thus far made it out to be. Wolf explains

In some happy families, people may treat each other with unfailing respect and consideration...In some happy families, though, people get angry; they raise voices, they fight, they cry, and then they apologize, they kiss and make up. The spirit of multiculturalism calls on us to celebrate this diversity. I just want to make sure that we have a similarly diverse repertoire of concepts to match.⁸⁹

In the above passage, Wolf helpfully points out that different families, groups, and friends might each handle moral conflicts differently. Some families stay calm and kind to one another in the face of moral conflict, while other families repress moral conflicts by refusing to even acknowledge them. Other families address moral conflicts directly with big outbursts of anger, perhaps even physical fighting. The variation contained within our interpersonal responses to moral wrongdoing is striking. For Wolf, getting angry and expressing blame towards her family members is the organic and unproblematic way her own family dynamic operates. According to her, we do not want to discount this method. Instead, Wolf argues that we ought to embrace the variety of ways we respond to wrongdoing as a testament to the diversity contained within in our relationships. Angry blame is one way among many of responding to moral

⁸⁹ Wolf, p. 26.

wrongdoing. Wolf further explains how her family uses anger and blame on a daily basis to bring them closer together:

(Consider) my daughter's repeated raiding of my closet to borrow clothes and shoes without permission...If you heard the slammed doors and raised voices or saw the dirty looks and tight jaws that accompanied the discussion of these events, you would not think twice about whether to describe these as episodes of blame. But do they indicate that the parties involved hold attitudes that impair their relations with each other?...I have a very close family, with deeply gratifying relationships, which I cannot imagine having anything like their actual character in the absence of episodes like these.⁹⁰

For Wolf, these daily familial conflicts certainly count as instances of blame. However, she maintains that there is no lasting damage done in expressing angry blame when these sorts of conflicts arise. Expressing blame is a fundamental way her family communicates and resolves conflicts. She simply cannot imagine her interactions with her family without the tool of blame.

Wolf concludes that "Getting angry and expressing it, and demanding a response, may bring people together and make them closer, rather than pushing them away. In part this is because such behavior encourages apology and remorse more than other shifts of attitude that reflect an impairment in relationships."⁹¹ R. Jay Wallace echoes this point, arguing that "a failure to experience (blaming emotions) when e.g. someone has wronged us would tend to indicate the absence of the sort of emotional investment in moral values that we generally take to be desirable."⁹² Wallace argues that a lack of blaming emotions implies a lack of emotional investment. Though this is not exactly Wolf's point, both Wallace and Wolf argue that the blaming emotions are highly useful indicators of our care for those close to us and of our interest in their moral growth. In this way, Wallace's and Wolf's views stand in sharp contrast to Nussbaum's view insofar as they see angry blame as a positive, healthy way to express and

⁹⁰ Wolf, p. 4.

⁹¹ Wolf, p. 13-14.

⁹² Wallace, R. J. (2011). "Internalism about Responsibility". Unpublished manuscript, p. 12.

resolve moral conflict, whereas Nussbaum (and I) see angry blame as a socially risky and retributive strategy.

B. Objections to Wolf

In light of Wolf's argument, I must examine whether I have exaggerated the extent to which angry blame serves a negative role in our interpersonal moral lives. I have two objections to Wolf's view. First, it might well be true that we can feel an overarching or generic sense of good will towards someone while still expressing angry blame towards them for a minor wrongdoing. However, it is hard to see how we can express angry blame towards someone without withdrawing our good will towards them, at least temporarily. When Wolf angrily blames her daughter for raiding her closet, Wolf certainly has an overriding, generic sense of good will and care towards her daughter. She might have a generic sense of good will towards her daughter no matter what events transpire between them, due to the inextricably close bond and relationship that they share.

That generic sense of good will is not impacted by Wolf's angrily blame about the specific instance of wrongdoing, her daughter raiding her closet. So, saying that Wolf can maintain a generic sense of good will towards her daughter while still angrily blaming her seems irrelevant to any hostile feelings or lack of good will exhibited in the exact moment of the wrongdoing. Because on the micro-level, when Wolf angrily blames her daughter, how can she not be withdrawing her good will, at least temporarily? A temporary withdrawal of good will towards another person is compatible with a generic or overall sense of good will towards another person. It would be an extreme view to think that the two—a temporary withdrawal of goodwill and a generic sense of good will—are incompatible. Furthermore, the fact that we can

retain generic feelings of respect and good will towards a wrongdoer does not mean our micro-scale withdrawal of good will is not deeply harmful. Wolf needs to say more about how the forcefulness of angry blame does not involve any withdrawal of good will in the present moment.

Indeed, many philosophers have written about the temporary sting or forcefulness involved in our expressions of blame. Summarizing this sort of view, Pamela Hieronymi writes, “the force of blame is located in the harms, burdens, demands, and expectations contained in or resulting from those blaming actions and attitudes that go beyond the simple belief that ill will or disregard was shown.”⁹³ Further, in “The Force and Fairness of Blame” (2004), Hieronymi argues that while there is a characteristic forcefulness to blame, that forcefulness is not unfair or unnecessarily harmful. Hieronymi argues for the general conclusion that “a characteristic force of a judgment of ill will cannot render one unfair in making that judgment; it could at most generate obligations of judgment management.”⁹⁴

Hieronymi further argues that the forcefulness (or lack thereof) of expressed blame “lies in the interpersonal significance of a judgment of it.”⁹⁵ Hieronymi means that the forcefulness of expressed blame depends on the context-sensitive dynamics between the agents and the ways in which they interpret the expressed judgment of blame. Expressed blame is not automatically or necessarily forceful, but it becomes forceful based on the interpersonal significance the respective agents ascribe to it. As is clear in Wolf’s account, she and her family do not ascribe excessive significance to their interpersonal expressions of angry blame towards each other. However, this is dependent on context-sensitive factors about the dynamics of their specific

⁹³ Hieronymi, Pamela (2004). “The Force and Fairness of Blame”. *Philosophical Perspectives*, 18 (1), p. 117.

⁹⁴ Hieronymi, p. 132.

⁹⁵ Hieronymi, p. 132.

family structure. That same type of angry blame might feel very retributive and harmful to other family or friendship dynamics.

Even if Hieronymi is right, there is a wide range of interpersonal cases and dynamics in which blame's forcefulness is harmfully active. I am not claiming that every single case of interpersonal moral wrongdoing involves actively forceful or hostile affective elements. Instead, I argue (in line with a common account of blame) that expressed blame involves a characteristic forcefulness, or sting.

IV. Flanagan

A. Flanagan Against Anger

In his recent book, *The Geography of Morals*, Flanagan draws upon work in anthropology, empirical moral psychology, behavioral economics, and philosophy to explore the role of anger in our current Western moral system. In line with Nussbaum's argument, Flanagan worries that anger is not the positive force that many Westerners think it to be. In order to be specific about what he is arguing against, Flanagan has created a taxonomy of anger, which I have summarized below:

1. *Anger* or *angry feelings* refers to the phenomenal state(s) that is/are experienced as anger, the phenomenal state(s) that feels/is angry. It includes whatever psychophysical states the angry person experiences, reddening, heat, and the impulse to strike out.
2. *Angry behavior* is any behavior in the world that results from anger, for example, strong words, criticism, gossip, shaming, striking.
3. *Anger norms*: permissions or recommendations about appropriate anger, both phenomenal, how angry one is allowed or supposed to feel in different circumstances, and behavioral ones, what it is legitimate to do given those feelings.
4. *Anger scripts*: socioculturally and normatively specific scripts for anger and angry behavior.
5. *Justified anger* is anger that is at least permitted by anger norms, normally the justification involves there being considerations that warrant or speak in favor of anger.

6. *Righteous anger* is anger, the phenomenal state and the psychophysical accompaniments, which is warranted or considered warranted by high values, especially justice violations.
7. *Righteous indignation* refers to the judgment that there has been disrespectful or insulting treatment of someone, others or me. The judgment involves feeling, the feeling that disrespect, insult, or injustice has occurred. It is not angry.
8. *Punishment, violence, war*. These are all practices that may or may not be motivated by anger, as are rudeness, sarcasm, snark, irony, and passive aggressiveness.
9. *Annoyance, frustration, resentment, disdain, and contempt*. These are all attitudes that seem to be, or are likely to be, in the anger family. They have negative valence, are heated, directed at an insult, obstacle, or impending threat, and typically come with a desire to harm the source, to kick the flat tire, tell you that you are racist slime.⁹⁶

I will first consider this comprehensive list in the specific context of angry blame. Each of these distinct types of anger can arise in a given case of moral wrongdoing, but most of the time one or two are dominant in a given situation. For example, imagine that I am about to have surgery and my sister has agreed to pick me up and drive me home afterwards. When my sister neglects to pick me up from my surgery, after she promised to do so, I am angry at her. She failed to live up to a promise she made to me, and as a result, I feel stranded and vulnerable while on painkillers without support. When I express this anger to my sister, I feel my anger is justified (type 5 in Flanagan's taxonomy). I might also feel annoyance and frustration (type 9) at her, and very likely also angry feelings (type 1), but in this case justified anger is dominant for me.

Consider another example. A colleague tells me she did not vote in the last election because she wanted to go home and watch *The Bachelorette* rather than stand in line for a few minutes in order to vote. In response, I might feel righteous anger (type 6). I might think that she made a morally bad decision, beyond just being lazy. She failed to live up to a norm of citizenship, and I am righteously upset. This is different from what Flanagan terms righteous *indignation* (type 7) in that indignation is when people, rather than our justice system in general, have been

⁹⁶ Flanagan, p. 160-1.

mistreated. I might express righteous indignation when my neighbor steals the FedEx package outside my doorstep for her own enjoyment before I get home to grab it. In this case, the wrong affected me specifically (rather than U.S. citizens as a group, in the voting case) and so my anger is expressed as indignation rather than as righteous.

Flanagan's taxonomy of anger is helpful to my own argument for a few reasons. For one, it reminds us of the variety of forms anger can take in our daily lives. Anger is not a singular or unchanging thing. As I mentioned at the start of this chapter, I am focused on heated, Senecan forms of anger. However, I am aware of the variety of forms of anger involved in our daily moral lives, and Flanagan's list helps make this variety more specific. Second, Flanagan's list narrows in on what is so bad about angry blame. He explicitly defines the attitudinal and behavioral accompaniments of angry states. In my focus on interpersonal blame, the types of anger that will most likely come up in our blaming practices are types 1 and 2 (anger and angry behavior), as well as types 7 and 9 (indignation and annoyance, frustration, resentment, disdain, and contempt). In what follows, I will continue to use the umbrella term *angry blame* and will be more specific if I am talking about other types of anger.

B. Eliminativist Views About Anger

Flanagan discusses the Stoic view of anger. In many ways, the Stoic mirrors Nussbaum's argument against traditional forms of anger. Seneca offers an eliminativist view about anger, arguing that anger is something best avoided at all costs. In "De Ira", Seneca argues that anger is not something we can or should moderate. Instead, Seneca thinks that we must banish dangerous emotions like anger entirely. He argues,

Let us be free from this evil, let us clear our minds of it, and extirpate root and branch a passion which grows again wherever the smallest particle of it finds a resting-place. Let

us not moderate anger, but get rid of it altogether: what can moderation have to do with an evil habit?⁹⁷

Because Seneca views anger as entirely evil, he thinks that there is no point in pursuing a more moderate path. This extreme view of anger includes expressions of anger that seem appropriate or fitting for the situation. Even if expressing anger seems justified for a given case of wrongdoing, *still*, one should not express it. Seneca's reasoning is that anger can, all too quickly, spin out of control.

Flanagan compares Seneca's extreme, eliminativist view to a Buddhist eliminativist view. For example, Santideva, an eighth-century master who wrote a book entitled "The Bodhisattva's Way of Life", writes about anger being the *worst* state of mind a moral agent can have. Buddhist arguments against anger involve a metaphysical world view that, due to limited space, I will not discuss at this time.⁹⁸ But, many Buddhists generally agree, like Seneca, that *anger is a poison* best avoided entirely. For example, Santideva writes

Those who wish to cause me suffering
Are like Buddhas bestowing waves of blessings
As they open the door for my not going to an
unfortunate realm,
Why should I be angry with them?⁹⁹

This passage reflects the attitude that even those who cause us suffering should not receive our anger. On this view, even when others harm us in significant ways, still, we should not respond with anger. This is in stark contrast to a typical Western approach to anger. Further, it seems very hard to enact, practically speaking.

⁹⁷ Seneca, Annaeus (1900). "Minor Dialogs". Translated by Aubrey Stewart. Bohn's Classical Library Edition (London): George Bell and Sons, p 140.

⁹⁸ For more information, see Flanagan (2016) pp. 163-165 for a brief overview.

⁹⁹ Flanagan, p. 165.

Strikingly, contained within Santideva's passage is the idea that when others harm us, it can provide us an opportunity to practice not getting angry. We can even be grateful to this person for the chance to practice responding to others without anger. Santideva's attitude is wholly eliminativist regarding anger, just like Seneca's view. With these two eliminativist views in mind, it is worthwhile to consider whether such an eliminativist position is overstated. According to eliminativists, anger's tendency to overreact provides a strong enough reason to eliminate it from our moral lives altogether. However, as I will now argue, a modified version of the eliminativist view is more realistic. My moderate eliminativism draws inspiration from the Stoic and Buddhist views, as well as Nussbaum's Transitional-Anger account, and is in stark contrast with Wolf's defense of angry blame, yet it allows some space for anger in our daily moral lives.

C. Modified Eliminativism

In the rest of this chapter, I will argue that a moderate eliminativist position is the best approach to anger. While we should not try to eliminate all anger from our daily moral lives, since it can occasionally serve morally useful purposes, we should be aware of the destructive after-effects of anger, and only express anger when both fitting and appropriate. A modified eliminativist view of expressing anger involves caution about (but not full elimination of) anger, because of anger's innate tendency to overact and spiral out of control, especially in moral contexts, i.e. angry blame.

There are several reasons why a modified eliminativism is attractive. For one, even if expressing angry blame functions in some positive ways, the excessive, spiraling nature of angry blame is an ever-present social risk. As an obvious example, a blamer's anger can inspire the

wrongdoer to express anger back. We see this happen every day. When a car cuts in front of us on the highway, many of us respond in kind with some degree of road rage. We manifest this by speeding up and cutting back in front of them, and perhaps also shouting obscenities out the window. We do this to get even. The first driver's aggressive driving spurred similarly aggressive driving in the other driver. This is clearly unsafe. The two drivers might continue to egg each other on, ending up in an accident.

The same type of exchange often occurs in interpersonal moral conflicts, too. When someone gets angry at us, it is entirely unremarkable for us to get angry back, perhaps even amplifying their anger out of defensiveness. When my friend steals my idea for a new article and writes and publishes it herself before I have a chance to, I might respond by getting angry at her. In turn, she might get angry back. Then, mere moments later, we engage in an escalated, aggressive shouting match. The problem with this dynamic, as Seneca tells us, is that anger almost always overreaches and overreacts. It is natural to match the emotional tone of the person blaming us. The blamer is riled up, so we get riled up, too.

Relatedly, the blamer's anger can encourage defensiveness from the wrongdoer. Being on the receiving end of such an aggressive emotional state can naturally lead to a defensive response. As demonstrated in Nussbaum's argument against anger, hostility is an unhelpful response to moral wrongdoing because it does not move either agent towards productive dialogue and moral problem solving. Instead, defensiveness causes one (or both) agents to further dig in their heels. Thus, defensiveness is not only a pernicious emotion to be on the receiving end of, but it also leads to a further moral problem: lack of motivation to make positive moral changes.

At the same time, angry blame can also function in positive ways. Wolf demonstrated that angry blame can, for some families and interpersonal dynamics, function productively. Sometimes, anger might instead serve as a useful signal, rather than as a spiraling outburst. In some cases, the angry blamer is signaling that she refuses to be treated this way, that the wrongdoer's behavior is unacceptable. It is possible that the wrongdoer will react to this by wanting to correct her own behavior. She might take the anger as a useful motivational force to get her to act differently next time. When angry blame serves as a signal, this seems useful. However, I maintain that all too often, the emotional charge of the anger breeds further escalated anger and defensiveness, instead of motivation. If we can achieve those same signaling and deterring aims without expressing anger, this seems, on the whole, less risky and thus preferable.

Contra the eliminativist position, I do not argue that we must eradicate angry blame. That may be impossible. We will, inevitably, sometimes angrily blame others and this is to be expected. In fact, there are some cases where anger is socially and morally useful. For example, when extreme and urgent injustice occurs in our daily moral lives, anger is sometimes the best response.

Relatedly, philosopher Myisha Cherry defends the role of moral anger in her article "Moral Anger, Motivation, and Productivity". Cherry's subject is *moral anger*, which she describes as *anger arising out of a moral wrong*. This is just how I have described angry blame throughout this chapter, so I take it that we are focused on the same subject, and the difference in word choice is negligible. Cherry defends the productive role moral anger plays in our moral conflicts, remarking that "Moral anger is productive when anger leads one to work towards pursuing justice in ways that are moral and focused on reaching the goal of justice."¹⁰⁰ On

¹⁰⁰ Cherry, Myisha (2012). "Moral Anger, Motivation, and Productivity". Presented at New Jersey Regional Philosophical Association Conference, Bergen Community College.

Cherry's view, moral anger can be a motivating force for agents to want to make moral progress. She views anger as a positive and productive catalyst for change.¹⁰¹

Cherry is right to think that *sometimes* moral anger can serve a positive function in working towards justice, despite its explosive tendencies. Even though anger can quickly escalate and become destructive, it does play a positive role in some of our moral practices. A moderate eliminativist position allows for this positive role anger plays, especially in the face of grave injustice, while still recognizing the destructive after-effects of anger. In sum, I propose that even if anger is *fitting* for a given situation, all things considered it is rarely *appropriate*, given its destructive tendencies.

V. Bommarito

A. Bommarito's Account

Finally, I will present a fourth account of anger, which serves as a contrast case to Flanagan's account and as support for my own case for the moderate eliminativist position. Bommarito argues that anger "is morally virtuous when it is a manifestation of concern for moral goods."¹⁰² This does not imply that anger is necessary for being virtuous. Instead, anger *can become morally virtuous or vicious based on the underlying concerns it manifests*. For Bommarito, anger can be virtuous or vicious, depending on the context. Bommarito's argument is compatible with a moderate eliminativist position because he only finds anger to be morally virtuous when it manifests a concern for moral goods. In this way, he presents a cautiously (and

¹⁰¹ For further discussion on the role of moral anger, see also works by Alison Jaggar and Elizabeth Spelman.

¹⁰² Bommarito, Nic (2017). "Virtuous and Vicious Anger". *Journal of Ethics and Social Philosophy* 11:3, p. 1.

limited in scope) positive account of anger.¹⁰³ Similarly, Bommarito thinks that anger *can be* retributive or non-retributive, depending on the circumstances. Bommarito explains,

Richard is angry that the Red Sox have lost the World Series because he cares about the team doing well. Julia is angry that she must work on Christmas because she cares about celebrating the holiday with her family. By contrast, someone who does not care about American sports or Christian holidays is unlikely to be angry in similar situations because these things just do not matter to her.¹⁰⁴

Bommarito is not saying that caring about others requires getting angry. Rather, he is saying that getting angry requires caring about *something*. When that care is about something morally good, one's anger is virtuous. He also holds that the inverse is true: "When anger manifests a lack of concern or immoral concern such as malice, contempt, or ill will toward them, then it is morally vicious."¹⁰⁵

Bommarito defends anger on the grounds that it is a mistake to assume reason involves being in control of our emotions and that anger rob us of that control.¹⁰⁶ This runs counter to Seneca's argument that anger prevents us from seeing things clearly, and for that reason should be banished from our daily lives. Differently, Bommarito suggests that anger need not be thought of something we fall prey to, or something that robs us of thinking clearly. Instead, anger can, at times, be something that inspires us, aids us, helps us be "more in control, less distracted, and less tempted."¹⁰⁷

Bommarito is right to think that anger can helpfully clarify what moral goods we care about. For example, I am hardly ever upset when a stranger breaks a promise to me, but I am almost guaranteed to be angry when my closest friend breaks a promise to me. In our

¹⁰³ Bommarito does not take a clear stance on whether anger is retributive. His account of anger is more context-dependent than Nussbaum's or Flanagan's.

¹⁰⁴ Bommarito, p. 12.

¹⁰⁵ Bommarito, p. 12.

¹⁰⁶ Bommarito, p. 10.

¹⁰⁷ Bommarito, p. 10.

interpersonal lives, we tend to expect *more* from those close to us, and are understandably let down when they do not live up to those expectations. In this sort of situation, anger can be a way I clearly manifest what I care about, i.e. our friendship. I get angry at my friend (and much more so than I would at a stranger) because I *value* our friendship, and feel that my friend's betrayal is not in keeping with the established bond of trust between us. I value the quality of trust within our friendship, and my anger arises in response to feeling that my friend has broken that moral norm and expectation of trust between us. Anger is the fuel for me to express what I care about. In this respect, Bommarito is surely right to say that *sometimes* anger makes more vivid the moral goods that we care about.

I will now present two arguments against Bommarito's account to demonstrate that while he is right to think that anger can be fitting when it is a recognition of moral goods, anger is in fact rarely appropriate. This argument will make further use of the fitting versus appropriate distinction I raised in chapter one. Further, this argument will support my moderate eliminativist position insofar as I grant that anger can sometimes serve a useful function in our daily moral lives, but that it is rarely appropriate.

B. First Objection: Restraint

First, Bommarito's account of anger ignores an important condition of *restraint*. Anger that manifests concern for moral goods through yelling, kicking, stomping off, rage, and any other aggressive attitudes and gestures is not very virtuous. Even if the intent behind such behavior is to manifest concern for moral goods, *the delivery of the message matters a great deal*.

It is not enough to deem any expression of anger that manifests concern for moral goods as virtuous. The presentation (tone, body language, facial expression, volume, etc.) of the anger matters. In this way, Bommarito's account is overly simplistic. What is missing from his account that anger can manifest concern for moral goods is the condition that the expression of that anger needs to be delivered in a morally appropriate way. I will not attempt to defend a full account of what that sort of restraint on delivery might look like, but I will offer some initial thoughts.

As Nussbaum persuasively demonstrated, an expression of anger should be delivered with an eye towards being *helpful* to the wrongdoer, rather than adding more emotional or psychological damage to her. Good intentions are not enough. Angry agents also need to be aware of the ways in which they communicate their anger, and *how that anger is received* by a particular audience. Shouting one's anger and throwing things around is unlikely to be morally helpful. In addition, expressions of anger might be more effective if not immediately delivered to wrongdoers. In the immediate aftermath of wrongdoing, tensions are likely to be high on both sides.

For example, imagine that I see someone cut in front of someone else in a line we are all waiting in at the grocery store. I express my anger over what I perceive to be a moral wrong by cutting in front of the person who had originally cut line. This qualifies as a virtuous response on Bommarito's account because my anger manifests my concern for the moral good of waiting one's turn. I am demonstrating my commitment to that moral good through my angry actions. However, the way I manifest my anger in this case is not virtuous, because I am doing the very thing that the wrongdoer did, cutting in front of others in line. My actions are hypocritical.

Bommarito might reply that there is an implied hypocrisy condition in his account, restricting my angry actions in this example from being virtuous on account of them being

hypocritical. Or, he might accept that my anger in this case did manifest concern for a moral good, and in this way, is virtuous, even though it is hypocritical. Either way, I propose that there is a big difference between anger in theory, and anger in our real, everyday interpersonal lives. Extreme sensitivity to the context-sensitive factors involved in our expressions of anger is crucial. Attention to this difference impacts our judgments of anger's virtue, and Bommarito's account ought to reflect this context-sensitivity. There are many ways to manifest concern for moral goods, and some ways are better—from a social risk perspective—than others. Thus, while anger can sometimes play a positive role in our moral lives, there are additional context-sensitive factors involved in determining anger's role, and these additional factors suggest that anger might not be appropriate, even when it manifests a concern for moral goods.

C. Second Objection: Anger Has Too Much Heat

Additionally, even if anger's delivery manifests clear concern for the wrongdoer, anger still may not be useful to express simply because it has *too much heat*. Even if angry forms of expressed blame result in the best short-term moral outcome (i.e. they successfully motivate wrongdoers to change in the immediate aftermath of the wrongdoing), perhaps this strategy is harmful for the long-term dynamic between wrongdoer and blamer. Stirring up all that heat and blinding intensity involved in Senecan varieties of anger is not worth the short-term, cathartic benefit of expressing heated blame. Again, the issue is largely with heated and blinding Senecan forms of anger, and less with cooler forms of anger.

Some commentators agree with me. As writer Rebecca Solnit argues in her article "Facing the Furies", "Much political rhetoric suggests that without anger there is no powerful engagement, that anger is a sort of gasoline that runs the engine of social change. But sometimes

gasoline just makes things explode.”¹⁰⁸ Contra the traditional view that heated forms of anger demonstrate one’s engagement in a given case of wrongdoing, Solnit explores how even if anger initially motivates a call for moral change, it might not create actual change. Instead, it might create too much heat. Like gasoline, a little bit is needed to fuel a car, but too much creates an explosion. This implies that sometimes anger is not effective in achieving the aim of actually creating lasting change. Implicit in Solnit’s remark is the idea that there are better, more accurately directed, engines of social (or moral) change beyond anger. If anger tends to make things explode, perhaps an alternative strategy is wise.

In a recent empirical study, Jennifer Lerner and Dacher Keltner found that feeling angry makes people as optimistic about the outcome of a situation as feeling happy would. In other words, “anger may make people miserable, but it also makes them more confident and obliterates other, more introspective miseries: pain, fear, guilt, uncertainty, vulnerability. We would rather be mad than sad.”¹⁰⁹ As this finding demonstrates, there is a real and compelling allure to anger. When we are wronged, anger masks other painful emotions like fear, guilt, or uncertainty that we might not want to feel or be ready to feel. Feeling angry when we are wronged can, at least in the moment, *feel good*. Rather than introspecting about why we feel those ways, sometimes we lash out and express anger instead. It can be easier and more rewarding in the short-term to express anger than to feel the fear, guilt, and uncertainty. But again, the problem with angry blame is how quickly it can escalate and spiral out of control, and how hard it can be to regulate it once it has begun spiraling.

¹⁰⁸ Solnit, Rebecca (2017). “Facing the Furies.” *Harper’s Magazine*, 1 May 2017. <https://harpers.org/archive/2017/05/facing-the-furies/4/>.

¹⁰⁹ Solnit, p. 4.

Returning to Solnit's gasoline example, it is clear that there is a very fine line between anger serving as productive moral fuel, and anger 'exploding', making it hostile to understanding.¹¹⁰ Additionally, knowing when we are expressing too much heat seems to require profound self-awareness. Caught up in the heat of the moment, this can be hard. Angry blame serves the purpose of clearly signaling that wrongdoing has occurred. But, there may be clearer and more effective signals. For example, angry blame does not convey to wrongdoers *how* to make moral changes, or *what* resources are available.

This is not to suggest that we should try to completely get rid of all angry impulses. It would be naïve to presume we have all that much control over our impulses. Instead, we have the power to harness our anger, and to fully feel it, but not need to express it as angry blame. One can fully feel the emotion of anger without inflicting it onto others. Indeed, we can recognize anger arising in us as an internal signal that something needs to change. Perhaps one can transform that anger into productive, solution-oriented energy.¹¹¹ Further, feeling obligated to get angry, and express it via angry blame, in the face of *every* daily instance of moral wrongdoing we encounter would be exhausting. None of us actually practice this, nor would it be advised to do so. Knowing that our moral emotions have a limited bandwidth, we can effort to be more focused in the cases we choose to get worked up over.

This insight provides support for the view that even if anger is fitting for a given situation, it still may well be inappropriate to express. While Cherry and Bommarito each provide positive arguments for anger, they have both drastically underestimated the potential destructive after-effects of anger. Rather than view anger as appropriate if it manifests a concern

¹¹⁰ Solnit, p. 4.

¹¹¹ See Emily McCrae (2015) for a detailed look at the Tantric Buddhist practice of transforming anger, a strategy with many practical applications.

for moral goods (Bommarito), or when it serves the goal of justice (Cherry), I propose that angry blame is rarely appropriate, given its destructive social after-effects.

Thus, a moderate form of eliminativism argues that even if angry blame serves some useful functions, it also comes with many troubling elements, elements which should, on the whole, dissuade us from expressing angry blame, for worry of it exacerbating expressed blame's already pernicious sting. Thus, in this section, I demonstrated that a more moderate version of eliminativist theories about anger (and by extension, angry blame) is more a plausible and promising path than pure eliminativism.

VI. Conclusion

In my discussion of these four authors' works, I have presented a critical discussion of several competing views on the role of anger in our daily moral lives. Nussbaum's and Wolf's accounts served as contrast cases about angry blame, while Flanagan's and Bommarito's accounts served as contrast cases about generic anger.

Nussbaum's account helped establish the case for regarding angry blame as retributive. Some readers might think that retributivism is a wholly appropriate response to being morally wronged, and thus would argue that the fact that blame is at least incipiently retributive is not a problem for our interpersonal moral lives. Some might even argue that blame's retributive nature is a good thing. But, Nussbaum's argument convincingly demonstrates the ways in which retribution does not solve our moral problems. Retributive modes of response may make one *feel* as if they are gaining closure after having been wronged. But in reality, retributive modes of

response do not offer productive moral resolution because they are premised on magical thinking or status down ranking (or both).

In addition, the Stoics' eliminativist view about anger offered an extreme ideal for our daily moral lives, entirely devoid of anger. Differently, I am proposing a more realistic, Stoic-inspired account involving a move away from angry blame due to its incipient social risks, rather than a complete elimination of angry blame from our daily moral lives.

My discussion of these various accounts of anger has led me to the following conclusion: anger's heat tends to exacerbate blame's retributive foundation. Angry blame, thus, seems to me to be the worst kind of expressed blame, and best avoided at all costs. Whereas civilized forms of blame seem at least *a few steps removed* from blame's primitive, retributive origin, angry blame seems much closer to blame's retributive origin. We can extrapolate from this that angry blame more easily devolves into retribution than more civilized, affectively 'cooler' forms of blame do, because angry blame is closer to the uncivilized, retributive root.

The aim of this chapter has been to highlight the worst-case scenario in our daily blaming lives: the toxic combination of anger, blame, and retribution. Even though not all forms of blame are this prone to devolve, being aware of the most unstable forms of blame should inform our understanding of and reliance on the practice of expressing blame as a whole. Thus, my new argument from this chapter is that anger exacerbates blame's nascent retributive impulses, and this finding lends itself to my overarching argument that given the social risks involved in expressed blame, it may fit a situation, yet is not appropriate to express. Given our proclivity towards expressing blame coupled with anger in our daily moral lives, we have another strong, *prima facie* reason to break the habit of expressing blaming so readily and automatically in our interpersonal moral lives.

Chapter Three. A Further Argument Against Expressed Blame: Epistemic Skepticism

The previous two chapters discussed constitutive elements of blame itself, whose nature provided strong prima facie reasons for refraining from expressed blame. This chapter approaches the topic from a different angle, epistemic reasons to refrain from expressed blame. In this chapter, I will build my case by drawing on an account of epistemic skepticism from Gideon Rosen (2002, 2004). Rosen is skeptical about our warrant to judge the blameworthiness of others. I will extend Rosen's skepticism regarding culpability in general to a skepticism regarding the moral status of wrongdoers. I will then argue that Rosen has exaggerated the scope of our epistemic skepticism. He is wrong to think we are never warranted to make positive judgments of culpability. Rosen's skepticism is too global. We are sometimes right to make positive judgments of culpability against wrongdoers. Yet, I will argue that even when we are warranted to make positive judgments of culpability, we might not be warranted to express that blame. The warrant to make positive judgments of culpability is distinct from the warrant to publicly express that blame.

In chapter one I argued that blame has retributive origins. In chapter two I argued that retribution is a key ingredient in angry forms of blame and that it is socially risky. Worse, I argued that anger exacerbates blame's retributive root. In this chapter, I shall add to those two arguments against blame a third line of argument that because of the high stakes of public blame and the epistemic uncertainty involved in our judgments of blame, expressions of blame are best avoided.

I. Rosen's Epistemic Skepticism

A. Rosen's Account

First, I will review the details of Rosen's account. Rosen is skeptical about moral responsibility altogether. In his paper "Culpability and Ignorance", Rosen considers two types of ignorance, factual and moral. An example of factual ignorance is when someone innocently trespasses on private property. Say this person is out on a hike on an unmarked patch of land and she winds up on someone else's property. Ordinarily, trespassers are liable to be blamed. But, Rosen counters, if one's trespassing occurs due to factual ignorance, one is not blameworthy for trespassing. In this case, there are no signs on the property telling the trespasser that she has overstepped her bounds. Without a clear sign, she has no way of knowing that she is trespassing. As a result, she inadvertently strays onto someone else's property without their permission.

Rosen argues that it would be wrong to blame the trespasser for this transgression since it was an innocent instance of factual ignorance.¹¹² Because there were no signs, she was not to know better. We cannot expect her to have acted any differently. On the other hand, when someone drives down a crowded street in Times Square furiously texting on her phone with her face glued to the screen and she runs into another pedestrian, she should have known better. She is ignorant of the practical risks involved in texting while driving, especially in a crowded metro area. This case, unlike the hiking case, is an instance of culpable ignorance because the Times Square texter was not ignorant about the facts of driving etiquette in a crowded metro area. We all know that texting while driving, especially in a crowded metro area, is risky. The Times Square texter knows her distracted driving is risky, but she does it anyways. The texter

¹¹² Rosen, Gideon (2002). "Culpability and Ignorance." *Proceedings of the Aristotelian Society*, vol. 103: 1, pp. 63-64.

intentionally engages in dangerous behavior, whereas the hiker does not intentionally engage in illegal behavior.

Then there are cases of moral ignorance. Rosen cites ancient Hittite slavery as an example of moral ignorance that is non-culpable. Imagine that an ancient slaveholder truly believes that it is morally permissible for him to buy and sell slaves and then force them to work. Rosen claims that because ancient slaveholding was so prevalent and not seen as a moral issue at the time, it was not a culpable moral wrong. He writes, “We may condemn the act. We may rail at the universe or at history for serving up injustice on so vast a scale. But in my view it makes no sense to hold this injustice against the perpetrator when it would have taken a miracle of moral vision for him to have seen the moral case for acting differently.”¹¹³

When we try to apply our current moral attitudes about slavery to this case of ancient slavery, of course we want to say that ancient slaveholding was terribly wrong. It was, certainly, wrong. Despite this correct intuition, Rosen maintains that ancient slaveholders were non-culpable for their moral ignorance. As he explains, ancient slaveholders lacked the correct moral concepts—or perhaps the capacity to employ the correct moral concepts—to see this practice as a moral wrong. Ancient slaveholders were acting in accordance with the prevalent moral attitudes of their time. Given the prevalent pro-slavery attitudes of the time, slave owners *were not to know* better. They were acting in line with their peers. Differently, if one were to engage in slaveholding today, one would be culpable, since at this point in history, one really must know better. It would be impossible to be ignorant about the moral impermissibility of slaveholding today. However, the prevalent moral attitudes about slavery today are vastly different than they

¹¹³ Rosen, p. 66.

were in ancient times. Thus, Rosen concludes, ancient slaveholders' moral ignorance is non-culpable.

In both the factual and moral cases of ignorance, Rosen's account relies heavily on our ability to determine whether a wrongdoer's *ignorance* is blameworthy or not. For Rosen, the non-culpability of the wrongdoer's ignorance determines the culpability of the resultant act.

Rosen expands these examples to the general claim:

If X does A from ignorance, then X is culpable for the act only if X is culpable for the ignorance from which X acts.¹¹⁴

When wrongdoers act from ignorance, they are only culpable if they are culpable for their ignorance. This is not to say that there will not be any *consequences* to an agent's actions done from ignorance. Rosen's view is not that extreme. Agents who commit crimes done from ignorance might well have to face punishment—social, legal, and so on. But, the issue of punishment is separable from the issue of judgments of culpability. A wrongdoer might be punished, by law, for a given action. But, if her morally wrong action was done from non-culpable ignorance, Rosen deems her non-culpable in a strictly moral sense. He further explains:

When I am passive with respect to an occurrence—when it merely happens in me or to me or around me—then I am responsible for the occurrence only if it is the (foreseeable) upshot of prior culpable activity on my part. I can be responsible for falling asleep at the wheel. I can be responsible for forgetting your name. But I am responsible for such things only if they result from some prior culpable act or omission on my part—taking sleeping pills before the long drive, for example, or neglecting to say your name over and over to myself when we were first introduced.¹¹⁵

This is a very striking passage, and one worth examining in greater detail. Rosen suggests that if a driver falls asleep at the wheel, she is responsible for doing so only if it is a result of a prior culpable act, such as taking sleeping pills. If she *knowingly* took sleeping pills and then got

¹¹⁴ Rosen, Gideon (2004). "Skepticism about Moral Responsibility". *Philosophical Perspectives*, vol. 8:1, p. 301.

¹¹⁵ Rosen, p. 303.

into her car, she is culpable because she knowingly risked falling asleep at the wheel and endangering other drivers on the road. However, if she falls asleep at the wheel without any prior culpable act—it just happens—then she is not culpable. This line of thinking runs counter to traditional intuitions about cases like this, where an asleep driver is likely culpable, barring any medical emergencies.

Yet, it is not clear what Rosen means in the first lines of the passage in which he writes “when it merely happens in me or to me or around me”. The statement is ambiguous. What does it look like for an occurrence to happen *in* me? For example, when anger arises in me and I express that anger, am I responsible for any uproar and harm that results from my expression of anger? What is the difference between me being angry and anger “merely happen(ing) in me”? Additionally, there is further ambiguity in what it means for an occurrence to happen *around* me. At what point am I centrally involved in an occurrence that happens around me, rather than passively involved in an occurrence that happens around me? Relating this ambiguity to Rosen’s case of the asleep driver, it is unclear what external factors happen *in* her versus what factors happen *around* her. If it gets very dark on the road and this causes the driver to feel drowsy, eventually falling asleep, is she culpable for falling asleep behind the wheel, on Rosen’s view? Likely not. However, intuitively, most people would describe the act of falling asleep behind the wheel as culpable, since the driver failed to pull over and stop driving as soon as she recognized she was getting drowsy. She did not take a preventative action—pulling over—and is culpable for this negligence.

Additional discussion from Rosen of what he means by ‘in me’ and ‘around me’ in this context is lacking from his text. Whichever way we choose to interpret those phrases, though, his argument remains bold. The standard interpretation of this case would be that the driver is

culpable for falling asleep, no matter why it happened. As I mentioned above, there may be some leniency if her falling asleep was beyond her control, such as in the case of a medical emergency. But, if she knowingly got into the car when she was very tired, or if she knew she tended to fall asleep on long drives, but kept driving anyways, I gather that most folks would consider this an act of negligence and deem her culpable.

Differently, Rosen argues that if there is no prior ignorance on the driver's part, then the wrong action that results—falling asleep at the wheel—is non-culpable. Rosen adds that

Unless you have some further exculpatory information, you should suspend judgment... You should embrace a version of what I shall call skepticism about moral responsibility. You should hold that confident positive judgments of responsibility are never justified.¹¹⁶

This is a radical claim. Rosen is not merely saying that we should make sure someone is responsible before making a positive exculpatory judgment about them. Instead, he is saying we can rarely know whether or not someone is culpable because we can almost never be sufficiently warranted in our judgment. The reasoning is that there are too many potentially relevant external factors at play in a given case of wrongdoing. We often cannot be sure in our judgments of culpability, so Rosen suggests suspending judgment altogether. For Rosen, most confident positive judgments of responsibility are unwarranted based on “current contingent and possibly temporary limits on our knowledge.”¹¹⁷ This means that Rosen is open to us one day having a better grip on how to effectively determine others' culpability. But given our current cognitive limits, judgments of culpability are largely unwarranted. In the next section, I shall respond to Rosen's radical view and present my own version of epistemic skepticism regarding expressed blame in particular.

¹¹⁶ Rosen, p. 295.

¹¹⁷ Rosen, p. 295.

B. Is Rosen's Account Overstated?

On the face of it, Rosen's account seems extreme. If we adopt his model, hardly anyone would ever be culpable. An obvious worry is that we would let many culpable wrongdoers go without moral address by following a Rosen-like methodology. I have two main responses to Rosen's view. The first response is that his skepticism is *too global*. We may well be warranted to make positive judgments of culpability. However, this warrant is distinct from warrant to express blame. We should treat these two warrants (private judgments and public expressions) in a differential way. Second, when it comes to determining whether one has warrant to express blame, I will argue that the *stakes of the case* and the *context* are crucially important, and Rosen's account is insufficiently sensitive to these important factors. While I agree with many aspects of Rosen's account, I will instead propose a more moderate skepticism than Rosen's, and one that takes stakes and context into account.

i. Against Wholesale Skepticism

First, Rosen's skepticism is too global. There are some cases of wrongdoing in which we can be quite sure who is culpable. Contra Rosen, we are sometimes warranted in judging another person to have engaged in clear-eyed akratic wrongdoing. Rosen argues that it is very difficult to identify genuine akrasia because "it is not readily distinguishable from an impostor: ordinary weakness of will."¹¹⁸ However, Rosen is overstating how little we can know about someone's culpability. It is too simplistic to think we lack warrant to judge others culpable, full stop. Sometimes we know with definitive certainty *who* the wrongdoer is, *that they are culpable*, and *for what* crime.

¹¹⁸ Rosen, p. 309.

Some cases of moral wrongdoing are clearer than others, and Rosen's account fails to acknowledge this fact. For example, imagine that my friend Alex punches our mutual friend John completely out of the blue. When I ask Alex why he did it, he replies, "I just felt like it. I wanted to hurt him." I have given Alex a chance to explain himself, but his explanation does not offer any mitigating circumstances or explanations. In this case, Alex is culpable. He has even admitted as much! His wrongdoing is clearly distinguishable from ordinary weakness of will, as his own testimony confirms. It is not that Alex tried not to punch John but simply could not help himself (weakness of will). Rather, it is that he chose, in a clear-eyed way, to punch John.

Therefore, Rosen is wrong to argue especially in cases like these, that we lack warrant to judge another person to have engaged in akratic wrongdoing. Of course, in our daily moral lives, cases of wrongdoing are rarely as simple as the case of Alex and John that I have just presented. Nevertheless, occasionally cases of interpersonal moral wrongdoing are this simple, and Rosen's view does not allow for these sorts of straightforward cases. Sometimes wrongdoers fully admit to their wrongdoing, and no mitigating circumstances can be found. As such, Rosen is wrong to embrace wholesale skepticism regarding culpability. The scope of his skepticism is too wide; there are some cases of wrongdoing in which we do have warrant to judge an agent blameworthy because the epistemic reasons behind the wrongdoer's actions are clear.

ii. Stakes

Second, even if we really do have the proper warrant to blame someone, there is still reason to be skeptical of whether or not to *express* blame to them. There is not necessarily an exact correlation between the warrant to judge someone culpable and the warrant to express blame publicly in light of that judgement. This distinction is important, and builds on the

arguments I have made in the first two chapters about the heightened social and practical risks involved in expressing blame, due to its retributive root and the ways in which anger exacerbates blame's retributive tendencies.

A similar epistemic argument can be made based on this concept of stakes. Even if an agent has warrant to privately judge another culpable, she might not have warrant to publicly express that blame. Expanding on Rosen's discussion of warrant, I argue that expressing blame publicly is a *higher stakes scenario* than is a private judgment of blame. Public expressions of blame involve all kinds of social factors that private judgments of blame lack. For one, public expressions of blame are said *out loud*. The risk of getting it wrong is patently worse if it is said out loud to another agent. Two, public expressions of blame risk being interpreted in a variety of (uncharitable) ways by their recipients. When we make private judgments of blame, we are our only audience, and so we always interpret our judgment of blame correctly.

We can categorize the social risks involved in public expressions of blame along (at least) three categories. First, there is the ethical risk: one's expressed blame may hurt the wrongdoer, lead them to rebel, or lead them to amplify their bad behavior. Second, there is a practical risk: one's expressing blame publicly might backfire and make the moral situation worse in various ways. Third, there is the obvious and practical epistemic risk: it is worse to be wrong if one has said it out loud to another person. Taking these risks together, the compound risk is quite high, certainly much higher than the negligible risk involved in keeping one's judgment of blame to oneself. As such, as I will now argue, the warrant we have to make private judgments of blame is different from the warrant we have to make public expressions of blame.

One reason for this is that public expressions of blame may not be interpreted by others correctly. This asymmetry in interpretation or reception between private judgments and public

expressions of blame effects our warrant to express blame. While I do not have room to go in depth into the nuances of these social factors, I will discuss what I have in mind by way of an example. In order to make this case, I will also draw upon literature on contextualism by Jason Stanley (among others). While Stanley's account is largely a semantic thesis, I will demonstrate that his account is highly relevant to Rosen's epistemic account. Further, my creative application of Stanley's account in this epistemic framework will highlight the context-sensitive conditions we ought to place on Rosen's account in order to make it seem more plausible.

First, a primer on Stanley's contextualism. Stanley explains that "contextualism is the semantic thesis that knowledge ascriptions, instances of '*x* knows that *p*', are context-sensitive in a distinctively epistemological way."¹¹⁹ Context-sensitive sentences express different propositions depending on their different usages. Contextualism claims that our knowledge ascriptions are similarly context-sensitive. A general idea of the view is that, "*S* 'knows' that *p* depends partly on something in the context of 'the attributor', and hence the view is often called *attributor contextualism*. Because an utterance is context-dependent, so too is whether the knowledge attribution is true context-dependent."¹²⁰ The crucial contextual feature in any given scenario is the attributor's *practical stake* in the situation. In the case above, the crucial contextual feature would be *S*'s practical stake in the truth of *p*. *S*'s practical stake in the truth value of *p* effects *S*'s knowledge of *p*.

Stanley's view is not, on the face of it, equipped to deal with, nor be applied to, my view about expressed blame. He is interested in what 'know' really means and his work carves out space to investigate this within the semantic realm. However, I propose a creative application of

¹¹⁹ Stanley, Jason (2005). *Knowledge and Practical Interests*. Oxford University Press, p. 19.

¹²⁰ <https://plato.stanford.edu/entries/contextualism-epistemology/#Bib>.

his semantic account of knowledge as applied to the stakes involved in expressed blame. As I will argue, Stanley's view is relevant to and supportive of my view.

Consider the following example. Say that Kayla hears second hand—from a mutual friend—that her assistant, Arjuna, has been secretly stealing money from her for months. Kayla is a very successful business woman. Beyond this second-hand information, Kayla had no inkling that Arjuna was stealing from her at all. Kayla fully trusted Arjuna in all areas of their working relationship, and has done so for several years. After hearing this information, Kayla certainly *suspects* that Arjuna is stealing from her. But, does she *know* that Arjuna is stealing from her? Kayla has no first-hand evidence. Her knowledge is limited to an external source of information. Moreover, the stakes for this case are high. If Kayla accuses Arjuna of stealing from her—whether or not it is true that Arjuna really is stealing from her—such an accusation will likely alter, if not end, their professional relationship going forward, since trust is such an essential aspect of a close working relationship.

Compare the stakes involved in Kayla's knowing that Arjuna has been stealing from her with the stakes involved in *my* knowing that Arjuna has been stealing from Kayla. I am merely a third-party observer. I do not know either agent well, and have no personal stakes in the situation. Given that Kayla is a public figure, the story has been in the news. When I hear about the case, I might interpret the facts in such a way as to conclude that Arjuna must have stolen from Kayla. In this case, my judgment of blame is private. I am not publicly declaring that Arjuna stole from Kayla, I am just thinking it to myself. I have no relevant party to tell about my judgment, since I do not know any of the agents involved. The stakes involved in my private, third-party judgment of blame are much lower than the stakes involved in Kayla publicly expressing blame to Arjuna. This is because no one will be hurt by my private judgment of

blame about Arjuna. In fact, no one will even know this is how I feel unless I verbalize it. The stakes involved in my private judgment are nearly zero. Differently, public expressions of blame have social consequences, and sometimes those consequences are quite high.

For example, Kayla's public blame of Arjuna could lead to both emotional and legal ramifications for Arjuna, especially since Kayla is in the public eye. There could be real, damaging consequences of Kayla's decision to express blame to Arjuna. Because of this, the stakes involved in Kayla's knowing are much higher than the stakes involved in my knowing. Before Kayla publicly accuses Arjuna of stealing from her, she will likely want to gain more sources of knowledge about the situation. Her private judgment (based on the second-hand information) that Arjuna has stolen from her might well be warranted, but the public expression of that judgment might not be warranted, since it could lead to a very negative consequence for Arjuna. As this case demonstrates, a public expression of blame has far higher stakes than a private judgement of blame, on the assumption that expressing blame publicly is more likely to lead to negative social consequences. In the case at hand, the negative social consequence is other people blaming Arjuna too, or treating him negatively, perhaps even punishing him for an action he may not have done. Therefore, a higher degree of justification in public expressions of blame than in private judgments of blame is called for.

Expanding this example out into a wider argument, the general point this case reveals is that the stakes involved in what counts as knowing depend both on *who* is doing the knowing and *how* that knowing is expressed (publicly or privately). Public expressions of blame are, generally speaking, higher stakes scenarios than private judgments of blame. This may seem quite obvious, since more people are involved in public expressions of blame than in private

conclusions of blame, but this distinction is seemingly absent from discussions of expressed blame in the literature, including from Rosen's view.

When the high stakes of a case are salient, we are best served not ascribing certainty about the case, unless we are fully certain, without a shadow of a doubt, because it is too socially risky. We should be nearly certain about our ascriptions of knowledge when the stakes of the moral wrongdoing are high. In public expressions of blame, the stakes are automatically higher than in private expressions of blame. Therefore, when the stakes are salient, we have to rethink the centrality expressed blame plays in our daily moral lives. This should remind readers of the fitting versus appropriate distinction I raised in chapters one and two. Because of the high stakes of public expressions of blame, I argue that even if blame is fitting for a given moral conflict, it might not be appropriate to express.

In sum, I am arguing:

Premise: Public blame is a higher stakes scenario than private blame

Conclusion: Therefore public blame requires greater warrant than private blame.

Rosen is equally worried about judgments of blameworthiness and expressions of blame. We can infer that he sees the two types of blaming practices as on a par, and further that he is a global skeptic about our warrant to blame (publicly or privately) in general. But Rosen's stance of treating the two on a par stops him from recognizing the different levels of warrant required in the two cases. Contra Rosen, I argue that some people really are blameworthy. In these cases, we sometimes can and should privately judge them blameworthy. However, this does not entail that we should express that blame. I argue that we ought to exert extreme caution over making these judgments of blame public. By publicly blaming too readily, we risk ignoring the specific and relevant context-dependent elements and stakes involved in cases of blame, which effect someone's warrant to make public expressions of blame.

Thus, Stanley's pragmatic contextualism about epistemic warrant is highly relevant to the view I am advancing about the different levels of warrant required respectively for judgements and communications of blame, even though on first pass you might not think so. I have offered a novel interpretation of his semantic account as it relates to Rosen's epistemic argument more broadly, as well as to my own narrower focus on expressed blame. I argued that the social stakes of a case of moral wrongdoing matter a great deal. Generally speaking, cases of public blame involve higher stakes than cases of private judgments of blame. There are further important distinctions to be made within the general set of cases of public blame. Some cases of public blame have higher stakes than other cases of public blame, due to any number of contextually-sensitive factors and interpersonal dynamics involved. Attention to this extreme context-sensitivity with regards to expressed blame helps highlight the ways in which most all expressions of public blame involve higher stakes, and this insight should inform how readily we express blame in our daily moral lives.

II. Harman's Reply to Rosen

I will now turn to another facet of Rosen's view, which will further clarify the differences between my view and Rosen's view. Rosen's view has garnered several objections. Among them is Harman's response in her paper "Does Moral Ignorance Exculpate?". Harman rejects Rosen's claim that positive judgments of culpability are unwarranted. On this, she and I agree. We can, she argues, often make positive judgments of culpability, and we can do so with certainty. However, Harman's account is otherwise a big departure from my own view. As I will explain,

Harman's view is just as radical as Rosen's view, though their views are radical in different ways.

According to Harman, even if someone's false moral beliefs stem from ignorance, they are still culpable for acting wrongly.¹²¹ Harman argues "We are morally obligated to believe the moral truths relevant to our actions (and thus not to believe false moral claims relevant to our actions), and we are often blameworthy for failing to meet these moral obligations, even if we have not been guilty of mismanagement of our beliefs, and even if our ignorance is not motivated."¹²²

By this, Harman means that we each have a moral duty to believe true moral claims relevant to our actions. Failing to meet this moral duty is culpable, even if the failure arises without an agent ever having mismanaged her beliefs. Even if an agent's failure to act in the morally correct way stems from ignorance, she can still be culpable if she has failed to believe the moral truths relevant to her actions. For Harman, this requirement trumps any excusing conditions, such as moral or factual ignorance.

This is a very strong claim, one that is extreme and implausible. Harman's view requires that we are pretty much always morally responsible for our wrong behaviors, despite any relevant mitigating circumstances. On her view, there are very few cases in which exempting conditions apply. For Harman, even if we have not mismanaged our beliefs, we are morally obligated to believe and act upon correct moral truths, and not doing so is culpable.

Harman is unrealistic to propose that we can reasonably expect the average moral agent to root out her false beliefs at all times. This seems impossibly demanding. In her proposal, Harman is essentially expecting each of us to be moral saints for whom constant diligence and a

¹²¹ Harman, Elizabeth. (2011) "Does Moral Ignorance Exculpate?" *Ratio*, vol. 24, p. 464.

¹²² Harman, p. 459.

scrupulous investigation of our beliefs is the norm. The epistemic demands of such a daily practice are too great for an average moral agent, myself included. This is not to say that we should place *no* epistemic demands on the agents in our shared moral community. That would be extreme on the opposite end of the spectrum. Instead, our epistemic demands should be reasonable such that they can plausibly be enacted by the average moral agents among us in our shared community.

The following example, which Harman discusses in her article, sharpens the significant differences between Rosen's view and Harman's view.

Example: A doctor forgets to double check a patient's chart before giving her a blood transfusion. As a result, he accidentally gives her a blood transfusion of the wrong blood type. He ordered type A blood for the transfusion but the patient has type B blood. Obviously, this mistake has severe repercussions for the patient. Although double-checking patients' charts is a standard medical procedure, the doctor forgets this one time.

On Rosen's view, the doctor's ignorance that the patient has type B blood means she is not culpable for her wrong transfusion order. She might be blameworthy for her failure to check the chart if this was indicative of some prior culpable behavior, like a routine failure to double-check a patient's chart. But if this case of forgetfulness was a one-off instance, she is not blameworthy.¹²³

Differently, Harman argues that the doctor is culpable because she does not have false beliefs. The doctor knows better, yet she fails to act in line with her correct knowledge. She does not put her true moral beliefs about how to properly prepare for a patient's blood transfusion into practice. She might have the correct moral beliefs about double checking charts as an important practice, but she fails to act in accord with this correct belief. Even if this is a one-off instance of forgetfulness, on Harman's view the doctor is still culpable. Of course, in my discussion of this

¹²³ See Harman, p. 447 for further discussion.

case, I am only considering moral responsibility, not legal or professional responsibility.

To emphasize the contrast between her view and Rosen's view, Harman brings up one of Rosen's own examples.

Example: Imagine a traditional, 1950's American father who has a son and a daughter. This father intentionally chooses not to save money to send his daughter to college, but does intentionally save money to send his son to college. This father does not see a need for his daughter to go to college, so he does not put aside any money for her. He is sexist in his treatment of his children.¹²⁴

If we apply Rosen's line of argumentation to this case, the sexist father is likely non-culpable.

For one, the father does not believe his actions are wrong. Second, he might not be conscious of how his decision to save for college for his son but not his daughter is sexist. If this is the case, his ignorance about this decision may be non-culpable. Third, even if the father is aware that his decision is sexist, he does not believe the action is morally wrong. He does not appear to have mismanaged his beliefs. The father is applying his beliefs correctly to his financial decision about saving for his son to go to college, but not saving for his daughter to go to college. His actions align correctly with his own beliefs (sexist as they may be).

So, on Rosen's view, we can predict that the sexist man is non-culpable because he is non-culpable for the beliefs from which his actions stem. Differently, Harman argues that this is implausible. The sexist father *must* be culpable. For her, even if one manages one's beliefs correctly, one can still be blameworthy for one's morally wrong actions. Recall Harman's strong argument that

We are morally obligated to believe the moral truths relevant to our actions (and thus not to believe false moral claims relevant to our actions), and we are often blameworthy for failing to meet these moral obligations, even if we have not been guilty of mismanagement of our beliefs, and even if our ignorance is not motivated.¹²⁵

¹²⁴ Harman, p. 457.

¹²⁵ Harman, p. 459.

As this passage reveals, underlying Harman's view is the claim that we are morally obligated to both believe and live up to our shared moral norms in our actions in every instance. Whereas Rosen's view gives a pass to anyone who acts wrongly but does so without any culpable mismanagement of their beliefs, Harman's view does not grant a pass to anyone who fails to meet our shared moral obligations.

Still, Harman needs to explain *why* agents are blameworthy for failing to live up to our shared obligations. On the surface, her account is a normative claim about how people should act. Harman expands on this point, remarking that

Beliefs (and failures to believe) are blameworthy if they involve inadequately caring about what is morally significant. Believing a certain kind of behavior is wrong on the basis of a certain consideration is a way of caring about that consideration.¹²⁶

As this passage demonstrates, Harman argues that wrongdoers' beliefs are blameworthy if they involve inadequately caring about the morally salient features of a given situation. However, Harman under-describes what the morally salient features of a given case of wrongdoing are. For one, it is not clear whether all agents involved in the wrongdoing need to agree about what the morally salient features of the situation are.

Returning to Rosen's ancient slaveholding example, recall that Harman claims that enslaving someone is blameworthy in virtue of the fact that the slave owner *knowingly* enslaves someone. Harman argues that the slave owner fails to care about the morally significant feature of the case, the moral wrongness of knowingly enslaving someone. This seems impossibly demanding of the average ancient moral agent given the prevailing cultural attitudes at the time. Recall that Rosen exculpates the slave owner on the grounds that the slave owner did not know that what he was doing was wrong, since his actions were in keeping with his peers' actions at

¹²⁶ Harman, p. 460.

the time. Rosen writes, “In my view it makes no sense to hold this injustice against the perpetrator when it would have taken a miracle of moral vision for him to have seen the moral case for acting differently.”¹²⁷ Harman thinks knowingly enslaving someone is blameworthy whether or not one believes it is wrong. For her, the only morally salient feature of the case is a clear distinction between what is morally permissible (not enslaving others) and what is not morally permissible (enslaving others). On Harman’s view, whether or not someone is *ignorant* about what is morally permissible is irrelevant.

Accordingly, it seems Harman would simply reject Rosen’s considerations about ancient slaveholders not being in the epistemic position to know better. She might say that Rosen’s consideration of the prevailing moral attitudes of the time is not a morally salient feature of the case. But, this is a normative claim without much support. I argue that both Rosen’s and Harman’s views are too extreme. Rosen’s view is too lenient, while Harman’s view is too demanding with respect to our epistemic limitations. My proposed view stops short of Harman’s excessively strong requirement that we are “morally obligated to believe the moral truths relevant to our action”¹²⁸, regardless of our epistemic limitations. Similarly, my proposed view stops short of Rosen’s excessively lenient suggestion that we nearly always lack the warrant to judge others blameworthy. A moderate view in between these two extremes best accounts for the fact that sometimes wrongdoers really are blameworthy and that we really do have the warrant to judge them blameworthy, while at the same time, we need not be so constantly and scrupulously judging each other’s moral characters.

George Sher’s epistemic account will help clarify my moderate view. In his book, *Who Knew?*, Sher rejects a common philosophical argument, the Searchlight View. After my

¹²⁷ Rosen, p. 66.

¹²⁸ Harman, p. 459.

discussion of Sher's view, I will propose that a moderated version of Rosen's skeptical view allows us to utilize the insights of all three of these philosophers' arguments into an account that remains skeptical about the utility of expressed blame, but nevertheless recognizes that agents do sometimes have real warrant to judge others blameworthy.

III. Sher on the Searchlight View

In his book *Who Knew?* Sher presents and then rejects what he deems a standard philosophical view on blame, the Searchlight View. There is a widely held and unspoken assumption in the literature in favor of a view akin to the Searchlight View. The Searchlight View proposes that we are only responsible for things that are within our awareness. Think of the searchlight a car emits. When it is very dark outside, a driver can only see those things that the car's searchlight reveals to her. Similarly, in our interpersonal moral lives, we can only be responsible for things that are on our radars, i.e. on our inner searchlights. Sher himself rejects the Searchlight View. In this way, Sher's diagnostic account supports Harman's view about the arduous requirements of our moral epistemic responsibilities.

As Sarah K. Paul (2017) notes, the Searchlight View is widely assumed in the philosophical literature, "though not often explicitly defended."¹²⁹ According to the Searchlight View,

An agent's responsibility extends only as far as his awareness of what he is doing. He is responsible only for those acts he consciously chooses to perform, only for those omissions he consciously chooses to allow, and only for those outcomes he consciously chooses to bring about.¹³⁰

¹²⁹ Paul, Sarah K. (2017). "Good intentions and the road to hell", *Philosophical Explorations*, pp. 44.

¹³⁰ Sher, George (2009). *Who Knew? Responsibility without Awareness*. New York: Oxford University press, p. 4.

This means that one is responsible for a given action only if one is aware of what one is doing. An agent's responsibility tracks their choices. Paul further notes the similarities between the Searchlight view and what Thomas Nagel (1979) has labeled the Control Condition on moral responsibility: that one should not be morally assessed for what is due to factors beyond one's control.¹³¹

Sher argues that the Searchlight View is a popular but wholly inadequate way of assessing responsibility. The Searchlight View suggests that responsibility is limited solely to acts (or omissions) that an agent chooses. The view does not dole out responsibility to agents who are not aware of what they are doing. As an alternative, Sher proposes a more complex account of responsibility. On his view, even agents who unwittingly act wrongly or foolishly can be responsible for a given wrong if 1. They have information that supports the conclusion that their acts are wrong or foolish, and 2. They fail to draw that conclusion based on prior information, and 3. The failure is caused by the constellation of psychological and/or physical features that make them the persons they are.¹³² On first glance, this view sounds quite similar to Harman's view. Harman claims that one can be culpable even when one does not consciously choose an action. Sher similarly argues that one can be culpable even when one does not consciously choose a wrong action.

In his paper "Out of Control", Sher makes the seemingly obvious, but important, descriptive claim that we often hold agents responsible for their actions "even though considerations unrelated to determinism strongly suggest that they cannot help performing them."¹³³ To illustrate the point, Sher provides the following example:

¹³¹ Paul, p. 45.

¹³² Sher, p.1.

¹³³ Sher, George (2006). "Out of Control". *Ethics*: 116 (2), p. 285.

Caught off Guard—Wren is on guard duty in a combat zone. There is real danger, but the night is quiet. Lulled by the sound of the wind in the leaves, Wren has twice caught herself dozing and shaken herself awake. The third time she does not catch herself. She falls into a deep slumber, leaving the compound unguarded.¹³⁴

Sher argues that Wren is definitely culpable and may be liable to punishment, even though when she falls asleep “she does not set her duty aside but ceases to be aware of it.”¹³⁵ It does not seem to be a problem of Wren’s *will* so much as it is a problem that something external (sleepiness) that overtakes her will. Sher assesses whether Wren’s sleepiness is something Wren has the capacity to control. If she does have the capacity to control it, but fails to control it, then it is a culpable case of negligence. Yet, if Wren does not have the capacity to control her sleepiness and it is not something she could have reasonably foreseen, then she is non-culpable.

This case bares strong resemblance to Harman’s case of the doctor and her failure to check the chart. In Sher’s case, the issue is about negligence—is Wren’s sleepiness something she ought to have predicted and have known well enough to avoid? In Harman’s case, the issue is about forgetfulness—is the doctor’s one-off forgetfulness something she ought to have predicted and have known well enough to avoid? The standard reading of both cases would be that both Wren and the doctor are culpable. Sher draws a general conclusion:

Agents are at best obligated to prevent the development of habits and traits that are markedly worse than normal.¹³⁶

In the case of Wren, Sher argues that there is no reason to think that Wren is any more irresponsible than the average person. Further, Sher notes that “we draw a blank when we look for wrongfully forgone opportunities for self-improvement in which to locate their responsibility for their later wrong acts.”¹³⁷ This is troubling for Sher because it appears to conflict with the

¹³⁴ Sher, p. 286.

¹³⁵ Sher, p. 286.

¹³⁶ Sher, p. 294.

¹³⁷ Sher, p. 294.

Searchlight View's assumption that a lack of control excuses an agent from responsibility. As I indicated, the Searchlight View has long been the established and default philosophical position on culpability. But, Sher maintains, there is a real tension between our intuition that Wren is culpable and the fact that she lacked control over her sleepiness.

Sher concludes, contra the Searchlight View, that agents can be responsible for acts whose wrongness they are unaware of. There is nothing Wren fails to realize or know as she is dozing off. She simply lacks control to not fall asleep. Still, on Sher's view, Wren is blameworthy even though she was unable to stop herself from dozing off. This is the same conclusion that Harman would reach, based on her conclusions about the doctor. In this way, Harman and Sher (in similar but not identical ways) reject the Searchlight View in favor of a more stringent approaches to judgments of culpability. For both authors, even if a given act is out of one's conscious control, one can still be culpable. Now that I have presented Rosen, Harman, and Sher's views, I will revisit the doctor case to further argue for a moderated version of Rosen's skepticism.

IV. Revisiting the Doctor Case: A Moderated Skepticism

The doctor case involves the doctor's failure to check a patient's chart. Without knowing all the particulars about the case, I can reasonably imagine that the doctor had to work very hard and pass many tests in order to get where she is. She surely spent many years in medical school studying and learning to prove her qualifications for her role as a doctor. Forgetfulness is not a quality that gets you very far as a doctor.

On one level, we would not ascribe any false beliefs to the doctor. She is not culpable, as Harman describes the case, for having a false belief. She had the correct moral beliefs, but she

made a mistake by failing to act on her correct moral beliefs. Intuitively, this seems less serious a wrong than having actual false beliefs about proper medical practice. The doctor likely knew better (through her years of medical training) than to not double check the chart. She had the right beliefs, she just failed to act on them. This is a different kind of error than having entirely wrong beliefs.

While the doctor should hardly be off the hook for giving her patient the wrong type of blood transfusion, this mistake is not necessarily culpable. We cannot be sure whether her forgetfulness was partially due to any number of external, mitigating factors. Indeed, this is precisely Rosen's view. Rosen and I agree on this aspect. Without knowing more specific details about the situation, Rosen's deep skepticism applies insofar as both private judgements of blame and public expressions of blame would be unwarranted.

This is not to say the doctor is not *causally* responsible for the mistake of switching up the blood types. Clearly, she is the one who mixed them up. Rather, I argue that Rosen's account gives us good reason to be skeptical about our warrant to judge the doctor culpable with any degree of certainty. If we lack sufficient warrant to deem the doctor's forgetfulness culpable, we lack sufficient warrant to express blame to her. Harman would likely respond by saying that even if the doctor's forgetfulness was a one-off case of forgetfulness, *still*, it cannot go unnoticed. Whether she was aware of her forgetfulness or not is irrelevant. Moreover, as Sher argues, we can be culpable for acts whose wrongness we are unaware. There was a clear failure on the doctor's part, and that failure is culpable.

Harman and Sher's points about the doctor and Wren are instructive. Intuitively, I agree with both of them that these agents are culpable, even though their omissions do not necessarily arise from a conscious choice. So, they are right to think that a Rosen-type argument is too

passive in the face of a real, and life-threatening case of wrongdoing. However, Harman and Sher's line of thought is that there may be some underlying beliefs, desires, or attitudes behind these wrong acts that are culpable. On this point, Rosen's skepticism is instructive. Without more information, we cannot know which beliefs, desires, and attitudes were involved behind the scenes in the doctor's negligent actions. As such, we (from an external perspective) do not have the warrant to declare, with certainty, that those underlying beliefs, desires, and attitudes can be traced to a culpable source. More importantly, because of the higher stakes involved in expressions of blame, especially public expressions of blame in places like a large hospital setting, we lack sufficient warrant to express blame to the doctor. We must practically assess the social risks of what publicly expressing blame to the doctor might look like, and what repercussions it could have.

There are many social risks in this case; the stakes are fairly high. Say that the doctor's forgetfulness does not result in any harm done to the patient having the blood transfusion. Despite this good and lucky outcome, say that the doctor gets publicly blamed in front of the patient's family. The patient's family learns that the doctor failed to check the chart before the blood transfusion, thereby endangering their daughter's life. The family could get unnecessarily worked up, angry, sad, etc. They might sue the doctor (or the hospital) for negligence. They might refuse to pay the bill for their daughter's treatment. They may tell their friends what a terrible hospital it is and suggest that others boycott the hospital. These outcomes are unnecessary and risky. Even if there is good reason to judge the doctor blameworthy, publicly expressing might result in unnecessary emotional harm for the patient's family. This, again, highlights the distinction in stakes between judgements of blame and public expressions of blame.

That being said, we ought not remain silent when wrongdoing occurs. This is where an extreme, Rosen-type view is insufficient. My moderated version of his skepticism attends to context-sensitive factors. For example, based on the context and persons involved, it might make sense to publicly blame the doctor. It might make sense for the doctor to lose her job or to have to retrain in certain aspects of her job. Given the dynamics of the situation and agents involved, expressing blame might be both fitting and appropriate. But, before expressing blame, we must question the default assumption that having warrant to make judgments of blame automatically grants one warrant to publicly express blame. Harman and Sher's more rigid views add instructive depth and context-sensitive reasons for why expressed blame might be more appropriate in some contexts than others.

V. Confucian Support for Epistemic Skepticism

To deepen my discussion of epistemic skepticism, I will now turn to an entirely different framework, Confucianism. The Confucian text, the *Analects*, raises very similar concerns about judgments of culpability as the concerns expressed above. First, let me provide some background on this text. The *Analects* is a collection of conversations between Confucius, his students, and colleagues, collected between the fifth to third centuries BCE. It discusses what it means to lead a positive life and what it means to be a leader. One central concern in Confucianism, throughout its various developments and permutations, is the notion of self-cultivation and becoming an excellent ethical being.¹³⁸ For example, one topic discussed in the *Analects* is how reluctant we should be to make negative judgments of others.

¹³⁸ See Ivanhoe (2000) for a great introduction to the central themes of the *Analects*.

One question raised in the *Analects* in relation to this topic is: Under what circumstances should we give others the benefit of the doubt? Confucian scholar Hagop Sarkissian explains that the text addresses these topics from a very particular perspective—one of a highly interconnected social world. In his article “When you think it’s bad it’s worse than you think”, Sarkissian explains that according to the Confucian perspective,

How any single person acts in any social occasion hinges greatly on the behavior of the other individuals at hand. Hence, whenever one wishes to explain or understand another’s behavior—that is, whenever one were to judge it in some way—one would look beyond a person’s motivations, goals, or traits of character.¹³⁹

Confucians believe that we can only understand human behavior by thinking about that human behavior in our larger, interdependent social framework. Sarkissian further remarks that Confucius viewed “behavior as highly interconnected, prompted and shaped by one’s social and environmental contexts...it would seldom be appropriate to discount or overlook such factors in accounting for the person’s behavior, as they might carry great explanatory weight.”¹⁴⁰

Given the social nature of the Confucian framework, it is no surprise that the Confucian recommendation in cases of moral wrongdoing is to first consider any relevant social or environmental factors that may have factored into the wrongdoing. These external social factors might include: the nature of the wrongdoer and wronged agent’s relationship, any sort of communication or moral norms the agents have already established between them, or any prior history of conflict between them. Perhaps the wrongdoer is having an off day. Perhaps she was up all-night caring for her sick child. Perhaps she is in a new environment and does not know all of the local moral norms in her new culture.

¹³⁹ Sarkissian, H. (2015). “When you think it’s bad it’s worse than you think: Psychological bias and the ethics of negative character assessments” in *The Philosophical Challenge from China*. Brian Bruya (ed.). MIT Press, pp. 4-5.

¹⁴⁰Sarkissian, p. 4.

There are any number of personal, attitudinal, environmental, and situational factors at play in each of our behaviors and moral decisions. The Confucians were highly aware of and sympathetic to these external factors in judgments of culpability. Sarkissian adds,

There is a distinct pattern in these passages that concerns how the moral exemplar is supposed to react when dealing with recalcitrant, disagreeable, or otherwise bad individuals—that is, when one has reason and opportunity to make negative judgments of others. The pattern is one of caution and restraint.¹⁴¹

This Confucian caution and restraint strategy reinforces my argument that expressed blame is socially risky and best avoided. Indeed, the caution and restraint strategy highlights the social risks involved in expressing an incorrect judgment of blame. As Sarkissian further explains,

It is not uncommon to misinterpret others' signals or to fail to convey our own intentions clearly. Wires get crossed, identities are mistaken, and unwarranted assumptions are made. Sadly, such miscues are often taken to be highly diagnostic of character and purpose, weighted accordingly, and thus reciprocated by real-life 'defection'—our tendency to have negative impressions harden into obstinate beliefs.¹⁴²

Because the chance of misfire is so high in our judgments of others, the suggestion we can extract from the *Analects* is that fostering the habit of giving others the benefit of the doubt is, on the whole, preferable to rushing to express blame without certainty.

Steering away from a default stance of expressing blame immediately can allow “fruitful, constructive, and productive relationships to unfold”¹⁴³ between the wronged agent and wrongdoer, because suspending judgment of others does not risk getting judgments of culpability wrong. The thought is that due to the inherent risk of misjudgment, we should maintain a *default position of epistemic skepticism* with regards to culpability, recognizing that we could always be mistaken about someone's culpability. This initial epistemic skepticism preserves the relationship between the wronged agent and wrongdoer, and allows for these agents to

¹⁴¹ Sarkissian, p. 6.

¹⁴² Sarkissian, p. 14.

¹⁴³ Sarkissian, p. 14.

communicate about the wrongdoing together, without a potentially wrong judgment of blame looming over their discussion. To be clear, this is not a strategy of doing nothing. The Confucian account suggests pausing before taking action, but not it does not endorse taking no action at all.

The Confucian account offers another compelling epistemic case for exhibiting caution in expressing blame. Even if we are able to make a positive judgment of culpability with a high degree of certainty, we still need to attend to the social risks of expressing blame. Up until this point, the Confucian account has paralleled my own view that we ought to avoid expressing blame, in order to avoid getting it wrong. However, Sarkissian explains that writers in the *Analects* allow for expressions of blame, as long as the case of moral wrongdoing has been put under deep reflective scrutiny. Though passages in the *Analects* suggest pausing and reflecting as an initial strategy, expressing blame later on in the interaction is appropriate if it is ultimately deemed warranted.

The Confucian account correctly asserts that there are cases of moral wrongdoing in which we can clearly identify who the wrongdoer is, and thus sometimes we can be confident, if not certain, in our expressions of blame. The Confucian account acknowledges that sometimes blamers are in a very good epistemic position to judge an agent culpable, and in these cases blamers, should feel free to express blame to that agent. In chapter four, I will discuss the Confucian strategy in greater detail as part of a proposal for an alternative strategy of moral response instead of expressed blame. For now, my goal in presenting Sarkissian's interpretation of the Confucian framework has been to lend support to my argument that there are several epistemic reasons (from varied philosophical sources) to support the view that we ought to exert extreme caution before expressing blame.

VI. Conclusion

In this chapter, I have argued for several epistemic reasons for thinking expressed blame is a socially risky response to moral wrongdoing. I argued for what I called a *moderated epistemic skepticism* regarding both the culpability status of apparent wrongdoers as well as the question of whether expressing blame is warranted in any given situation. I argued that Rosen's account exaggerates the scope of epistemic skepticism regarding culpability, on the grounds that there surely are some cases in which we can be highly certain of one's culpability. Rosen is mistaken to think that we are *never* warranted to blame others. Sometimes, we really do know that a judgment of culpability would be warranted and whether it would be warranted to express that judgment through blame.

Further, I discussed Stanley's epistemic contextualism to argue that the stakes of a given case of wrongdoing can change our warrant for judgment in that particular context. I argued that expressions of blame involve higher stakes than private judgments of blame do. Those higher stakes involved in public blame include social risks such as over-exaggeration and escalation, retributive or hostile emotions, unfair or false accusation, and so on. The warrant we have to express public blame is determined by the specific context, stakes, and agents involved in a given case of wrongdoing. Even if we have the proper warrant to make a given private judgment of blame, that warrant does not automatically extend to the public realm. We might not have the warrant to express that judgment of blame because of the higher stakes involved in public expressions of blame.

Before I move on to chapter four's discussion of alternative responses to moral wrongdoing, I will first summarize the main arguments in these first three chapters. There are (at least) three

reasons to proceed with extreme caution before publicly expressing blame. First, I argued that blame is retributive at root, and even civilized versions of blame are not so many steps evolved from this primitive root. Second, I drew upon Nussbaum's account to argue that following our retributive impulses is irrational, because our retributive impulses are either premised on magical thinking, or on an excessive focus on one's status-rank, neither of which are helpful responses to wrongdoing. Relatedly, I presented a further worry about expressed blame when it is coupled with Senecan anger, an exacerbating feature of blame's retributive root. Third, assuming that blame ought only to be expressed to those who are actually blameworthy, and assuming that we often lack sufficient warrant to make judgments of culpability, expressing blame is a high-risk strategy. I have therefore proposed that we ought to make a much greater epistemic effort before expressing blame publicly.

Returning to the distinction between fittingness and appropriateness I raised in chapters one and two, I maintain that while deeming someone blameworthy may be *fitting*, it is often not *appropriate* to express that blame given the high social stakes of public expressions of blame. Similarly, I proposed that an initial Confucian stance of pausing and reflecting, rather than a default mode of blame, is best. If, after the blamer has paused and reflected on the nature of the wrongdoing and has decided blame is still appropriate, she has another hurdle to pass: the social risks of expressing blame. She can get a sense of these risks based on what she knows about the wrongdoer and the dynamics of their relationship.

My modified epistemic skepticism requires that all would-be blamers need to pass both the epistemic and social risk hurdles before making the decision to express blame. These hurdles are significant. We ought to adopt a differential approach between private judgments of blame and public expressions of blame, since they carry different levels of risk. The epistemic risk of

getting our private judgments of blame wrong, in addition to the heightened social risks of expressing blame make expressed blame, on the whole, inadvisable.

Chapter Four. Answerability Without Expressed Blame

I. Introduction

In this final chapter, I will present concrete and detailed ways to hold others responsible without expressed blame. I will also advocate for the importance of context-sensitivity in response to moral wrongdoing. It would be silly to think that there is only one way to respond to moral wrongdoing. Expressing blame is one option we have in our moral toolbox, among several others. Context determines which of the various modes of moral response best fits a given case of wrongdoing.

There are several reasons for this. For one, we each know the details of our own interpersonal lives as well as the nuances of our close relationships better than anyone else does. Some people are very defensive when they are blamed, others receive moral feedback well. Some agents we know very well, and thus deeply understand (and perhaps also sympathize with) the reasons behind their moral transgressions. Because of this close relationship, perhaps we are more willing to excuse or overlook their wrong behavior. Other agents we know less well, and in these cases we are perhaps are less willing to grant them the benefit of the doubt. These interpersonal nuances all impact whether expressing blame makes good sense or not in a given situation.

Second, the complexities of our interpersonal relationships are constantly changing; they are not static. As each of us grows and gains more moral knowledge over time, this affects the dynamics of each of our interpersonal relationships, as well as what sort of response to moral conflict best serves the current interpersonal dynamic of the relationship. Imposing artificial structures on our interpersonal relationships will not capture the variety and depth contained

within our interpersonal relationships as well as the varying degrees of psychological and emotional sensitivity among us. A one size fits all approach to wrongdoing will not sufficiently meet our diverse and evolving interpersonal moral needs.

With this context-sensitivity in mind, this chapter serves as a toolkit of non-retributive moral responses to wrongdoing. It remains an open question which of them is the *best* response to wrongdoing in a given situation. Expressed blame remains *an* option for responding to moral wrongdoing, but I argue that we also have several other options at our disposal that might be more efficacious.

There is a pervasive line of thought in the literature¹⁴⁴ that a wrongdoer cannot fully grasp the nature of her wrongdoing unless she feels appropriate *remorse* or *guilt* for her actions. Relatedly, many argue that remorse and guilt further serve vital motivating functions for wrongdoers. Vargas explains

Guilt and the process of repentance have restorative functions: they provide us with the impetus to undertake courses of action that repair or restore that status. In their absence, genuine acknowledgment of wrongdoing is difficult.¹⁴⁵

However, contra this standard view regarding guilt and remorse, Blustein argues that “It is simply not always the case that when someone acts wrongly, the most powerful or effective way to show him what it does to another person (including oneself), to elicit remorse, and to defend the moral order he violated is to punish or threaten to punish him for it.”¹⁴⁶ One of my overarching goals in this chapter is to demonstrate that Blustein’s assertion is correct. I will extend Blustein’s statement and argue that wrongdoers can grasp the nature of their wrongdoing without being made to feel guilty or remorseful by the blamer, and that doing so is preferable.

¹⁴⁴ For example, see works by Vargas; Fricker; Bennett; and Duff.

¹⁴⁵ Vargas, p. 263.

¹⁴⁶ Blustein, Jeffrey (2018). “Forgiveness and the Moral Psychology of Sadness”. Unpublished manuscript, p. 20.

Indeed, I do not think receiving blame (and subsequently feelings of guilt or remorse) is a necessary aspect of wrongdoers gaining the required motivation to make positive moral changes in their lives. We can hold wrongdoers responsible (construed as answerability, which I explain below), and they can have the proper motivation to change their bad behaviors, without directly blaming them and inciting remorse or guilt in them. This is not to say that the wrongdoer should feel no remorse or guilt about what they did. Negative feelings from the wrongdoer, blamer, or both seem an inevitable part of responding to wrongdoing, with or without blame. To suggest we can do away with negative feelings altogether in cases of wrongdoing seems unrealistic and unhealthy.

Instead, I argue that although finding fault (in the normative sense) in others is, to some degree, an inevitable part of living in a shared moral community, expressed blame and the subsequent cycle of inducing guilt and remorse in wrongdoers is not also an inevitable part of living in a shared moral community. Because there are alternatives to expressed blame, and because expressed blame comes with a host of negative, affective elements, it seems prudent to shift our moral responses away from expressing blame and towards these alternatives.

It is important to not only understand these alternatives to blame in theory, but also make sense of their practical applicability. Alternatives to blame, understood as legitimate responses to moral wrongdoing, have not been given enough credit in the moral philosophical literature; this chapter aims to fill that gap.

The alternatives I present in this chapter are each consistent with (but do not require) finding fault in a wrongdoer's behavior. There is a way for us to find someone blameworthy but not express blame to them. Our moral lives and interpersonal relationships are, and should be, dynamic and ongoing conversations. They are not static. We are each constantly growing and

expanding our moral knowledge and understanding, and this is reflected in our interpersonal relationships. In responding to moral wrongdoing, we will not get it right every time, and should not expect ourselves or others to. That is impossibly demanding. Instead, I aim to steer us away from our habitual mode of blaming as the default response to wrongdoing.

The three strategies I will discuss in detail in this chapter are:

1. Responsibility Without “Affective” Blame
2. Emotional Conversion
3. Confucian Answerability

The first strategy takes inspiration from Hanna Pickard and Nicola Lacey’s work, in which they use the term “affective” blame. As I will explain in greater detail, I take the term “affective” blame to mean roughly the same as the term expressed blame that I have been using throughout this dissertation to refer to affective, publicly communicated forms of blame. Further, as I will demonstrate, these first two strategies can work well alongside the third strategy, which is Confucian Answerability. In fact, Confucian Answerability serves as a *unifying* framework to understand each of these modes of alternative moral response.

II. Responsibility as Answerability

Before I discuss the various alternative responses, I ought to clarify how I am understanding the idea of *holding others responsible*. Since I argue that we can hold others responsible without expressed blame, it behooves me to get clear on what, exactly, I take holding others responsible to mean. There is widespread discussion in the philosophical literature of what it means to hold others morally responsible.¹⁴⁷ For example, in “Two Faces of Responsibility”,

¹⁴⁷ See Smith (2012); Smith (2005); Shoemaker (2011)

Gary Watson introduced a distinction between responsibility as attributability and responsibility as accountability.¹⁴⁸ Then, in “Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility” David Shoemaker expands on Watson’s argument and makes the case that there are three distinct conceptions of moral responsibility: attributability, answerability, and accountability. Finally, Angela Smith (2012) argues that answerability is the core of what is at stake in our practices of holding others responsible, and in what follows I will adopt her view.

Accountability typically refers to an agent being an appropriate object of the Strawsonian reactive attitudes.¹⁴⁹ Shoemaker argues that to hold someone to account is “precisely to sanction that person, whether it be via the expression of a reactive attitude, public shaming, or something more psychologically or physically damaging.”¹⁵⁰ Accountability, defined as Shoemaker describes, involves expressing blame publicly through various sanctioning practices and the reactive attitudes.¹⁵¹ Without public sanctioning, according to this view, a wrongdoer is not fully held to account for their actions. This is in line with Strawson’s view that our moral responsibility practices are inherently social. It is through publicly expressing the reactive attitudes “that we communicate to fellow members of the moral community our interpersonal expectation for a reasonable degree of goodwill.”¹⁵² Thus, according to the strategy of accountability, publicly expressed blame (through the reactive attitudes) is required for holding wrongdoers responsible. As should already be clear, I disagree with this account, since I do not

¹⁴⁸ Watson, Gary (1996). “Two Faces of Responsibility”. in *Agency and Answerability*. Oxford: Oxford University Press, pp. 260-86.

¹⁴⁹ See <https://plato.stanford.edu/entries/moral-responsibility/> for a more detailed account.

¹⁵⁰ Shoemaker, David (2011). “Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility,” *Ethics*: 121, p. 623.

¹⁵¹ For further support of this view, see Bennett (1980) and Wallace (1994).

¹⁵² <https://plato.stanford.edu/entries/moral-responsibility/>.

think we need to publicly sanction wrongdoers via the reactive attitudes in order to hold them responsible.

Answerability, on the other hand, refers to the way a given agent held to account as an appropriate object of justificatory challenge. Angela Smith argues that to hold someone responsible “is to say that that agent is open, in principle, to demands for justification regarding that thing.”¹⁵³ For her, the practice of holding someone responsible *just is* to hold them answerable for their actions. The blamer must demand reasons from the wrongdoer, and in response, the wrongdoer must provide those reasons for her actions. Smith further states that to blame an agent “always embodies (at least implicitly) a demand to her to justify herself.”¹⁵⁴ Relatedly, Shoemaker describes answerability as being “susceptible for assessment of, and response to, the reasons one take to justify one’s actions. The sorts of answers one gives will reveal one’s ends, the things one takes to be important.”¹⁵⁵

Answerability requires that we hold others responsible by tasking them with justifying their actions. This justification can happen via public blame that demands response from the wrongdoer, or via a process of moral inquisition that need not be based in public blaming at all. Recall McGeer’s argument from chapter one for the dialogical exchange between wrongdoer and blamer. This sort of back and forth dialogue lends itself well to the conditions of answerability. In the inquisitive dialogue session, the blamer demands justification from the wrongdoer, and the wrongdoer is given a chance to explain herself—to state her reasons for action—and potentially exonerate herself. However, it remains to be seen whether that inquisitive dialogue can effectively function without an underlying foundation of blame. In the next section, I will

¹⁵³ Smith, Angela (2012). “Attributability, Answerability, and Accountability: In Defense of a Unified Account”. *Ethics*, Vol: 122 (3), p. 578.

¹⁵⁴ Smith, p. 578.

¹⁵⁵ Shoemaker, p. 623.

analyze whether inquisitive dialogue can serve the goals of answerability *without expressed blame*.

Finally, attributability refers to cases in which a given action is attributable to a given agent. Attributability does not require sanctioning or justification. Instead, attributability accounts offer a “ledger” view of moral responsibility, according to which “the practice of ascribing responsibility involves assigning a credit or debit to a metaphorical ledger associated with each agent.”¹⁵⁶ In other words, an agent is responsible if a fault or credit is properly attributable to her; *she* did it, so the attributability is on her.

Shoemaker argues that “certain morally significant things may be properly ‘attributable’ to an agent, and therefore open her to certain distinctive forms of moral appraisal, without that agent being ‘answerable’ for those things.”¹⁵⁷ Intuitively, it makes sense that we can only blame a case of moral wrongdoing on a given agent if that wrongdoing is properly attributable to her. However, that a given wrong is attributable to a given agent is not sufficient to conclude that the given agent is responsible for the given wrong. For example, say that I was late to a business meeting due to an unforeseeable delay on the train. Though my tardiness is appropriately attributed to me, this is not sufficient reason to hold me responsible for my tardiness. I planned as best as I could to be on time, and circumstances beyond my control—unforeseeable transit delays—prevented my being on time.

At the same time, it seems that attributability is a baseline condition for accountability. Working backwards, it is logical that a given wrong action must be attributable to a given person in order to make her accountable and/or answerable for that wrong action. For example, a

¹⁵⁶ <https://plato.stanford.edu/entries/moral-responsibility/>. See also Watson (1987) and Frischer & Ravizza (1998) for further discussion.

¹⁵⁷ Smith, p. 578.

standard account of responsibility concludes that “it would seem unfair to hold someone accountable for an action via reactive attitudes such as resentment or indignation, if the action was not properly attributable to the agent--say, because she succumbed to a genuinely coercive psychological compulsion.”¹⁵⁸ Just because a given wrong action is attributable to a given agent does not imply or require that the agent is accountable or answerable for that action.

Shoemaker, then, argues that the concepts of accountability, answerability, and attributability come apart, both analytically and also extensionally. However, Smith responds to Shoemaker’s article in “Attributability, Answerability, and Accountability: In Defense of a Unified Account”, by arguing that there are not, in fact, three separate notions of responsibility. Instead, Smith argues that answerability is “the only kind of moral responsibility there is”.¹⁵⁹ Despite the clear differences in meaning among these three terms, Smith argues that our moral practices “do *not* embody three distinct conceptions of moral responsibility...our moral responses hang together in a theoretically unified and satisfying way around the notion of answerability.”¹⁶⁰ For Smith, answerability is a sufficiently unifying way to hold others responsible.

It is clear that attributability is not a unifying conception of moral responsibility, since its requirements are fairly thin. So, in order to investigate Smith’s claim that answerability is a unifying conception of moral responsibility, I will examine the distinctions between answerability and accountability to assess whether the latter can be collapsed into the former. Shoemaker defends the distinction between answerability and accountability on the grounds that accountability involves “nonsanctioning modifications to one’s actions and attitudes toward the

¹⁵⁸ <https://plato.stanford.edu/entries/moral-responsibility/>.

¹⁵⁹ Smith, Angela, p. 576.

¹⁶⁰ Smith, p. 589.

agent, while accountability licenses more explicit forms of sanctioning activity.”¹⁶¹ Smith agrees that we respond to moral wrongdoing with varying degrees of sanctioning activities, but does not think this fact means that accountability and answerability are two distinct forms of holding them responsible. Instead, both forms are unified under the umbrella of answerability, which can encompass both sanctioning and non-sanctioning modes of justificatory challenge.

I agree with Smith that answerability is the core of what is at stake in our moral practices of holding others responsible. Her view is entirely consistent with my argument that there are both blaming and non-blaming ways to hold others morally responsible. Some ways of responding to wrongdoing involve obvious sanctioning activities, while others do not. Put another way, some responses to wrongdoing involve accountability in the sanctioning sense, while some lack this sense of accountability. But, whether a response is sanctioning or not, when we hold someone responsible for a moral wrong, at a bare minimum we do hold them to a certain level of justificatory challenge.

Thus, Smith is right to think that our moral responses hang together in a reasonably unified way through the notion of answerability. At the very least, the notion of answerability provides a major foundational aspect of what it means to hold others responsible, because it demands that the wrongdoer *explain herself* in the fully normative sense of offering reasons. Even if there is no sanctioning activity involved in the response, holding others responsible still requires this normative sense of offering reasons. Accordingly, I will apply Smith’s argument about answerability to the issue of how we can hold each other morally responsible without expressed blame.

¹⁶¹ Smith, p. 586-87.

In what follows, I will discuss several alternatives to blame that meet Smith's answerability requirements without requiring expressed blame. I will argue that the process of moral inquisition through justification need not involve expressed blame. The mode of answerability can be consistent with expressed blame, but it does not require that expressed blame remain in the picture. In the three alternatives I present, I will demonstrate how each fits the requirements of answerability.

III. Responsibility Without (Affective) Blame

The first blame alternative draws on the account Lacey and Pickard present in their article "From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm". Building on Pickard's extensive experience in clinical psychology and Lacey's extensive experience in the legal field, the authors argue that we are better off addressing wrongdoers without affective blame. It should be noted that while Lacey and Pickard's model does not map onto the answerability framework perfectly, still the two models have a great deal in common. Further, the answerability framework is consistent with their focus on therapeutic and rehabilitative benefits to wrongdoers. In both of our accounts, the wrongdoer is actively involved in her own treatment and rehabilitation. I will first provide a critical summary of Lacey and Pickard's account, before assessing the feasibility and utility of their account as a non-blaming strategy.

They describe affective blame as:

The range of hostile, negative attitudes and emotions that are typical human responses to blameworthiness. Blame can include, for instance, hatred, anger, resentment, indignation, disgust, disapproval, contempt and scorn, and can be manifest in any number of ways,

including seeking retaliation, retribution, and vengeance, rejection and banishment from the community, and the withdrawal of basic respect.¹⁶²

As we can see from the list above, Pickard and Lacey pick out retaliation, retribution, and vengeance as fundamental aspects of blame. They note that while we tend to feel that these negative blaming attitudes and emotions are justified and appropriate ways to feel towards wrongdoers, these attitudes and emotions can be excessive, even irrational.¹⁶³ As I argued in chapters one and two, expressed blame very often involves retributive elements as well as the negative blaming emotions. In this way, Lacey and Pickard's term affective blame closely parallels what I have referred to as expressed blame throughout this dissertation, because we both identify retributive and hostile emotions as central to our blaming practices. Lacey and Pickard's overall strategy is to steer our habits of moral response to wrongdoing away from our traditional hostile and retributive methods and instead towards rehabilitative models. Their discussion remains focused on what rehabilitation for wrongdoers would look like in the legal realm. In this section, I will translate their focus in the legal realm into the non-legal realm of moral wrongdoing and posit what rehabilitation would look like in interpersonal moral contexts.

In "Responsibility without Blame for Addiction" (2016), Pickard cautions against the damaging effects of affective blame, remarking that

Blame is understood within clinical practice to undermine the capacity of responsibility and accountability to enable change and empower, because of its propensity to make patients feel rejected, worthless, ashamed and uncared for, thereby rupturing the therapeutic relationship as well as damaging any sense of hope for the future they might otherwise have, and, correspondingly, any motivation or belief that they really can overcome their difficulties.¹⁶⁴

¹⁶² Lacey, Nicola and Hanna Pickard (2012). "From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm." *Oxford J Legal Studies* vol. 33:1, p. 3.

¹⁶³ Lacey and Pickard, p. 19.

¹⁶⁴ Pickard, Hanna (2016). "Responsibility without Blame for Addiction". *Neuroethics* (Special Issue on Marc Lewis' *The Biology of Desire: Why Addiction is not a Disease*), p. 7.

Therapeutic patients who feel shame about what they did often feel less hopeful about the future. Pickard argues that this hopelessness largely prevents patients from gaining the motivation they need to want to make positive moral changes in their behaviors and actions. Without this motivation, these agents fail to change their bad behaviors, creating a loop of hopelessness in their minds. Simply put, it does not work.

Even though this empirical finding stems from a specific setting (a clinical therapeutic context) I will discuss Lacey and Pickard's work in the context of interpersonal wrongdoing more generally. I predict that affective blame has all the same flaws in daily interpersonal cases of wrongdoing as it does in clinical therapeutic settings. The two contexts are not dissimilar. Lacey and Pickard also recognize their account's wider applicability, explaining

Nobody likes to be blamed. We all have some grip on what it is like to feel the object of another's hostile, negative attitudes and emotions, and why this might be detrimental to one's sense of self-worth, one's relationships with others, and one's hopes for the future and motivation to change aspects of oneself that are difficult to face.¹⁶⁵

Instead of receiving affective blame, Lacey and Pickard suggest that wrongdoers need to be encouraged to recognize their capacity for change, and to be praised for their progress. Responsibility in this context is not concerned with a negative moral evaluation, but instead with the wrongdoer's motivation and capacity to change.¹⁶⁶ This is a forward-looking, hopeful approach. Yet, this is not to suggest that wrongdoers are not held to account for their actions. Part of the work of the therapist (in the clinical therapeutic model) "involves pointing out or indeed imposing consequences should they fail."¹⁶⁷ On Lacey and Pickard's model, wrongdoers are still held responsible without expressed blame. If patients fail to learn from their transgressions, the therapist's job is to point this out and create appropriate consequences for

¹⁶⁵ Lacey and Pickard, p. 14.

¹⁶⁶ Lacey and Pickard, p. 16.

¹⁶⁷ Lacey and Pickard, p. 13.

their patients. Similarly, outside of therapeutic contexts, we can hold others responsible (by pointing out consequences to wrongdoers) without expressing hostile or retributive forms of blame. In this way, Lacey and Pickard advocate for responsibility without affective blame, and their view can reasonably extend to non-therapeutic contexts, too.

As Lacey and Pickard's strategy demonstrates, we can hold others responsible (construed as answerable for their actions) without affective blame. Holding others responsible can be *rehabilitative* and need not involve the charges of retribution, anger, indignation, resentment, etc. For example, when my student failed to turn in his final paper, he failed my class. This is an obvious punishment for an obvious failure to meet established expectations. I did not get angry or blame him for his choice, he simply suffered a negative consequence for his choice not to write the final paper. Similarly, in our moral lives, there can be consequences for our actions that do not involve such affectively-heated elements, as is the case most strongly in angry blame.

Lacey and Pickard's clinical therapeutic model replaces affective blame with what they call *detached blame*, which provides an attitude of "concern, respect and compassion for the person, while nonetheless questioning, challenging and reproofing their conduct."¹⁶⁸ Detached blame involves holding others responsible without the hostile, incipiently retributive elements of blame. Their detached blame model demonstrates that the attitudes of concern, respect, and compassion, rather than the typical, hostile and angry blaming emotions can effectively hold wrongdoers responsible. With the detached blame therapeutic model, the communication towards the wrongdoer involves concern and compassion, alongside basic respect and decency, even in the face of blameworthiness.¹⁶⁹ Rather than respond to wrongdoers with anger, indignation, and resentment, Lacey and Pickard argue that we ought to respond to wrongdoers

¹⁶⁸ Lacey and Pickard, p. 19.

¹⁶⁹ Lacey and Pickard, p. 19.

with care and compassion. The difference in emotional tone between the two approaches (hostile versus compassionate) is striking and runs completely counter to our traditional way of dealing with moral conflict.

Some readers might worry that Lacey and Pickard's approach, while appealing in theory, will not provide a practical or realistic way to hold wrongdoers to proper account, especially in a legal context, for their actions. However, Lacey and Pickard are not opposed to punishing wrongdoers (even in a legal sense) for crimes they have committed. Lacey and Pickard are not opposed to punishing wrongdoers in appropriate and proportionate ways given the severity of their wrongdoing. But, any punitive consequences ought to be imposed for "psychiatric improvement, not out of retaliatory vengeance."¹⁷⁰ Cast in this light, holding wrongdoers responsible consists of a rehabilitative, not retaliatory aim.

Lacey and Pickard's account extends well beyond the narrow confines of their clinical therapeutic context. Indeed, the overall goal of Lacey and Pickard's model, applied to everyday cases of moral wrongdoing, is to integrate the rehabilitative ideal and justice model in order to "move us away from backward-looking retaliation and wrath, and towards a more human and forward-looking attitude towards offenders."¹⁷¹ Lacey and Pickard's account makes good sense as a theory of punishment in the legal realm, the context they are focused on. If a wrongdoer is sentenced to prison, then it should be a prison sentence that involves work experience, clinical therapeutic forms of rehabilitation, positive reward schemes for good behavior, educational opportunities, etc.

Further, Lacey and Pickard's model is unlike traditional forms of expressed blame in two senses. For one, it removes the negative affective element. For another, its aim is explicitly

¹⁷⁰ Lacey and Pickard, p. 19.

¹⁷¹ Lacey and Pickard, p. 28.

rehabilitative, not retaliatory. Both of these aims strongly support my own view which is that affective and retributive forms of blame are unnecessarily hostile and socially damaging, and that we can hold wrongdoers to account without these elements. Agreeing with Lacey and Pickard, Blustein writes

The suffering or distress of the offender may not be necessary to accomplish (blame's goals). Perhaps offenders can be brought to fully acknowledge the moral significance of what they have done and to atone for it without needing to use distress as the medium through which to bring this about, by instead showing them love and compassion or by exposing them to positive role models.¹⁷²

Blustein helpfully adds that not all wrongdoers need to receive affective blame in order to acknowledge the significance of their actions and be motivated to change. This further supports my argument that in cases of interpersonal moral wrongdoing *context matters* a great deal. There is a great degree of variation among agents' temperaments, psychological states, and values. Some wrongdoers are more emotionally sensitive than others, so much so that they are able to *self-correct* their moral transgressions.¹⁷³ These agents likely do not need to be publicly blamed in order to have the motivation to make positive moral changes to their behaviors. Other agents need more prodding.

A. Answerability Without (Affective) Blame

Lacey and Pickard's argument about rehabilitation in the legal realm is highly applicable to the interpersonal realm, too. Providing rehabilitation for wrongdoers in interpersonal settings is

¹⁷² Blustein, Jeffrey (2018). "Should We Get Rid of Blame?" Unpublished manuscript, p. 21.

¹⁷³ Blustein also mentions the possibility of exposing wrongdoers to positive role models as one way to bring them to acknowledge the moral significance of their actions. That strategy removes the emotional distress involved with receiving affective blame. Instead, it aims to rehabilitate and uplift wrongdoers and have them be influenced by and mirror the morally superior behavior of a role model. This seems promising in some cases of wrongdoing, especially when the wrongdoer does not have positive role models in her life already. Looking up to someone can equally serve as a motivation to change one's moral behavior.

consistent with holding them answerable to their wrongs. Further, this can happen without affective (used interchangeably with “expressed” here) blame. Say that my partner and I have gotten into the same argument about him failing to complete his household chores dozens of times. We repeatedly engage in this same conflict and each time, no new information is expressed from either of us. It happens again. I recognize that my blame—expressed in the exasperated tone of anger and resentment—is not motivating him to change his behavior and actually complete his chores. My expressions of blame have repeatedly failed to motivate him to do the dishes. I have a choice in front of me. I need to either give up my hope that he changes his behavior, or choose another strategy of response to get him to understand and be motivated to change. If I continue to express blame about this same issue, it will likely lead to heightened resentment, anger, or defensiveness, from one or both of us.

Instead, say that I adopt Lacey and Pickard’s strategy—an approach grounded in curiosity and compassion, without affective blame. Perhaps this involves me asking him questions about his choice not to take out the trash, do his dishes, etc. I get curious. I aim to get more information from him and to deeply listen to his point of view about our conflict. In getting more information from him, I prompt him to justify his actions. In doing so, I try to get to the root of the problem, to better understand why this behavior keeps happening. Maybe my expressing blame to him makes him extremely defensive and self-righteous, encouraging him to dig in his heels about this point of conflict further. Perhaps neither of us really understands the other person’s needs about this issue, and we need to engage in non-blaming, inquisitive moral dialogue in order to get to the root of one another’s needs.

With this strategy of inquisitive moral dialogue, I do not aim to point fingers at him or scold him. Instead, I aim to gain information and understand his own reasons for action. He

surely is aware that his behavior hurts me, and that he is failing to live up to a prior agreement we made. My expressing blame towards him is not uncovering any new insight he did not already have. By engaging in this dialogue, I might be able to uncover the root of his behavior, and we can jointly figure out how to solve the problem. Even though I do have hurt feelings about this repeated issue, I can recognize that a hostile and retributive expressed blame approach is simply not productive. Furthermore, my inquisitive dialogue meets the conditions of answerability insofar as it seeks a justificatory response from the wrongdoer.

I will now extend a Lacey and Pickard-type detached blame model into my own, more targeted account of how to hold others answerable without affective blame in our interpersonal moral lives. The specific factors involved in this method are: 1. Compassion, 2. Curiosity, and 3. Inquisitive dialogue. On my account, prioritizing a sense of compassion and real curiosity about the wrongdoing is consistent with still holding wrongdoers responsible. In addition, as I will argue, compassion and curiosity are not on a par with the third factor, inquisitive dialogue. Instead, compassion and curiosity are useful features that support the primary goal of fostering inquisitive dialogue between the agents as means to hold the wrongdoer answerable for her actions.

B. Inverting the Paradigmatic Response: Foregrounding Compassion

i. Compassion

Compassion may underlie expressions of blame and non-blame alike. For example, angry blamers may think that despite the harshness of their expressions of blame, underlying that harshness is compassionate, warm concern for the wrongdoer's moral growth and wellbeing. As the saying goes, we often get most angry at those we care about (or love) the most. As we saw in

chapter two, that is certainly Susan Wolf's view. However, it can be hard for receivers of angry blame to *feel* that underlying compassion and love, in any real sense, because what is expressed is often emotionally heated and hostile. We have all heard the expression "It is not what you say it is *how* you say it".

The forcefulness of the blaming emotions involved in most expressions of blame likely mask any underlying senses of care and compassion. Of course, some cases of wrongdoing call for harsh responses. Sometimes, if the blamer communicates in a calm and relaxed manner, the wrongdoer might not get the moral message forcefully enough and might do serious harm to herself or others. However, in our interpersonal moral lives, not all cases are so urgent or grave.

In these less urgent and grave cases, a Lacey and Pickard-type approach, which puts compassion at the forefront of the communication, is a viable blame-alternative. In this strategy, the blamer's compassion is not hiding in the background, underneath the harshness of expressed blame, but is instead leading the charge of the communication. Specifically, I propose *inverting* the communication such that care and compassion are clearly expressed up front. Further, any responsibility measures are held secondary to the primary overt focus of verbally expressing and demonstrating care and compassion for the wrongdoer. Foregrounding compassion and care does not restrict one from finding fault in the wrongdoer's behavior and holding them responsible later.

This inversion has two significant benefits in our interpersonal moral lives. First, it can reduce the wrongdoer's tendency to feel defensive or attacked. Second, it can create space for and foster productive dialogue between the two agents. This moral dialogue can lead to moral understanding between the agents as well as rehabilitation of the wrongdoer. The wrongdoer is

able to feel overtly cared for despite her wrongdoing and can heal and commit to acting in morally better ways in the future.

By reducing the manifest emotional heat and hostility involved in typical expressions of blame, the inquisitive dialogue between blamer and wrongdoer is more likely to go better. This argument is supported by empirical work on anger. In “The Positive and Negative Effects of Anger on Dispute Resolution: Evidence From Electronically Mediated Disputes”, authors Friedman, Anderson, Brett, Olekalns, and Goates argue that “an expression of anger lowers the resolution rate in mediation and that this effect occurs in part because expressing anger generates an angry response by the other party.”¹⁷⁴ Again, this is a modest proposal, and one that is consistent with judging others’ behavior blameworthy and also expressing blame to them if it is deemed appropriate after the compassion-based inquisitive dialogue. In sum, this paradigmatic inversion seeks to provide a safe and more caring space for moral dialogue.

ii. Curiosity

Second, curiosity is a central aspect of non-hostile and non-retributive responses to wrongdoing. Specifically, curiosity involves the inquisitive dialogue aspect I mentioned in the example of my unmotivated, messy partner. Before expressing blame, a blamer can simply aim to get curious about why the wrongdoer acted as she did. She can ask questions in a non-judgmental, non-blaming tone. This does not prevent her from expressing blame to the wrongdoer later, it might just delay the process. Alternatively, it might lead to a new insight that

¹⁷⁴ Friedman, Ray, Cameron Anderson, Jeanne Brett, Mara Olekalns, and Nathan Goates (2004). “The Positive and Negative Effects of Anger on Dispute Resolution: Evidence From Electronically Mediated Disputes” *Journal of Applied Psychology*: 89 (2), p. 369.

¹⁷⁵ See also: Tiedens, L. Z. (2001). “Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral.” *Journal of Personality and Social Psychology*: 80(86).

prompts her to conclude that expressing blame to the wrongdoer is not appropriate for this particular situation. Either way, the benefit of getting curious is that it helps mitigate defensiveness from the wrongdoer. By foregrounding curiosity, the blamer is neutrally seeking information, rather than actively seeking to criticize the wrongdoer. Even if the same information is conveyed in the curiosity strategy as in the typical expressed blame paradigm, the difference in tone is striking. In getting curious, the blamer delivers her questions without a loaded judgment of culpability towards the wrongdoer. The affectively-neutral delivery of the curiosity strategy better allows for authentic dialogue to occur.

Indeed, in getting curious, the blamer aims to get clarity about why the wrongdoer acted as she did, and the wrongdoer has a chance to explain herself. The blamer does not aim to accuse; she first seeks to understand and to gain more information. In doing so, this process of moral inquisition acts as a way to hold wrongdoers answerable for their actions. Wrongdoers are called to justify their actions and provide reasons through the back and forth dialogue. Further, this process meets the conditions of answerability. Foregrounding curiosity will not work for every case of moral wrongdoing. For example, when I witness my friend shouting racial slurs at a stranger, getting curious is clearly not the best initial approach. Time is of the essence in this case. But in lower stakes, less urgent cases of blame, there will be more space and time for an approach grounded in curiosity.

Additionally, this curiosity strategy has theoretical support from Rosen's account as well as the Confucian epistemic account discussed in chapter three. Indeed, as Rosen highlighted, there are limits on our epistemic capacities such that getting curious is a useful step towards gaining more information about the case that we are unlikely to already have complete insight about. In getting curious, the blamer seeks to better understand the wrongdoer's motives,

perspective, and any external factors involved in the given case of wrongdoing. In doing so, she gathers useful information and meets the conditions of answerability insofar as she seeks reasons (in the normative sense) from the wrongdoer. The curiosity strategy is consistent with the blamer deciding to express blame after the process of moral inquisition. However, expressed blame is not required. We can hold wrongdoers answerable for their wrongs through a focus on moral dialogue, because in the inquisitive dialogical exchange, each agent has a chance to offer their own reasons for action.

Finally, this curiosity strategy has additional support from Lacey and Pickard's therapeutic model of holding others responsible, in which they endorse an attitude of "concern, respect and compassion for the person, while nonetheless questioning, challenging and reproofing their conduct."¹⁷⁶ Expressed blame has largely become the default first response to wrongdoing in our culture. I propose, by contrast, that we cultivate a curious, dialogue-based first response in the aftermath of moral wrongdoing. This proposal requires that we invert our paradigmatic response to wrongdoing. The inquisitive dialogue that forefronts compassion might reveal something unknown, which might alter the need for blame. Moreover, the inquisitive dialogue might better foster authentic communication and resolution between the agents. In either of these cases, the outcome is less affectively-heated and retributive-based than traditional forms of expressed blame.

iii. Answerability

Third, we can hold wrongdoers responsible (construed as answerability) largely through the process of moral inquisition itself. I have already argued this in my discussions of

¹⁷⁶ Lacey and Pickard, p. 19.

compassion and curiosity above, but I will now make this point more explicitly and in greater detail. Once the blamer has questioned the wrongdoer and expressed compassion, she has more information about the nature of the wrong. Then, the next step is to decide how to hold the wrongdoer responsible. Lacey and Pickard advocate holding wrongdoers responsible through traditional punitive means, i.e. legal liability, but without the hostility and retribution involved in affective styles of blame.

Differently, I argue that what it means to hold a wrongdoer responsible depends on the nature of the relationship between the two agents, and the nature of the specific wrongdoing. Like deciding whether or not to express blame, deciding what it means to hold others responsible is similarly dynamic and context-sensitive. There is not a one-size fits all approach to holding others accountable, especially in our evolving interpersonal relationships. In our interpersonal lives, there are some people who we are more willing to be lenient towards in our judgments of moral responsibility, because we know personally that they are going through a hard time or are dealing with a something that warrants special consideration. This variation is to be expected. There is, inherently, variability in the ways that moral responsibility is administered within our close, personal relationships.

How to hold an agent responsible in a given situation will depend, for one, on what has transpired through the inquisitive dialogue between blamer and wrongdoer. There are a variety of ways to hold others to account, as well as a variety of ways to express it. The position I am proposing here offers a way to meet the demands of answerability without expressed blame, through the process of inquisitive dialogue itself. Getting curious and asking the wrongdoer questions in a compassionate manner is consistent with demanding a justificatory response from her.

Putting these three elements together, I have proposed a combination of Lacey and Pickard's detached blame model combined with a new element: a heightened emphasis on prioritizing inquisitive dialogue. I am not suggesting that this method will work in every case of interpersonal wrongdoing. It might require emotionally engaged, vulnerable, and sensitive humans to engage in this sort of dialogue. It is hard emotional work, and I deeply recognize that. Nevertheless, it is valuable to carve out space for this sort of response to moral wrongdoing as an alternative option to expressing blame.

IV. Transforming Hostile Emotions: The Emotional Conversion View

Chapter two surveyed Nussbaum's account of Transition-Anger, in which she argues that we ought to transform our heated anger into Transition-Anger, as quickly as possible. In chapter two, I discussed why angry blame in particular is especially pernicious. Now, I will argue for a similar emotional transformation process, with respect to our hostile blaming emotions more generally. This argument is focused on especially heated forms of angry blame and on converting angry blame into cooler emotions. One way to respond to wrongdoing without expressed blame is to *transform* our hostile blaming emotions into their cooler, less aggressive counterparts. Instead of getting angry when someone wrongs us, perhaps we can channel and convert that anger into, for example, disappointment, grief, or sadness. As we shall see, there is reason to regard this emotional transformation as preferable not only from a social perspective, but further from the point of view of answerability. While this strategy of emotional transformation is often not a sufficient response to wrongdoing on its own, I argue that it is an

effective strategy when used in tandem with other strategies such as getting curious and engaging in inquisitive dialogue.

In this section, I will first discuss what these cooler emotions consist of, then I will argue why we can meet the conditions of answerability through cooler modes of response such as sadness, disappointment, or grief. In doing so, I argue that the emotional conversion view is a strong alternative response to moral wrongdoing.

A. Emotionally Cooler Modes of Response

Recent empirical work demonstrates the utility of responding to wrongdoing by converting angry emotions into cooler emotions, such as sadness. Before discussing this research, I will first present recent work by Jeffrey Blustein on sadness in particular. Although his work remains focused on sadness, I take it to be the case that the same argument holds true for other cool emotions such as disappointment, grief, and even compassion. So, this more focused discussion of sadness supports my overall goal of arguing that emotionally cooler modes of response can be effectively converted from the retributive or otherwise hostile emotions that tend to accompany expressed blame.

Blustein (2017) argues that we can effectively respond to wrongdoing with sadness instead of heated forms of blame, such as angry blame. Sadness is a very different emotion than anger.¹⁷⁷ For one, payback is not a conceptual feature of sadness the way that payback (as I have argued in chapters one and two) is a conceptual feature of expressed blame and anger. Second, sadness is an inward-looking emotion, whereas anger often spills outside of us and tends to be directed at others. Some philosophers claim that the fact that sadness lacks this payback aspect

¹⁷⁷ For a more detailed discussion, see: Bonanno, George (2009). *The Other Side of Sadness: What the New Science of Bereavement Tells Us About Life After Loss*. Hatchett Book Group, p. 30.

automatically disqualifies it from the range of appropriate responses to moral wrongdoing.

Christopher Evan Franklin (2013) holds this view. Franklin argues that

Expressing sadness does not indicate to others that an object of moral value was treated in an objectionable way and thus does not help to safeguard the value against further mistreatment.... (thus) sadness is not an apt substitute for blame since it does not carry the condemnatory aspect embedded in blame.¹⁷⁸

On Franklin's view, we *must* respond to wrongdoing with condemnation, and blame is the best way to do so. Sadness does not meet this condemnatory requirement. In addition, unlike anger, indignation, or resentment, sadness is generally thought of as a softer or gentler emotion. Sadness, generally speaking, does not sting its intended recipient in the way anger does. Sadness does not seem as antagonistic as anger. Yet, sadness still (like anger, indignation, etc.) has a negative valence, which may lend itself well to responding appropriately to moral conflict. Responding to a serious moral wrongdoing with joy, for example, would be intuitively odd and discomforting. But because sadness is negatively charged, it is not so far removed from traditionally negative responses so as to seem out of place.

David Goldman (2014) argues that sadness can play a functional role similar to resentment or indignation in response to moral wrongdoing, and that sadness is in fact preferable to resentment or indignation. One might worry, however, that sadness is too passive a response to serious moral wrongdoing. This worry is inaccurate. Sadness, and similarly cool emotions need not be passive nor announcements of resignation. Instead, the reflective and introspective space cooler emotions bring can help an agent clarify how she feels about the wrongdoing. Cooler emotions provide an active, emotionally rich resource to process and respond to moral wrongdoing insofar

¹⁷⁸ Franklin, Christopher Evan (2013). "Valuing Blame." In *Blame: Its Nature and Norms*, eds. D. Justin Coates and Neal A. Tognazzini. New York: Oxford University Press, p. 222.

as they can relay a clear sense of disappointment to the wrongdoer, while simultaneously allowing space for introspection from both agents.¹⁷⁹¹⁸⁰

Further, Blustein argues that sadness “does not necessarily prevent one from holding the wrongdoer responsible for his actions or from deliberating about and taking these further steps: it may in fact sustain the effort to do so. And importantly, it can do so even in the absence of anger or some other vengeful passion.”¹⁸¹ At the same time, a sad response to moral wrongdoing is consistent with finding fault in wrongdoers’ behaviors. Emotions like sadness, disappointment, and grief are thus very flexible and helpful first responses to moral wrongdoing, because (among other things) they are non-hostile, non-retributive, and do not prohibit further modes of response.

B. Empirical Research on Anger and Sadness

Recent empirical work further supports the view that emotionally cooler modes of response are helpful first responses to moral wrongdoing. In particular, recent findings suggest that cooler emotions like sadness can blunt, or prevent, our experience of heated emotions like anger. Karen Page Winterich, Seunghee Han, and Jennifer S. Lerner (2010) have termed this effect the ‘Emotional Blunting Hypothesis’, which predicts that a specific emotion can carry over to blunt the experience of a subsequent emotion. The authors’ studies confirm that inducing sadness

¹⁷⁹ Relatedly, disappointment has many characteristics in common with sadness. In “The Moral Emotions”, Jesse Prinz argues that disappointment is “just a form of sadness.” On Prinz’s view, disappointment would not be considered a separate emotion, but instead a strand of sadness itself.

¹⁸⁰ For further literature on sadness, see specifically: Kurt Gray and Daniel Wegner (2011). “Dimensions of Moral Emotions”. *Emotion Review*: 3(3), p. 258; Phoebe Ellsworth and Craig Smith (1998). “From Appraisal to Emotion: Differences Among Unpleasant Feelings,” *Motivation and Emotion*, vol. 12, no. 3, pp. 271-302; Galen Bodenhausen, Lori Sheppard, and Geoffrey Kramer (1994). “Negative affect and social judgment: the differential impact of anger and sadness,” *European Journal of Social Psychology*, vol. 24, pp. 45-62.

¹⁸¹ Blustein, p. 15.

blunts subsequent anger, and inducing anger blunts subsequent sadness.¹⁸² This finding, if true, offers substantial support for my own view, the Emotional Conversion View.

For example, say that I find out that my coworker has taken credit for a project I completed all by myself. Our mutual boss praises my coworker for her excellent work, and as a result gives her a raise. I consider this betrayal morally wrong. If I experience sadness in response, the Emotional Blunting Hypothesis suggests that my experience of sadness will likely blunt my experience of anger. I cannot feel both sad and angry at the same time (or at least not both sad and angry in full force at the same time). Additionally, Winterich, Han, and Lerner suggest that sadness is characterized by situational agency, whereas anger is characterized by appraisals of human or individual agency.¹⁸³¹⁸⁴

If I were to get angry about the situation, that anger could quickly escalate, and I might lash out and say something I do not mean to my coworker or, even worse, our boss. As I argued in chapter two, a strongly negative feature of Senecan anger is its tendency to spiral out of control. My heated anger, the Emotional Blunting Hypothesis suggests, will also blunt, if not prevent, me from experiencing the event through a cooler emotion, like sadness. Knowing this, it seems less socially risky and therefore, on the whole *preferable*, to aim to respond to this case with sadness instead of angry forms of blame. It is less socially risky because experiencing the wrongdoing through sadness blunts my chance of

¹⁸² Winterich, Karen Page, Han, Seunghye, and Lerner, Jennifer S. (2010). "Now That I'm Sad, It's Hard to Be Mad: The Role of Cognitive Appraisals in Emotional Blunting". *Personality and Social Psychology Bulletin*: 36(11), p. 1467.

¹⁸³ Winterich, Lee, and Lerner, p. 1469.

¹⁸⁴ This finding is confirmed by Haidt (2003) who showed that anger is specifically aimed at regulating behavior and opinions towards *people* not situations. We can make sense of this finding in the case of expressed blame by understanding sadness as an affective response that arises due to negative feelings about the situation (the wrongdoing), and understanding anger as an affective response that arises due to negative feelings about the individual, (the wrongdoer herself).

expressing anger, and potentially losing my temper, escalating the situation, and inciting lasting emotional damage.

That being said, I am not suggesting we can (or should) cherry-pick our emotional responses to being wronged. In the heat of the moment, I might not be able to decide whether I want to experience the case of wrongdoing through sadness or anger. I am not under the delusion that we can (or should) choose our emotional responses. Nevertheless, armed with the knowledge of the Emotional Blunting Hypothesis, trying to experience the situation through sadness, to whatever degree possible, is a worthy, and socially useful aim.

Beyond anger's tendency to escalate, there are additional reasons why sadness is a socially preferable response. For one, anger is an intrinsically unpleasant and negative emotion to be on the receiving end of. Second, anger is instrumentally dysfunctional as a response to wrongdoing because it promotes defensiveness and is unlikely to motivate the wrongdoer to open up and feel encouraged to make positive moral changes in her behavior. If the goal of expressing angry blame to a wrongdoer is to motivate her to change her behavior, anger is likely not the most effective way to achieve this goal. Instead, by actively attempting to *blunt* our natural angry responses and substitute them for a cooler emotion, we can transform our habits of moral response towards more effective and compassionate ways of engaging with wrongdoers.

In her article "What's the Point of Blame? A Paradigm Based Explanation" Miranda Fricker describes several pathologies of blame. As she argues, one of these pathologies is when blame is not administered at the appropriate *register*. Fricker states

Blame should be appropriately contained in its proper remit, both temporally and in terms of the relationship(s) it affects. This means that blame's expression should not be allowed

to go on too long, and should not be allowed to migrate into regions of the relationship where it does not belong, or indeed to wantonly damage other relationships into the bargain. Blame should be allowed neither to fester nor to spread.¹⁸⁵

In the passage above, Fricker argues that blame should not linger excessively long, nor should it seep into parts of the agents' dynamic where it does not belong. Blame needs to be contained, temporally and in terms of which agents it affects. One clear advantage of responding to wrongdoing with sadness, over anger, is that sadness starts the conversation at an emotionally-cooler register. Because of this, the chance of the communication festering where it does not belong is already lessened. The troubling spiraling and escalating characteristics of angry blame I have discussed throughout this dissertation do not apply to sad expressions of blame. While sadness may be partially contagious in that we are sometimes negatively affected by others' sad moods, its spread is not nearly as pernicious as anger's is. Further, I am positing, based on Winterich, Lee and Lerner's findings, that sadness is less likely to fester and spread in emotionally damaging ways than anger is. Therefore, as this empirical research demonstrates, we have good reason to think that sadness may be a good antidote to anger, and that we can transform our angry blaming responses to sad responses.

C. Answerability Through the Emotional Conversion View

In "Should We Get Rid of Blame", Blustein suggests that perhaps no blame whatsoever is needed in order for wrongdoers to understand the moral significance of their behaviors. He explains,

Perhaps offenders can be brought to fully acknowledge the moral significance of what they have done and to atone for it without (distress), by instead showing them love and compassion or by exposing them to positive role models... There is no reason to think that moral understanding and self-correction could not be accomplished by measure that do not

¹⁸⁵ Fricker, Miranda. (2016). "What's the Point of Blame? A Paradigm Based Explanation". *Noûs*, 50, p. 169.

reprimand, measures that might in fact be preferable to blame because there is less of a risk of backlash and defensiveness from the offender who stands accused of wrongdoing.¹⁸⁶

The view Blustein expresses in the passage above is fairly close to the view I will now propose. I argue that an emotionally cool response to wrongdoing can sometimes be a *better* affect to accompany the inquisition of answerability than more hostile responses like anger.

Some might argue that a response of sadness, disappointment, or grief is incompatible with holding wrongdoers answerable for their actions because these cooler emotions are largely outcome or situation-oriented, rather than individual focused, and there is a real worry that this situation-focus prevents us from holding specific individuals responsible for their actions. Put another way, the worry is that an attitude of generalized sadness, disappointment, or grief does not pick out a specific person as having done something morally wrong, and thus is not a sufficient way to hold others answerable for their actions.

We can resolve this worry. Emotionally cooler responses to wrongdoing are not incompatible with holding a specific person answerable for her morally wrong action(s). For one, the emotional component of a wronged agent's response is not the only mechanism by which answerability happens. As mentioned above, the emotional conversion technique can work *in tandem* with getting curious and asking the wrongdoer questions about her actions—and in doing so making her a proper object of justificatory challenge.

For example, I can feel generalized sadness about the wrong having happened while also feeling and expressing that a specific agent is culpable. The emotional component of expressed blame need not be directed exclusively on the wrongdoer herself. That exclusivity is needlessly limiting. Assuming that responses like sadness, disappointment, and grief are antithetical to

¹⁸⁶ Blustein, p. 21.

holding wrongdoers responsible (construed as answerability) makes it seem as if our (heated) emotional reactions are the only way we hold others responsible. Further, not only are affectively cooler responses consistent with the aim of answerability, they can be a *better* way to achieve the aim of answerability. For one reason, a major benefit of converting our hostile and retributive blaming emotions into emotionally cooler substitutes is that doing so does not risk escalating the situation and dynamic between the agents. As Blustein puts it in the above passage, with sadness there is “less of a risk of backlash and defensiveness from the offender who stands accused of wrongdoing.”¹⁸⁷ Second, these emotionally cooler modes of response fundamentally lack the retributive edge that is so troubling about angry forms of blame.

Third, it may be easier to demand justification from the wrongdoer when that demand is delivered in an emotionally cooler tone. The blamer can more easily probe the wrongdoer for her reasons for action with an emotionally cooler tone. With transforming anger into emotionally cooler substitutes, there is less tendency of defensiveness from both agents and a better chance for productive dialogue between the agents. Fourth, responding to moral wrongdoing in emotionally cooler ways does not prohibit further modes of response, including expressing blame. Other things being equal, then, it seems best to respond to moral wrongdoing first with sadness, grief, or disappointment, rather than more hostile responses like anger.

V. Confucian Answerability

A third strategy of responding to wrongdoing without expressed blame stems from the Confucian epistemic skepticism I surveyed in chapter three. This method additionally provides

¹⁸⁷ Blustein, p. 21.

us with a practical psychological technique for converting retributive or hostile blaming emotions into calmer ones. I have termed this third strategy Confucian Answerability. Like the first two strategies, this third strategy also meets the conditions of answerability, and does so without expressed blame. Moreover, this third strategy may well provide a *unified* way to meet the conditions of answerability and combine elements of the first two strategies surveyed in this chapter. In this way, Confucian Answerability provides a comprehensive account of how to achieve answerability without expressed blame.

A. Sarkissian's Commentary

In his papers “When you think it’s bad it’s worse than you think: Psychological bias and the ethics of negative character assessments” and “Virtuous Contempt in the *Analects*”, Sarkissian offers an interpretation of the *Analects*, a collection of sayings of Confucius and his followers. Sarkissian’s interpretation of the *Analects* is one interpretation out of many interpretations of the text. Sarkissian’s view is not necessarily *the* Confucian view, but rather is a compelling interpretation of the text, and one that supports my own argument.

In chapter three I questioned the epistemic capacity of blamers to truly know, with certainty, whether a given wrongdoer is fully culpable. One interpretation of the Confucian view regarding culpability is that we ought to, whenever possible, give others the benefit of the doubt. When wrongdoing occurs, our first response should be no-response. Instead of responding with blame immediately, the suggestion is to simply pause. Blamers are invited to investigate the nature of the wrongdoing and reflect on it, before taking any external action. The strategy is to initially avoid finding fault on either side, and to see what arises from deep reflection about the nature of the wrong before taking further action.

Then, on Sarkissian's interpretation of the Confucian perspective, if one has truly engaged in this reflection and still deems the wrongdoer culpable, it is permissible to blame her. Sarkissian explains that the Confucian approach of giving others the benefit of the doubt is

A strategy with a limited shelf life; the cognitively demanding act of staving off blame and resentment can be expected to last only so long. The injunction to give others the benefit of the doubt...will prove effective only when others provide evidence of the transitory or contingent nature of their initial disagreeable behavior.¹⁸⁸

In this passage, Sarkissian argues that the strategy of giving others the benefit of the doubt only last so long. There must be evidence of the wrong behavior being explainable by transitory or contingent circumstances. If nothing is found to exculpate the agent, then it is permissible to blame her. On Sarkissian's interpretation, the Confucian view does not encourage giving other agents the benefit of the doubt in an open-ended, enduring fashion. Rather, the benefit of the doubt strategy is a good *initial* strategy, with a limited shelf life. But then, without evidence to the contrary, expressing blaming is permissible.

Indeed, this interpretation of the Confucian view deems expressing blame appropriate if the wronged agent (blamer) decides, after thorough reflection on the given case, that he is not to blame for the action and that the culpability in fact lies at least partially with the other person.¹⁸⁹ Sarkissian explains, "We find an analogue to the Tit for Two Tats strategy: Pause not once but twice before retaliating with a defection—in this case, with a negative character assessment."¹⁹⁰ If the wronged agent has truly paused and reflected on the wrong twice, and both times concludes that the wrongdoer is culpable and worthy of blame, then expressing blame is appropriate. This strategy certainly seems like a reasonable method.

¹⁸⁸ Sarkissian, H. (2015). "When you think it's bad it's worse than you think: Psychological bias and the ethics of negative character assessments" in *The Philosophical Challenge from China*. Brian Bruya (ed.). MIT Press, p. 16.

¹⁸⁹ For example, the blame may lie *solely* with the other person, or the two agents might be mutually blameworthy for the situation.

¹⁹⁰ Sarkissian, p. 16.

In this section, I will consider what it would look like to take Sarkissian's interpretation of the Confucian view to its logical conclusion. Is there a way to both hold wrongdoers responsible for their actions *and* continue to give them the benefit of the doubt? On the face of it, these two elements are at odds with one another. However, careful analysis of Sarkissian's account helps resolve this tension. It is important to note that the Confucian strategy is not a wholly passive response to wrongdoing. It is not endorsing pausing and reflecting in isolation. Instead, it is an active process of seeking out the causal and epistemic factors involved that might explain the wrongdoer's behavior.

Sarkissian argues that on the Confucian interpretation, the wronged agent's immediate impulse should be to see how he might have engendered the bad moral conduct himself: Was he indiscrete or unkind? Did he lack patience or resolve? The first step is to try to come to a proper or complete understanding of what may have caused the person to act as they did.¹⁹¹ Similarly, the wronged agent's first thought should be to reflect on how he might be implicated in the wrong. Has he, in any way, participated in or created the dynamic between himself and the wrongdoer that might have impacted the wrongdoing directly?

Second, if the wronged agent feels confident that he did not contribute in any way to the wrong, then the next step is to investigate any relevant external factors that may have influenced the wrong. Perhaps the wrongdoer is tired. Perhaps she is having a bad day, is in pain, or is otherwise preoccupied. These external factors include not only being tired or distracted, but also include the impact that other peoples' actions and moods have on our own actions. Indeed, the Confucian view is profoundly sensitive to the impact we have on one another and the ways in which our interpersonal dynamics inform our own behaviors. This is not to say any of these

¹⁹¹ Sarkissian, p. 16.

external factors are necessarily *mitigating* factors, only that they might be relevant to understanding the nature of the wrongdoer's actions, and therefore should be thoroughly investigated.

Third, Sarkissian explains that the wronged agent is encouraged to consider whether his emotions appropriately signal the severity of the wrong. Are his emotions excessive, given the situation? Are they deficient, given the situation? The wronged agent should reflect on ways in which his own emotional responses to the wrongdoing might be interfering with his ability to have a clear and rational view of the case. Is his anger blocking his ability to recognize that the wrongdoer is not (fully) at fault? Is his degree of anger, indignation, or resentment appropriate for the situation? There are a variety of mitigating circumstances that might be revealed through this reflection.

It might be the case that the wrongdoing triggered an emotional response in the wronged agent that has nothing to do with the wrong itself, and instead has to do with the wronged agent's unique psychological wiring or past painful experiences. Further, the wronged agent should consider whether his emotions are reliable indicators of the wrongdoing. It could be the case that his current emotional state is colored by his own psychological wiring. He should consider whether he is currently reacting in a way that he will feel good about in the future. On the Confucian view, these are all questions wronged agents should consider and reflect upon seriously before deciding whether to express blame.

Still, it can be hard to make sense of this seemingly passive approach in the face of serious moral wrongdoing.¹⁹² For example, In *Hard Feelings* (2013), Macalester Bell suggests that while

¹⁹² In a similar vein, writer Rebecca Solnit argues, "Sometimes there are good reasons for a strong response, including the prevention of further harm. But more often lashing out is a way to avoid looking inward...anger may make people miserable, but it also makes them more confident and obliterates other, more introspective miseries: pain,

we generally have good reasons to be civil towards one another, not everything deserves a civil response.¹⁹³ Bell argues that to respond civilly to persons who publicly express superiority “is to risk condoning the superbia they express, thereby further damaging moral relations.”¹⁹⁴ Bell argues that contempt plays a vital role in response to wrongdoing insofar as it helps put wrongdoers in a position of reflection. For example, to respond with contempt to someone who just uttered a racist remark serves to helpfully allow that person to step back and reflect on her racist attitudes. Contempt, Bell argues, is the best way to object to this person’s racist attitudes and to encourage their self-reflection. Responding with civility, Bell claims, might not enable such an outcome.¹⁹⁵

Like Bell, I recognize that harsh emotional responses like contempt are natural responses to being wronged, and I do not deny that they serve as helpful signals in certain contexts, as in the case of the racist utterance. However, as Rebecca Solnit argues, strong negative emotions like contempt and anger can sometimes *mask* the deeper emotional internal work we each need to do in the aftermath of moral wrongdoing. Put simply, we would rather be mad than sad. But, as we have already seen, being mad and expressing it through our reactive attitudes has a negative, and sometimes irreversible, impact on others. So, the message we ought to gather from both Sarkissian and Solnit is to introspect about the various internal and external factors that might be involved in both the wrongdoer’s actions, and in the blamer’s response to his actions. Again, this is not a passive process. Rather, it is an active, dynamic interrogation of the situation and one’s

fear, guilt, uncertainty, vulnerability. We’d rather be mad than sad.” Solnit, Rebecca. “Facing the Furies.” *Harper’s Magazine*, 1 May 2017. <https://harpers.org/archive/2017/05/facing-the-furies>.

¹⁹³ Bell, Macalester (2015). *Hard Feelings*. Oxford University Press, p. 14.

¹⁹⁴ Bell, p. 219.

¹⁹⁵ For similar defenses of the harsh emotions (anger, contempt, etc.) see also Jaggar (1989) and Spelman (1989).

involvement in it in order to gain a better understanding of the situation and to respond effectively.

Though I have not discussed contempt yet in this dissertation, Bell's work is highly relevant as a critique to the seemingly passive Confucian view just proposed. Furthermore, contempt strikes me as *worse* to express than blame. Bell defines contempt as having four distinguishing characteristics: 1. A judgment concerning the status of the object of contempt, 2. An affectively unpleasant way of regarding or attending to the object of contempt, 3. A comparison to the object of one's contempt, resulting in a positive self-feeling as compared to the object of one's contempt, 4. Psychological withdrawal or distance regarding the object of one's contempt.¹⁹⁶ As this definition demonstrates, contempt involves a complex, psychologically-deep scorn of another person, as well as a withdrawal from them. Though expressed blame *can* involve this sort of deep and intentional scorn or withdrawal, it need not to. In fact, sometimes in expressing (especially angry forms of) blame, we engage in the exact opposite of withdrawal insofar as we get right up in the wrongdoer's face and intimately convey to them our displeasure with their behavior. So, contempt comes with many of the social risks (escalation, spiraling) I have argued that expressed blame has, and likely additional and more troubling risks, too.

In "Virtuous Contempt in the *Analects*", Sarkissian argues that if an agent has properly concluded that another agent is truly contemptible, then she ought to publicly *signal* her contempt for the wrongdoer. Sarkissian strongly argues that "a disposition to despise is a feature of morally exemplary persons in the *Analects*."¹⁹⁷ A major benefit of publicly signaling one's contempt for a contemptible person stems from the motivation to protect other,

¹⁹⁶ Bell, Macalester (2005). "A Woman's Scorn: Toward a Feminist Defense of Contempt as a Moral Emotion". *Hypatia* Vol. 20: 4, pp. 83-4.

¹⁹⁷ Sarkissian, Hagop (2017). "Virtuous Contempt in the *Analects*". *The Oxford Handbook of Chinese Philosophy*, ed. Justin Tiwald. Forthcoming, p. 12.

innocent agents who might be duped by the wrongdoer in the future. Sarkissian further explains that “part of what it means to be a morally exemplary person is to feel and display genuinely negative, aversive feelings toward a range of appropriate targets.”¹⁹⁸

It is important to keep in mind the deeply social framing of the *Analects*. The context is less related to the interpersonal context I am discussing in this dissertation, and more of a broader social concern. From this broader social vantage point, Sarkissian argues that we owe it to each other to publicly express contempt when someone wrongs us, in order to protect our fellow community members. For example, if I feel that my coworker is a truly contemptible person and that I have done a thorough and fair investigation in order to come to this conclusion, and I am worried that if others do not know then they too will be victimized, it would be a moral failing of mine not to signal contempt.¹⁹⁹ On this view, if one feels contempt in one’s heart about a given agent’s actions, one should not hold back that contempt for moral reasons, because one needs to signal to and warn their shared community that the wrong has occurred.

This is clearly demonstrated in a passage in the *Mencius*, another classical Confucian text expanding upon Confucius’s teachings:

Suppose someone were to be harsh in their treatment of me. A junzi (noble person) would, in such a case, invariably examine himself, thinking “I wasn’t benevolent; I lacked propriety. How else could such a thing have come about?” But if, after examining himself, he discovers he had been benevolent, he had acted with propriety, and yet the person still treats him harshly, then the junzi will again invariably examine himself, thinking “I must have lacked commitment.” But if he discovers that he was, in fact, committed, and the person still treats him harshly, only then would the junzi say, “I suppose he is the incorrigible one.”²⁰⁰

In the above passage, the noble person has been a victim to disagreeable conduct directed his way. As Sarkissian comments on the case,

¹⁹⁸ Sarkissian, p. 12.

¹⁹⁹ I’ve paraphrased this example from email correspondence with Sarkissian on the topic of contempt in the *Analects*.

²⁰⁰ Sarkissian, p. 16.

His first impulse is to see how he might have engendered the conduct himself: Was he indiscrete or unkind? Did he lack patience or resolve? Here he is merely trying to come to a proper or complete understanding of what may have caused the person to act in such a fashion.

Only after arriving at a more definite understanding of the situation is the noble person content and justified in blaming the other person. Again, this Confucian interpretation encourages agents to freely express blame to wrongdoers, but only after the situation has been thoroughly investigated and the blamer feels that she has a complete understanding of the case and the factors involved. Sarkissian's recommendation is thus seemingly in tension with my argument that it is preferable not to express blame even after the time of reflection has made one's culpability clear. However, as I will argue in the next section, we can maintain most of Sarkissian's interpretation of Confucianism without the moral tool of expressed blame, and further, can still meet the conditions of answerability.

B. Confucian Answerability

I propose that the Confucian recommendation to pause, reflect, and investigate, in tandem with inquisitive moral dialogue, meets the requirements of answerability, and can do so without expressed blame or expressed contempt. By adding the strategy of inquisitive moral dialogue to the Confucian strategy of pause, reflect, and investigate, I present a more nuanced account of what taking the Confucian recommendation might look like, practically speaking, in our interpersonal moral lives. I have termed this strategy Confucian Answerability, and argue that it is a comprehensive way to achieve the demands of answerability in response to moral wrongdoing without expressed blame.

The first reason for this is that, like the Emotional Conversion technique, the Confucian Answerability strategy requires an awareness of one's own emotional state and the practical need

to temper hot and reactive emotions. Part of this initial non-reactive response includes recognizing the ways in which we often try to use our strong, negative emotions as self-protective defense mechanisms, rather than as morally efficient communications. This desire is understandable, yet not all that useful. Prioritizing the need to enter into an emotionally-cooler affective state when in dialogue about a moral conflict increases the chances of said dialogue between the two agents functioning more effectively.

Second, the Confucian Answerability strategy embraces the highly social-nature of our moral lives. This insight encourages a more nuanced and context-specific understanding of cases of moral wrongdoing. Focusing on the deeply social and interconnected nature of our moral lives, we can see the complex relationships and interpersonal dynamics involved in each case of wrongdoing. In addition, initially suspending our judgments of blame allows for a more critical epistemic inquiry into the nature of the wrongdoing. I take it as given that most of us care about and value our friendships, and do not want to publicly blame those close to us without epistemic warrant.²⁰¹

Third, the Confucian Answerability strategy encourages blamers to direct their reflection in the aftermath of wrongdoing upon themselves first. On this view, blamers are prompted to introspect about whether they themselves have been culpable in any way for the wrong that has just occurred, rather than merely considering the ways in which the wrongdoer is solely culpable. This deep self-reflection reduces the social harm of expressed blame because the blamer is coming to the process of moral inquisition from a place of self-awareness and recognition of her own potential involvement in the wrong.

²⁰¹ For discussion of epistemic partiality in friendship, see Stroud (2006) and Keller (2007).

Finally, the Confucian Answerability strategy is highly compatible with both of the strategies I surveyed above. After first pausing, investigating, and reflecting, the blamer is free to pursue either of those strategies as well as expressing blame, if deemed appropriate. For example, the blamer can begin an inquisitive dialogue with the wrongdoer about the nature of the wrongdoing. The blamer can, and should, still demand reasons for action from the wrongdoer. After each agent has taken the time to independently investigate and reflect on the nature of the wrongdoing, that period of reflection can strengthen the efficacy of their shared dialogical exchange. After each agent has thoroughly reflected and introspected, then productive and inquisitive moral dialogue, grounded in curiosity and compassion, can more organically occur. If the introspection and reflection does not amount rendering the wrongdoer non-culpable, then the onus remains on the wrongdoer to justify her reasons for action.

The Confucian Answerability strategy can be utilized in a wide array of interpersonal moral wrongdoing scenarios, and in conjunction with a variety of other strategies. For this reason, it is a unifying and comprehensive strategy of response to moral wrongdoing, and one that can demand justification from wrongdoers without expressed blame.

C. When Immediate Response is Required

In response to the reflective, non-urgent strategy of Confucian Answerability, one might argue that sometimes an immediate response to moral wrongdoing is required. There is good reason to think that urgency is sometimes required in responding to moral wrongdoing and that the Confucian Answerability strategy is ill-equipped to meet these demands. Perhaps these sorts of cases require special consideration. If the blamer waits to see if her emotions cool off over time, the time for moral reform and dialogue with the wrongdoer might have passed. It might be

too late. If the goal of blaming others is to promote reform and moral correction, time may be of the essence. If one waits too long to express blame, it might not be as meaningful and motivating to the wrongdoer.

This is a legitimate worry. However, moral reform is not exclusively available in the immediate aftermath of wrongdoing, and thinking that urgency is always required is a needlessly limited view. Moral wrongdoing does not always require an immediate response. We might be tempted to think that moral course correction is most likely to happen in the direct aftermath of wrongdoing because emotions are so heated and seemingly primed for action and moral response. However, we can also motivate wrongdoers to make positive moral changes in the days, if not months, after the wrongdoing. It is incorrectly limiting to think that reform must immediately follow bad moral action. Wrongdoers understand the wrongness of their actions at different times and in different contexts. Sometimes it takes an unbiased external party to get the moral point across. Sometimes it takes additional life experience to motivate wrongdoers to change.

Related to this worry about urgency is the distinction between immediate and settled emotions. In the direct aftermath of wrongdoing, our emotions might not be authentic indicators of our emotional response to the wrongdoing. After a few days of reflection, our heated emotions about the wrongdoing might be much weaker or cooler, or perhaps even resolved. At best, our emotions correctly capture how we feel about the situation at that exact moment. But who is to say that those emotions will persist over time? There is an important distinction between our hasty, immediate emotional responses to being wronged, and our more considered, settled emotional responses to being wronged.

Joseph Butler makes a similar distinction in the case of resentment between settled (or deliberate) resentment and hasty (or sudden) resentment. Hasty resentment, Butler argues, is an immediate response to “mere harm without appearance of wrong or injustice.”²⁰² He argues that hasty resentment can be excessive or misguided. It might not be informed of all the relevant facts. Differently, settled resentment takes injustice as its object. It has a broader scope than hasty resentment. Settled resentment aims at pursuing justice and is not excessive or misguided.²⁰³

In a similar vein, our immediate blaming emotions risk being excessive or misguided. Because our immediate blaming emotions can often be misguided, there is no sense in hastily expressing a misguided view. In the initial aftermath of wrongdoing, our emotions might be overly focused on the injury done to oneself, rather than on correcting the injustice more generally. There might be more the blamer needs to learn about the situation, and her initial emotional responses likely lack this information.

In light of this, I argue that we ought to restrain or control our negative emotions in the direct aftermath of a moral conflict to avoid an immediate response dictated by our baser, hostile emotions. Indeed, this is the same conclusion I presented with the Emotional Conversion technique, which argues that responding to moral wrongdoing with cooler emotions like sadness is preferable to responding to moral wrongdoing with heated emotions like anger. Assuming we want to express blame to wrongdoers in a way that accurately depicts how we feel about their moral actions over time, our more thought-out, considered emotions and attitudes are the blaming emotions we want to express, if we decide to express blame at all. Attention to this potential shift in emotional intensity is important in our discussion of blaming emotions, because

²⁰² Butler, J. (1726). *Fifteen Sermons Preached at the Rolls Chapel*. London: J. and J. Knapton, Sermon VIII.

²⁰³ See <https://plato.stanford.edu/entries/butler-moral/> for a more detailed account of Butler’s distinction between hasty and settled resentment.

what the ‘blaming emotions’ refer to might naturally shift (and lessen in intensity) over time. With this in mind, taking a pause before responding, in order to respond authentically, may (at least in some cases) be more important than any felt urgency about responding.

V. The Moral Failure to Not Express Blame

Even if I am right to think that we can hold others responsible without expressed blame, there is an additional remaining worry. That worry is that even if we *can* hold others accountable without expressed blame, the act of blaming itself is necessary to avoid condoning the wrongs that are going unblamed. Some may further argue that it is a *moral failure* to not express blame publicly when wrongdoing occurs, even if there are alternative methods of holding wrongdoers responsible. In this final section of the chapter, I will consider and reject arguments of this kind to demonstrate that my view does not imply or assume that we should condone the wrongs that are going unblamed.

One form of this argument for the necessity of blame is that blamers *owe it* to their wrongdoers to publicly express blame to them. Another form of this argument is that the process of holding others accountable is *incomplete* without the public charge of blame. Moreover, some argue that expressed blame serves as a statement of respect, in two ways. First, by expressing blame, the blamer is demonstrating that she respects herself and refuses to be treated wrongly by the wrongdoer or by anyone else. Second, by expressing blame, the blamer is indicating her respect for the wrongdoer as an agent in our shared moral community who should be held accountable for her wrongdoing. The blamer respects the wrongdoer as a moral agent who can, and should, act in morally better ways. Without the moral tool of expressed blame, the worry is

that wronged agents (i.e. blamers) have no way to assert their rights, and thus are essentially condoning the wrongs that are going unblamed.

A. The Moral Duty to Express Blame

Several authors²⁰⁴ argue for some version of the view that we have a moral duty to express blame when wrongdoing occurs. For example, Angela Smith proposes a view of blame as *moral protest*. Protest, in this context, refers to an “outward statement of disapproval, whose explicit aim is to draw public attention to some serious moral wrong.”²⁰⁵ Smith argues that to morally blame another agent is to “register in some significant way one’s moral protest of that agent’s treatment of oneself or others.”²⁰⁶ According to Smith, when someone morally wrongs us, we protest their actions through expressing blame, thereby “*challenging* the moral claim implicit in the wrongdoer’s action.”²⁰⁷ Moreover, Smith proposes that blame expressed as protest aims to prompt moral recognition and acknowledgment not only from the wrongdoer, but also from others in the community.²⁰⁸ Smith further argues that the ability to express blame as protest is necessary and sufficient for being a morally responsible agent.²⁰⁹ So for Smith, publicly expressed blame serves the vital function of protesting wrongdoers’ behaviors to challenge their actions and prompt acknowledgment both from the wrongdoer and from the greater moral community.

John Kekes makes a related argument in defense of blame in his article “Blame Versus Forgiveness” (2009). Kekes argues that forgiveness is not an appropriate response to

²⁰⁴ Bennett (2008); Hieronymi (2001); Smith (2013); Kekes (2009); Boxill (1976)

²⁰⁵ Smith, Angela (2013). “Moral Blame and Moral Protest”. in *Blame: Its Nature and Norms*, (eds.) D. Justin Coates and Neal A. Tognazzini. Oxford University Press.

²⁰⁶ Smith, p. 1.

²⁰⁷ Smith, p. 22.

²⁰⁸ Smith, p. 22.

²⁰⁹ Smith, p. 22.

wrongdoing. Forgiveness, in this context, can be thought of as the failure to blame.²¹⁰ For Kekes, forgiveness is not an appropriate response to wrongdoing because wrongdoers *should be blamed continuously*, and forgiveness is not compatible with that aim. According to Kekes, a failure to express blame is inconceivable as a response to moral wrongdoing. Kekes explains

Forgiving people have a very low opinion of themselves. They find it natural to be abused because they doubt that they are worthy of anything else. They lack self-respect...In refusing to blame wrongdoers who have inflicted moral injury on them, habitually forgiving agents are colluding in the violation of moral requirements.²¹¹

In this passage, Kekes argues that forgiveness is not only a vice, but, more strongly, that forgiving people fundamentally lack self-respect.²¹² Kekes argues that we must express blame to wrongdoers, and that a failure to do so is a violation of our moral duty. On Kekes's view, to not find fault with wrongdoers' behaviors and publicly express it to them is a failure to respect them as moral agents. Worse, this failure does not give wrongdoers the motivation to act in morally better ways in the future.

Smith's and Kekes's arguments each defend the important role public expressions of blame play in our moral lives. Consider the following example. Say that my sister borrows my car without asking and then gets into an accident and totals my car. When I find out, I express blame towards her. Part of my expression of blame involves me valuing of my rights and my property (my car). My sister stole something of mine without asking. When I express blame to her, I challenge her actions and signal to her (and our shared moral community) that I will not

²¹⁰ Though deciding not to blame someone and forgiving them are different concepts, they are closely related, so for the sake of this argument I will use them interchangeably. When we forgive someone, we no longer blame them. We have decided to let go of the hard feelings that we have had towards them over their wrongdoing, and make amends. Alternatively, perhaps we never blamed them to begin with, and we were able to forgive them right away.

²¹¹ John Kekes (2009). "Blame Versus Forgiveness," *The Monist*, vol. 92, no. 4, p. 493.

²¹² In "Servility and Self-Respect", Thomas Hill makes a related argument about self-respect. Hill argues that a *servile person* is one who tends to deny or disavow his own moral rights because he does not understand them or has little concern for the status they give him. For Hill, a failure to understand and acknowledge one's own rights is a moral defect, equivalent to accepting servility.

tolerate that kind of behavior. If I fail to express blame to her, this might give her the idea that she can steal any of my things without asking again in the future. She might even start stealing things from others, too. Her morally bad behavior could escalate and expand quickly. If I fail to make it clear to her immediately that her behavior is morally unacceptable, perhaps I am not asserting or valuing my own rights. Further, perhaps my failure to publicly blame her is a sign that I do not respect her as a moral agent capable of acting better.

B. Response to the Moral Duty to Express Blame Objection

I have two responses to the general argument that we have a moral duty to express blame to wrongdoers. Before those responses, however, I would like to make clear that my dissertation has largely assumed that our interpersonal relationships are *valuable* and *worth fostering*, even in the face of serious moral conflict. This is a value ranking that not all authors on this subject share. There seems to be a very strong Kantian attitude of individualism and autonomy operating in the background of arguments about our moral duty to express blame. It might be the case that some authors, such as Kekes, do not fundamentally peg any of the value of one's life onto the value of one's relationships. As such, if you value the strong importance of individualism, my view throughout this dissertation might be less compelling. If one does not value the maintenance and growth of our interpersonal relationships, it might seem less urgent to not rupture those relationships with expressed blame. Beyond this basic distinction, I have two additional objections to the view that we have a moral duty to express blame.

i. Manifestations of Respect

Many advocates of blame²¹³ cite expressed blame as a key way to show the wrongdoer that the blamer respects her and her capacity to act in morally better ways next time. Even if blame-alternatives exist, some philosophers maintain that the very act of blaming itself is a way to demand respect, and for this reason, blame *must* remain a central feature of our moral lives. Respect can certainly be made manifest through expressed blame, but respect also has many other outlets. For example, we can respect wrongdoers and their capacities as moral agents to act in morally better ways by initiating the type of compassionate, curiosity-based, inquisitive dialogue I have presented throughout this chapter. This inquisitive dialogue creates space for learning more about how the wrongdoer understands her own wrong actions. For example, the wrongdoer might already feel remorse, guilt, or shame. She might already be involved in thoughtful, moral self-correction. In these cases, expressed blame is likely superfluous. This is especially true for wrongdoers with a high degree of emotional sensitivity and self-awareness. These agents already understand what about their behavior needs correcting. They do not need further moral instructions from the blamer.

Moreover, respect can sometimes involve giving a wrongdoer the benefit of the doubt. When someone wrongs us, we can keep their actions in mind, but do not necessarily need to express blame to them in this particular instance. We can forgive an agent for their wrongdoing right away, and doing so is consistent with maintaining respect for them. Certainly, this is not to say that their past wrong actions will not alter the ways we interact with them in the future. For example, we might keep a mental note of their actions as an instructive guide for how we interact with them in the future. But this does not require that we publicly express blame to them.

²¹³ In addition to Smith and Kekes, see also work on this topic by Christopher Bennett and Antony Duff.

We do not need to so diligently express blame about others' moral transgressions. This constant effort is exhausting and impractical. Additionally, we may choose not to express blame to a given wrongdoer simply because *we may not care that much* about the harm they caused us. We may be able to see that in the greater scheme of things, their moral wrong in this instance is not worth getting worked up over or even commenting on. As these considerations reveal, failing to express blame in response to a case of moral wrongdoing is not necessarily a violation of our moral duty, nor does it mean that one has a fundamental disrespect for our shared moral norms.

ii. Dignity

Relatedly, we do not need to constantly assert or demand self-respect from others in order to feel a deep-seated sense of self-respect. Having self-respect is at least partially an *internal practice*, and one that need not be continuously exhibited in a public setting in order to be authentic. This strikes me to be precisely what dignity is. Although dignity is traditionally defined as “a sense of pride in oneself” or simply “self-respect”²¹⁴, it connotes something over and above self-respect. Rather than needing to publicly declare that someone has self-respect, one who has dignity feels grounded in their self-worth regardless of their moral treatment by wrongdoers. Of course, there are times in which public affirmation of respect is symbolically important, and there is a real need to publicly assert one's rights. Extreme and systemic cases of wrongdoing certainly call for public response. Indeed, there is nothing theoretically wrong with wanting to continuously affirm and announce one's self-respect in a public setting.

However, it is not necessary for cultivating and maintaining self-respect. Failing to blame others does not pigeonhole one into the role of a doormat or lacking self-respect. As such,

²¹⁴ <https://www.merriam-webster.com/dictionary/dignity>.

responding to moral wrongdoing in a calm and kindly fashion, without expressed blame, does not imply that one 1. Condone the wrongs that are going unblamed, 2. Lacks self-respect, or 3. Disrespects the wrongdoer or their shared moral norms.

VII. Conclusion

In this chapter, I presented three methods of responding to moral wrongdoing that each depart from our traditional Western, hostile blaming practices. Agreeing with Smith, I argued that answerability, defined as a requiring reasons (in the normative sense) from the wrongdoer, is the leading conception of what it means to hold wrongdoers morally responsible. Given that answerability is the correct conception of moral responsibility, I further argued that we can go in for justificatory challenge without expressed blame, and that it is *preferable* to do so without expressed blame. That is, expressed blame is not required to meet the conditions of answerability, and, more strongly, that we have reason to think that the interaction between wrongdoer and blamer will go better without expressed blame.

In my discussion of the various alternative methods to respond to moral wrongdoing, I have drawn upon the work of Lacey and Pickard, Blustein, and Sarkissian, as well as some empirical findings on the ways in which certain emotions have the capacity to blunt other emotions from arising. I have not claimed that any of the alternative methods I have presented will always serve the exact function of blame. Further, it is obviously not possible to hold every wrongdoer responsible in our interpersonal lives. For some things, it is best if we just let go. It is inaccurate to think that we can, or should, try to hold every wrongdoer responsible for their wrongs, big and small. Attempting this would likely involve the majority of our days spent regulating others'

behavior. This is not consistent with the demands and joys of our lives. Instead, we are putting less strain on our interpersonal relationships by steering away from our daily blaming practices. Assuming that preserving and growing our interpersonal relationships is valuable, I have argued that this is a preferable outcome to expressing blame.

Throughout the dissertation, I made use of the idea that even if blaming a wrongdoer is *fitting* for a given situation, it might not be appropriate to express. Because of the higher stakes and social risks that the hostile and retributive blaming emotions present, it might not be *appropriate* to publicly express blame, even when expressing blame fits a given moral situation. I am confident that we have the philosophical tools necessary to resolve moral conflicts in our interpersonal daily lives without any manner of expressing blame. My hope with this project, more generally, is to have expanded the range of useful responses to interpersonal moral wrongdoing that we can take. Put simply, I have aimed to shift our default interpersonal strategies of moral response to wrongdoing away from expressed blame and towards alternative practices.

Bibliography

- Anderson, N. H. (1965). "Averaging versus adding as a stimulus-combination rule in impression formation". *Journal of Personality and Social Psychology*: 2, pp. 1-9.
- Bagley, Benjamin (2017). "Properly Proleptic Blame". *Ethics* 127 (4): 852-882.
- Bakhurst, David, Brad Hooker, and Margaret Olivia Little (eds.) (2013). *Thinking About Reasons: Themes from the Philosophy of Jonathan Dancy*. Oxford: Oxford University Press.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5, pp. 323-370.
- Bell, Macalester (2005). "A Woman's Scorn: Toward a Feminist Defense of Contempt as a Moral Emotion". *Hypatia* Vol. 20: 4, pp. 80-93.
- Bennett, Christopher (2008). *The Apology Ritual*. Cambridge, UK: Cambridge University Press.
- Bennett, Christopher (2017). "The Expressive Function of Blame" in *Blame: Its Nature and Norms*, (eds.) D. Justin Coates and Neal A. Tognazzini. Oxford University Press.
- Benson, Paul (2001). "Culture and Responsibility: A Reply to Moody-Adams". *Journal of Social Philosophy* vol. 32: 610-620.
- Bittner, Rudiger (1992). "Is it Reasonable to Regret Things One Did?" *The Journal of Philosophy*, 89:5, pp. 262-273.
- Blustein, Jeffrey (2017). "Forgiveness and The Moral Psychology of Sadness" in: Anna Gotlib (ed.). *The Moral Psychology of Sadness*. Rowman & Littlefield International: London.
- Blustein, Jeffrey (2018). "Should We Get Rid of Blame?" Unpublished manuscript.
- Bodenhausen, Galen, Lori Sheppard, and Geoffrey Kramer (1994). "Negative affect and social judgment: the differential impact of anger and sadness," *European Journal of Social Psychology*, vol. 24, pp. 45-62.
- Bommarito, Nic. (2011). "Bile & Bodhisattvas: Santideva on Justified Anger". *Ethics*, vol. 18: 386-481.
- Bommarito, Nic (2017). *Inner Virtue*. Oxford University Press.
- Bommarito, Nic. (2014). "Patience and Perspective." *Philosophy East and West*, vol. 64 no. 2, pp. 269-286.
- Bommarito, Nic (2017). "Virtuous and Vicious Anger". *Journal of Ethics and Social Philosophy*

11:3, p. 374-5.

Bonanno, George (2009). *The Other Side of Sadness: What the New Science of Bereavement Tells Us About Life After Loss*. Hatchett Book Group.

Boxill, Bernard R. (1976). "Self-Respect and Protest". *Philosophy & Public Affairs* 6 (1): 58-69.

Butler, J. (1726). *Fifteen Sermons Preached at the Rolls Chapel*. London: J. and J. Knapton.

Carbonell, Vanessa (2009). "What Moral Saints Look Like". *Canadian Journal of Philosophy* 371 vol. 39(3), pp. 371-398.

Chapman, Benjamin, Kevin Fiscella, Ichiro Kawachi, Paul Duberstein, and Peter Muennig (2013). Emotion suppression and mortality risk over a 12-year follow-up. *Journal of Psychosomatic Research*: 75 (4), p. 381-385.

Cherry, Myisha (2012). "Moral Anger, Motivation, and Productivity". Presented at New Jersey Regional Philosophical Association Conference, Bergen Community College.

Choi, James; Laibson, David; Madrian, Brigitte C. and Metrick, Andrew (2003). "Defined Contribution in Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance," in James M. Poterba (ed.), *Tax policy and the economy*, Vol. 16. Cambridge, MA: MIT Press, pp. 67-113.

Cozort, Daniel (1995). "Cutting the Roots of Virtue: Tsongkhapa on the Results of Anger," *Journal of Buddhist Ethics* vol. 2, pp. 83-104.

D'Arms, Justin and Daniel Jacobsen (2000). "The Moralistic Fallacy: On the 'Appropriateness' of Emotions". *Philosophy and Phenomenological Research*, Vol. 61:1, pp. 65-90.

Darwall, Stephen (2004). "Respect and the Second Person Standpoint". *Proceedings and Addresses of the American Philosophical Association*, Vol. 78, No. 2.

Darwall, Stephen (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

Edelglass, William, and Jay L. Garfield (2009). *Buddhist Philosophy: Essential Readings*. Oxford University Press.

Ellsworth, Phoebe and Craig Smith (1998). "From Appraisal to Emotion: Differences Among Unpleasant Feelings," *Motivation and Emotion*, vol. 12, no. 3, pp. 271-302.

Feldman Barrett, Lisa (2017). "The Theory of Constructed Emotion: An Active Inference Account of Interoception and Categorization." *Social Cognitive and Affective Neuroscience*.

- FitzPatrick, William J (2008). "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge." *Ethics*, vol. 118, no. 4, pp. 589–613.
- Flanagan, Owen J. (2011). *The Bodhisattva's Brain: Buddhism Naturalized*. Cambridge, MA: MIT.
- Flanagan, Owen (2016). *The Geography of Morals: Varieties of Moral Possibility*. Oxford University Press.
- Flynn, J. J., Hollenstein, T., & Mackey, A. (2010). The effect of suppressing and not accepting emotions on depressive symptoms: Is suppression different for men and women? *Personality and Individual Differences*, 49(6), 582-586.
- Fowler, James H. and Cindy D. Kam (2006). "Patience as a Political Virtue: Delayed Gratification and Turnout". *Political Behavior*, Vol. 28, No. 2: 113-128.
- Franklin, Christopher Evan (2013). "Valuing Blame." in *Blame: Its Nature and Norms*, eds. D. Justin Coates and Neal A. Tognazzini. New York: Oxford University Press, pp. 207-223.
- Fricker, Miranda (2013). "Styles of Moral Relativism—A Critical Family Tree". In: Crisp, R., (ed.) *Oxford Handbook of the History of Ethics*. Oxford Handbooks in Philosophy. Oxford University Press: Oxford, pp. 793-817.
- Fricker, Miranda (2010). "The Relativism of Blame and Williams' Relativism of Distance". *Proceedings of the Aristotelian Society Supp.* Vol. (July).
- Fricker, Miranda. (2016). "What's the Point of Blame? A Paradigm Based Explanation". *Noûs*, 50, pp. 165–183.
- Friedman, Marilyn (2013). "How to Blame People Responsibly". *The Journal of Value Inquiry* 47, pp. 271-284.
- Friedman, Ray, Cameron Anderson, Jeanne Brett, Mara Olekalns, and Nathan Goates (2004). "The Positive and Negative Effects of Anger on Dispute Resolution: Evidence From Electronically Mediated Disputes" *Journal of Applied Psychology*: 89 (2), pp. 369-376.
- Gilbert, Margaret (2006). "Who's to Blame? Collective Moral Responsibility and its Implications for Group Members." *Midwest Studies in Philosophy*, vol. 30 (1), pp.94-114.
- Gilovich, Thomas; Griffin, Dale, and Kahneman, Daniel (2002), eds. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, U.K.: Cambridge University Press.
- Ginet, Carl (2000). "The Epistemic Requirements for Moral Responsibility". *Nous* 34: 267-277.

- Goldman, David (2014). "Modification of the Reactive Attitudes". *Pacific Philosophical Quarterly* 95(1), pp. 1-22.
- Goodman, Charles (2002). "Resentment and Reality: Buddhism on Moral Responsibility". *American Philosophical Quarterly*, vol. 39: 4, pp. 359-372.
- Gray, Kurt and Daniel Wegner (2011). "Dimensions of Moral Emotions". *Emotion Review*: 3(3), pp. 258-260.
- Greene, J. & Cohen, J. (2004). "For the Law, Neuroscience Changes Nothing and Everything". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*: 359, pp. 1775-1785.
- Haidt, Jonathan (2001). "The Moral Emotions." In *Handbook of Affective Sciences*, ed. Richard Davidson, Klaus Scherer, and H. Hill Goldsmith. Oxford: Oxford University Press, pp. 852-870.
- Hampton, J. (1984). "The Moral Education Theory of Punishment". *Philosophy & Public Affairs*, 13(3), 208-238.
- Harman, Elizabeth (2011). "Does Moral Ignorance Exculpate?" *Ratio*, vol. 24, pp. 443-468.
- Hieronymi, Pamela (2001). "Articulating an Uncompromising Forgiveness". *Philosophy and Phenomenological Research*, vol. 62:3, pp. 529-555.
- Hieronymi, Pamela (2007). "Rational Capacity as a Condition on Blame". *Philosophical Books*, 48(2):109-123.
- Hieronymi, Pamela (2008). "Sher's Defense of Blame." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, vol. 137:1, pp. 19-30.
- Hieronymi, Pamela (2004). "The Force and Fairness of Blame". *Philosophical Perspectives*, 18(1): 115-148.
- Hill, Thomas E. Jr. (1973). "Servility and Self-Respect". *The Monist*, vol. 57(1): 87-104.
- Ivanhoe, Philip J. (2000). *Confucian Moral Self Cultivation*. Hackett Publishing: Indianapolis.
- Kahneman, Daniel and Tversky, Amos (2000). eds. *Choices, values, and frames*. Cambridge, U.K.: Cambridge University Press.
- Kawall, Jason (2013). "Friendship and Epistemic Norms". *Philosophical Studies*, vol. 165, pp. 349-370.
- Kekes, John (2009). "Blame Versus Forgiveness," *The Monist*, vol. 92, no. 4, pp. 488-506.

- Keller, Simon (2007). *The Limits of Loyalty*. Cambridge University Press.
- Kugler, M. B., F. Funk, J. Braun, M. Gollwitzer, A. C. Kay and J. M. Darley. (2013). "Differences in Punitiveness across Three Cultures: A Test of American Exceptionalism in Justice Attitudes." *Journal of Criminal Law & Criminology* 103, pp. 1071–1201.
- Lacey, Nicola and Hanna Pickard (2012). "From the Consulting Room to the Court Room? Taking the Clinical Model of Responsibility Without Blame into the Legal Realm." *Oxford J Legal Studies* vol. 33:1, pp. 1-29.
- Levy, Neil (2014). "Consciousness, Implicit Attitudes and Moral Responsibility". *Noûs*, Vol.48(1), pp.21-40
- Malle, Bertram F., Guglielmo, Steve, and Monroe, Andrew E (2017). "A Theory of Blame". *Psychological Inquiry*, 25, pp. 147-186.
- Mason, Michelle (2011). "Blame: Taking it Seriously". *Philosophy and Phenomenological Research* 83:2.
- McCrae, Emily (2012). "Emotion, Ethics and Choice: Lessons from Tsongkhapa," *Journal of Buddhist Ethics*, 22: 19.
- McCrae, Emily (2015). "Metabolizing Anger: A Tantric Buddhist Solution to the Problem of Moral Anger," *Philosophy East and West*, 65: 3.
- McGeer, Victoria (2010). "Co-reactive attitudes and the making of moral community". In *Emotions, Imagination and Moral Reasoning*, eds., C. MacKenzie and R. Langdon. Macquarie monographs in Cognitive Science. Psychology Press.
- McGeer, Victoria (2013). "Civilizing Blame". in *Blame: Its Nature and Norms*, (eds.) D. Justin Coates and Neal A. Tognazzini. Oxford University Press.
- McGeer, Victoria and Friederike Funk (2017). "Are 'Optimistic' Theories of Criminal Justice Psychologically Feasible? The Probative Case of Civic Republicanism." *Criminal Law and Philosophy*. Vol. 11, pp. 523-544.
- Mele, Alfred (2011). "Moral Responsibility for Actions: Epistemic and Freedom Conditions". *Philosophical Explorations* 13, pp. 101-111.
- Murphy, J. G. (1985). "Retributivism, Moral Education, and the Liberal State." *Criminal Justice Ethics* 4(1): 3–11.
- Murphy, Jeffrie, and Jean Hampton (1988). *Forgiveness and Mercy*. New York: Cambridge University Press.
- Nussbaum, Martha (2016). *Anger and Forgiveness*. Oxford University Press.

- Nussbaum, Martha C. (2016). "Beyond Anger". *Aeon*, accessed 26 July 2016, aeon.co/essays/there-s-no-emotion-we-ought-to-think-harder-about-than-anger.
- Nussbaum, Martha (1994). *The Therapy of Desire*. Princeton University Press.
- Nussbaum, Martha (2015). "Transitional Anger". *Journal of the American Philosophical Association*, pp. 41-56.
- Nussbaum, Martha (2001). *Upheavals of Thought*. Cambridge University Press.
- O'Donoghue, Ted and Rabin, Matthew (2003). "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes". *American Economic Review (Papers and Proceedings)*, 93:2, pp. 186-91.
- Paul, Sarah K. (2017). "Good intentions and the road to hell". *Philosophical Explorations*, pp. 40–54.
- Peeters, G., & Czapinski, J. (1990). "Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects". In W. Stroebe & M. Hewstone (Eds.) *European review of social psychology* (Vol. 1, pp. 33-60). New York: Wiley.
- Pereboom, Derek (1995). "Determinism al Dente". *Noûs* vol. 29:1, pp. 21-45.
- Pettigrove, Glen (2012). *Forgiveness and Love*. Oxford University Press.
- Pettigrove, Glen (2012). "Meekness and 'Moral' Anger," *Ethics*, vol. 122, no. 2, pp. 341-370.
- Pickard, Hanna (2016). "Responsibility without Blame for Addiction". *Neuroethics* (Special Issue on Marc Lewis' *The Biology of Desire: Why Addiction is not a Disease*):1-12.
- Plutchick, Robert (2001). "The Nature of Human Emotions: Human emotions have deep revolutionary roots, a fact that may explain their complexity and provide tools for clinical practice." *American Scientist* vol. 89 (4), pp. 344-350.
- Prinz, Jesse (2010). "The Moral Emotions" in P. Goldie (ed.) *The Oxford Handbook of Philosophy of Emotion*. New York: Oxford.
- Radzik, L. (2011). "On Minding Your Own Business: Differentiating Accountability Relations within the Moral Community", *Social Theory and Practice*, 37, pp. 574–598.
- Rosebury, Brian (1995). "Moral Responsibility and Moral Luck". *The Philosophical Review* 104: 499-524.
- Rosen, Gideon (2003). "Culpability and Ignorance". *Proceedings of the Aristotelian Society* Vol. CIII, Pt. 1: 61-84.

- Rosen, Gideon (2004). "Skepticism about Moral Responsibility". *Philosophical Perspectives*, vol. 8:1, pp. 295–313.
- Rosenberg, Marshall B. (2015). *Nonviolent Communication: A Language of Life*. PuddleDancer Press.
- Śāntideva. (1997). *The Way of the Bodhisattva: A Translation of the Bodhicharyāvātāra*. Boston: Shambhala.
- Sarkissian, Hagop (2017). "Virtuous Contempt in the *Analects*". *The Oxford Handbook of Chinese Philosophy*, (ed.) Justin Tiwald. Forthcoming.
- Sarkissian, Hagop. (2015). "When You Think It's Bad It's Worse Than You Think: Psychological Bias and the Ethics of Negative Character Assessments". in *The Philosophical Challenge from China*, (ed.) Brian Buya. MIT Press.
- Scanlon, T. M. (2008). *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, MA: Harvard University Press.
- Seneca, Annaeus (1900). "Minor Dialogs". Translated by Aubrey Stewart. Bohn's Classical Library Edition (London): George Bell and Sons, pp. 115-161.
- Shapland, Joanna et al. (2006). "Situating Restorative Justice Within Criminal Justice". *Theoretical Criminology* II:505-532.
- Sher, George (2005). *In Praise of Blame*. New York: OUP.
- Sher, George (2006). "Out of Control". *Ethics*: 116 (2), pp. 285-301
- Sher, George (2009). *Who Knew? Responsibility Without Awareness*. New York: OUP.
- Shoemaker, David (2011). "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility". *Ethics* 121 (3): 602-632.
- Skowronski, J. J., & Carlston, D. E. (1989). "Negativity and extremity biases in impression formation: A review of explanation". *Psychological Review*, 105, pp. 131-142.
- Smith, Adam (1759). *The Theory of Moral Sentiments*. Gutenberg Publishers.
- Smith, Angela (2015). "Attitudes, Tracing, and Control". *Journal of Applied Philosophy*, 32 (2):115-132.
- Smith, Angela (2013). "Moral Blame and Moral Protest". in *Blame: Its Nature and Norms*, (eds.) D. Justin Coates and Neal A. Tognazzini. Oxford University Press.
- Smith, Angela (2007). "On Being Responsible and Holding Responsible". *The Journal of Ethics*

11: 367-392.

- Smith, Angela (2005). "Responsibility for Attitudes: Activity and Passivity in Mental Life". *Ethics* 115 (2): 236-271.
- Smith, Holly (1983). "Culpable Ignorance". *The Philosophical Review* 97: 543-571.
- Solnit, Rebecca. "Facing the Furies." *Harper's Magazine*, 1 May 2017.
<https://harpers.org/archive/2017/05/facing-the-furies/4/>.
- Stanley, Jason (2005). *Knowledge and Practical Interests*. Oxford University Press.
- Strawson, P. F. (1962). "Freedom and Resentment". *Proceedings of the British Academy* (48), pp. 1-25.
- Stroud, S. (2006). Epistemic Partiality in Friendship. *Ethics*, 116(3), pp. 498-524.
- Talbert, Matthew (2012). "Moral Competence, Moral Blame, and Protest". *Journal of Ethics*, 16 (1), pp. 89-109.
- Thurman, Robert (2005). *Anger*. Oxford: OUP.
- Tiedens, L. Z. (2001). "Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral." *Journal of Personality and Social Psychology*: 80(86).
- Vargas, Manuel (2013). *Building Better Beings: A Theory of Moral Responsibility*. Oxford University Press.
- Wallace, R. J. (2011). "Internalism about Responsibility". Unpublished manuscript, pp. 9-10.
- Wallace, R. J. (1998) "Responsibility and the Moral Sentiments". Harvard University Press.
- Watson, Gary (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," in F.D. Schoeman (ed.), *Responsibility, Character and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press, pp. 256-286.
- Watson, Gary (2012). "Standing in Judgment". in *Blame: Its Nature and Norms*, (eds.) D. Justin Coates and Neal A. Tognazzini. Oxford University Press.
- Watson, Gary (1996). "Two Faces of Responsibility". in *Agency and Answerability*. Oxford: Oxford University Press, pp. 260-86.
- Williams, Bernard (1989). "Internal Reasons and the Obscurity of Blame". *Logos* 10, pp. 1-11.
- Williams, Bernard (1981). *Moral Luck*, Cambridge: Cambridge University Press.

- Winterich, Karen Page, Han, Seunghee, and Lerner, Jennifer S. (2010). "Now That I'm Sad, It's Hard to Be Mad: The Role of Cognitive Appraisals in Emotional Blunting". *Personality and Social Psychology Bulletin*: 36(11), pp. 1467-1483.
- Wolf, Susan (2011). "Blame Italian Style". in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, (eds.) Wallace, R. Jay, Rahul Kumar and Samuel Freeman. Oxford: Oxford University Press.
- Wong, David (2006). *Natural Moralities*. Oxford University Press.
- Zimmerman, Michael. (1988). *An Essay on Moral Responsibility*, Totowa, NJ: Rowman & Littlefield.
- Zimmerman, Michael. (1997). "Moral Responsibility and Ignorance". *Ethics* 107, pp. 410-426.