

2016-08-05

Moral Injury in Contemporary Ethics: The Application of a Socratic Idea

Heleana Theixos

University of Miami, h.theixos@umiami.edu

Follow this and additional works at: https://scholarlyrepository.miami.edu/oa_dissertations

Recommended Citation

Theixos, Heleana, "Moral Injury in Contemporary Ethics: The Application of a Socratic Idea" (2016). *Open Access Dissertations*. 1734.
https://scholarlyrepository.miami.edu/oa_dissertations/1734

This Open access is brought to you for free and open access by the Electronic Theses and Dissertations at Scholarly Repository. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

UNIVERSITY OF MIAMI

MORAL INJURY IN CONTEMPORARY ETHICS:
THE APPLICATION OF A SOCRATIC IDEA

By

Heleana Theixos

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida
August 2016

©2016
Heleana Theixos
All Rights Reserved

UNIVERSITY OF MIAMI

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

MORAL INJURY IN CONTEMPORARY ETHICS:
THE APPLICATION OF A SOCRATIC IDEA

Heleana Theixos

Approved:

Kenneth W. Goodman, Ph.D.
Professor, Ethics Programs

Risto Hilpinen, Ph.D.
Professor of Philosophy

Wilson H. Shearin, Ph.D.
Associate Professor of Classics

Guillermo Prado, Ph.D.
Dean of the Graduate School

Jeff McMahan, Ph.D.
White's Professor of Moral Philosophy
Oxford University, Corpus Christi College

THEIXOS, HELEANA
Moral Injury in Contemporary Ethics:
The Application of a Socratic Idea.

(Ph.D., Philosophy)
(August 2016)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Kenneth W. Goodman.
No. of pages in text. (161)

This dissertation is an investigation of the Socratic moral claim that being a moral wrongdoer is worse for the wrongdoer than it is for the victim. Chapter One investigates this moral claim in the context of Plato's *Gorgias* and the historical Gorgias' *Encomium of Helen*. Chapter Two brings the ancient conception into contemporary ethics into a typology of what the ancient concept looks like in contemporary ethics. Chapter Three prepares for an investigation of a certain moral wrongdoer, the military interrogator/torturer and wrongdoing, by first examining a standard torture hypothetical, the Ticking Time Bomb. All three chapters work together as they revolve around my explication and characterization of the ancient and contemporary phenomenon of *moral injury*.

Table of Contents

Expanded Table of Contents	1
Introduction	7
Chapter One	
What did Socrates mean when he claimed that moral wrongdoing is bad for the moral wrongdoer?	13
Chapter Two	
Is the Socratic claim that moral wrongdoing is bad for the moral wrongdoer evident in empirical studies? In what way? How can the concept of bad for the moral wrongdoer, dubbed moral injury, be a useful concept for moral philosophy?	69
Chapter Three	
Checking the moral intuitions we derive from the Ticking Time-Bomb Hypothetical with resources and constraints in revisionist just war theory. Does the interrogator's susceptibility to moral injury change our moral intuitions?	127
Conclusion	148
Bibliography	152

Expanded Table of Contents

Chapter One

What did Socrates mean when he claimed that moral wrongdoing is bad for the wrongdoer?

Preface

§1.1.1. Rhetoric in the Classical Context

Socrates says that craftsmen fit together what is appropriate to a certain order, according to each component's own function. There is a standard to which these things can be evaluated, at *Gor*, 504d. If the analogy of rhetoricians with craftsmen holds, then rhetoric is, despite Callicles' objection, eligible for our scrutiny, and so Socrates is setting the stage to talk about how rhetoric ought be as well as how the good rhetorician ought be.

§1.1.2. My Examination of Gorgias on Rhetoric in Plato's *Gorgias*

Premise 1: The rhetorician teaches his students to be powerful orators.

Premise 2: The teacher is not responsible for the mis-use of rhetoric. Gorgias makes the analogy to a boxer who is trained in the ring and then later assaults his father. We would not blame the trainer for the boxer's mis-use of the craft.

Conclusion: rhetoric is a morally neutral tool, and so the rhetorician, the teacher, is not to blame for its misuse.

§1.1.3. My Examination of Gorgias on Rhetoric in his Encomium of Helen

Premise 1: The rhetorician relies upon the listener's own *doxa* as a primary instrument. The *doxa* is the presupposed psychic state, a pre-existing state of opinion, which the rhetorician uses to his own purposes.

Premise 2: The rhetorician is concerned with communication and persuasion. It makes sense that he uses the *doxa* to aid in the persuasion.

Conclusion: Belief is a combination of persuasion of the rhetor and the persuadeable capacities of the listener.

Gloss: For example, great literature rests on a deception, *απάτη*, which is not moral or immoral but the means to communicate. Rhetoric is not moral or immoral, in part because belief is co-created by the listener's own *doxa*.

§1.1.4. My Examination of Socrates on Rhetoric in Plato's *Gorgias*

The dialogue presents us with two "teachers" who have different opinions on their roles

as teachers and on the meaning of what they practice. I see Socrates alluding to the interlocutors who are unjust agents wielding a moral *techne* without recognizing their own complicity in a wrongdoing, and that they should accept the punishment, in the form of the truth of the matter, which Socrates is doling out. This is a system which is alluded to at the end of the dialogue, *Gor.* 517a, leading the discussion back to the (moral) nature of rhetoric, which is where the discussion began.

§1.1.5. My Examination of Rhetoric as a Moral *Techne*

I characterize what I think Socrates is arguing for, ethically, in Plato's *Gorgias*, and how his argument fits into his overall moral ontology. My focus is on Socrates and Gorgias' debate about the nature of rhetoric, *Gor.* 448e6-461b2. I characterize what I take to be Socrates' overall moral proposition, what I call his fulcrum argument, and how that bears upon the questions of what rhetoric is but more to the ethical point, who an unjust rhetorician is.

§1.1.6. My Examination of The Ethical Psyche

I explore what I take to be Socrates' argument that rhetoric is a kind of moral τέχνη, which ushers in questions about what sorts of reasons we have to pay attention to the implications of a moral *techne* like rhetoric. Ultimately the question of why should I be moral is answered by Socrates with, your εὐδαιμονία depends on it. But how is that connection made? This section explores the foundational assumptions Socrates makes about ethics and the *psuche* in order that ethics is connected, via the *psuche*, to *eudaimonia*.

Chapter Two

Exploring the Socratic concept of moral injury in contemporary ethical theory. Does the Socratic concept hold? In what way? A Typology of moral injury.

Preface

§2.1.1. An Introduction, Overview, and Preparatory Discussion of the Subsequent Sections

An introduction, overview, and preparatory discussion of the subsequent sections
In this section I motivate my entire project, that there is such a thing as a particular set of moral beliefs which certain moral wrongdoers judge of themselves, which are separate from the moral beliefs a witness or an innocent bystander holds. I begin with a look at the psychological material, to provide an empirical survey of the unique phenomenology and beliefs which are formed in being a moral wrongdoer. The significance is not just important to psychology, it is important to philosophy, because it regards a legitimate, narrow, unique kind of moral belief-state.

I will show that the philosophical community has good reason to also be interested in this phenomenon, as moral injury is a new moral category, that of a particular belief-state, the

phenomenology of being a certain kind of morally culpable wrongdoer. The phenomenon of moral injury entails new and unique judgments, beliefs, and impact to well-being.

§2.1.2. Moral Wrongdoers with Moral Injury: Narratives

In this section I incorporate actual first-person narratives by a certain subset of moral wrongdoers: soldiers who have perpetrated serious moral transgressions and are sensitive to moral reasons. This section serves to explore the phenomenology of being a certain kind of moral wrongdoer.

Chapter Two, Section Two

§2.2.1. Typology of Moral Injury: Moral Trauma

The concept of moral trauma as I am using it has been used by moral theorists, feminist theorists, and psychologists, but the descriptions are varied. My working definition of moral trauma, then, is: a significant transgression against the moral beliefs, moral codes, moral expectations, held by most people. My characterization and definition of the term moral trauma is born from the conjunction of a traumatic experience and the significant breach of strong moral code. Moral trauma is thus the event, an event which is interpreted as having violated the codes of the moral community, and/or the moral sense of the agent.

§2.2.2. Typology of Moral Injury: Moral Injury

Moral injury is the term I use for the psychological phenomenon of being sensitive to having caused moral trauma. This is a term which is used loosely in contemporary philosophy and psychology, in a variety of ways depending on the theorist, but which I argued in Chapter One is what the ancient Greek thinkers meant about moral wrongdoing harming the moral wrongdoer. The moral trauma is the traumatic event, the moral injury is the particular kind of psychological experience, the phenomenology, of being the person who is responsible for committing, or failing to prevent, the moral trauma.

§2.2.3. Common Morality

This section is not a clinical analysis but a discussion of how moral injury as I define it can be useful in reference to the affect to the moral identity of agents who are causally involved in moral trauma. I combine insights from the empirical psychological research with the insights of moral philosophy, especially as it relates to the psyche and character of the moral transgressor.

I make an appeal to Gert's characterization of the on-the-ground moral systems in order to contextually place the moral wrongdoer. The wrongdoer has transgressed against this public, other-regarding system, and her claims about breaking moral codes are primarily claims about codes of behavior which others hold her to. Thus the moral wrongdoer has some sense of the moral rules of her community, evidenced by her sensitivity to her actions being deemed moral transgressions, which her community will hold her to

account. She is an informed public citizen participating in an informal public moral system, where disagreements arise and can be resolved, which shows that these are rational agents in conversation with a moral system.

Chapter Two, Section Three

§2.3.1. Moral Repair for Moral Injury

What the moral wrongdoer is engaged in is contemplating the moral reactive attitudes of her peers, and whether she ought accept their judgments. This involves her taking seriously those reactive attitudes, and how they bear upon her conception of herself as a moral being. By establishing that there exists such a concept as common morality which takes seriously social moral criticism, there is then a moral concept from which to begin talking about what the moral wrongdoer believes to be true about herself.

§2.3.2. The Morally Injured can Legitimately Pursue Moral Repair

I offer the description of holding oneself accountable to the moral reactive attitudes, to ground much of what is interesting in moral injury, that the wrongdoer holds herself accountable to the moral reactive attitudes to such a degree that her own moral goodness is implicated. The designation of holding oneself accountable is a two-part relation, where the moral community holds the wrongdoer accountable (as expressed by the moral reactive attitudes), and the wrongdoer holds herself accountable to those attitudes (as expressed by taking them seriously, including downgrading her own moral standing).

§2.3.3 Mitigating Blame for the Morally Injured

In their research, theorists such as Walker and Gert have emphasized the importance of the moral dimension of our social reality – how normative expectations shape both our self conception and our relations to others in fundamental ways. In this section, I offer an exposition of Gert and Walker’s Strawsonian notions of moral relationships and moral repair. Giving examples to illustrate these ideas, I show the crucial role that moral expectations play in our self conception, relationships, and greater community. Fully explaining the nature and processes of moral repair enables me to move towards my conclusion – that even wrongdoers deserve the chance to heal their moral character and damage done to their moral relationships.

Chapter Three

Checking our moral intuitions elicited from the Ticking Time Bomb hypothetical against constraints in revisionist just war theory. Incorporating moral concern for the torturer.

Preface

In this chapter I check the moral intuitions that the classic Ticking Time-Bomb

hypothetical elicits against moral constraints through the lens Jeff McMahan's reexamination of traditional just war theory, revisionist just war theory.

§3.1.1. The Power of the Ticking Time-Bomb Hypothetical in Intuitive Moral Thinking and in Practice

There exists a phenomenon known as torture-analysis error, what Darius Rejali describes as commitment to the "folklore of torture with social scientific legitimacy," of which the TTB hypothetical plays a pivotal, and privileged, epistemic role.

Section II. Checking our TTB Intuitions Against Constraints in Revisionist Just War Theory

§3.2.1. Investigating Principles of Liability: The TTB Hypothetical Stipulates That The Detainee is a Legitimate Target, Who Makes Himself Liable to Defensive Harms

The primary justification for the permissibility of torture as a mode of defensive harm in the TTB hypothetical is that the detainee has, "through his own culpable action, made himself liable to be tortured in defense of his innocent potential victims." In the TTB case, because we accept that the detainee is morally culpable if the bomb detonates, then that so long as he refuses to divulge the bomb's location he makes himself liable to some degree of defensive harm.

§ 3.2.3. Checking our TTB Intuitions With Principles of Necessity & Proportionality

That which must be done, that which is necessary, is that which will cause the least amount of harm to the detainee in the genuine pursuit of saving the innocents. The necessity constraint forbids the use of harms that are unnecessary to the aim sought, meaning, are egregious, gratuitous, and thus punitive or sadistic torture is outside the parameters of the TTB case.

Section III. Checking what harms to the interrogator are morally permissible

§ 3.3.1. The interrogator is liable to psycho-social effects of torture like moral injury

§ 3.3.2. Checking the interrogator's liability to moral injury

A simple reading of liability would say that the interrogator is the source of the threat to her own well-being, and thus she has some degree of liability for the possible traumatic effect. A more robust reading of liability would ask questions as to whether she really, as an agent of the state, can decline orders to interrogate. A more robust reading could also ask her to weigh the effects to her own well-being against the risks of the bomb's detonation.

§ 3.3.3. Checking whether moral injury to the interrogator is proportionate

The harm to the interrogator would be unintentional, not wrongful in terms of liability, and foreseeable. In this way the interrogator becomes complicit in her own suffering, a disconcerting perspective to have when discussing the moral considerations relevant to torture.

§ 3.3.4. The Disconcerting Conclusion

Introduction

In Chapter One I investigate Socrates' striking moral claim, that being a moral wrongdoer is worse than being a victim of moral wrongdoing, *Gor.* 474b4, and that to do a moral wrong is the greatest of evils, *Gor.* 469b8-9. Socrates alludes to his moral claim in various dialogues including in Plato's *Republic* I 354a9, where he says that injustice is never more profitable than justice. This Socratic claim is made most explicit and is most thoroughly analyzed throughout Plato's *Gorgias*, so my focus is on that dialogue.

I take Plato's *Gorgias* to be a moral dialogue exemplifying the Socratic moral philosophy that the moral wrongdoer is not free from the bad effects of moral wrongdoing. Socrates is not precise about his claim, he says that moral wrongdoing is bad in some way—to the psyche, *psuche*, soul, conscience, practical well-being, happiness/*eudaimonia*—to the moral wrongdoer, and I investigate what he could mean despite his ambiguity.

Adding to the difficulty of Socrates' ambiguity is that the dialogue's interlocutors are explicitly incredulous. Socrates invites his interlocutors, who are all rhetoricians, to examine how the moral wrong of manipulating audiences against what is true can be bad for the interlocutors themselves. Both Gorgias and Socrates offer an account of what rhetoric is; for the character of Gorgias in Plato's *Gorgias*, rhetoric is, for the most part,

morally neutral. In contrast, the historical Gorgias in his Encomium of Helen is much more explicit about the moral neutrality of rhetoric.

In Plato's dialogue, Gorgias admits, perhaps begrudgingly, that while unjust rhetoric ought be the moral concern of the rhetorician the rhetorician is only nominally blameworthy if his rhetoric leads to bad ends. Gorgias' students take a more strident perspective, and offer examples of moral wrongdoers who are clearly fairing well in life, directly contradicting the Socratic claim. The dialogue ends with the interlocutors seemingly unconvinced. As such the dialogue is often seen as a failure, an example of bad argumentation on Socrates' part, or an example of the trivial claim that rhetoric can be used unjustly. My exploration delves into this question: what did Socrates mean when he claimed that moral wrongdoing is bad for the moral wrongdoer? In this Chapter I hope to investigate and clarify the seemingly failed moral claim.

In section One of this Chapter I examine Socrates' conversation with Gorgias about the nature of rhetoric in Plato's Gorgias. Section One compares and contrasts the historical Gorgias in his own work, The Encomium of Helen, with the character in Plato's dialogue Gorgias, regarding unjust rhetoric being an example of moral wrongdoing.

In Section Two I examine what I take Socrates to be doing overall, which is arguing that moral wrongdoing is bad for the moral wrongdoer, of which (unjust) rhetoric is the example in Plato's Gorgias. To do this I pursue my hypothesis that Socrates views unjust rhetoric as a kind of moral *techne*, which can be implemented (morally) wrongly, which ushers in moral questions about its use and implications. Is the Socratic claim that moral wrongdoing is bad for the moral wrongdoer evident in empirical studies? In what way?

How can the concept of bad for the moral wrongdoer, dubbed moral injury, be a useful concept for moral philosophy?

In Chapter Two I explore and offer a typology of moral injury. The *Journal of Traumatic Stress* has published several articles linking perpetrating a serious moral transgression as having unique and elevated psycho-social effects compared to having witnessed serious moral transgressions. Killing in combat is one area of research: studies indicate that soldiers who killed others in combat experienced high instances of shame, regret, remorse, and debilitating psycho-social disorders as compared to soldiers in combat who did not kill. What is emerging in these studies is a unique set of moral beliefs which certain moral wrongdoers hold about their own moral goodness, character, and belonging, in very sharp contrast from soldiers who witnessed, but did not perpetrate, serious moral wrongs.

One such study, the Maguen study of US soldiers in 2009, shows that killing was a significant predictor of PTSD, alcohol, and substance abuse, as compared to soldiers who had witnessed killing, but who had not killed anyone. In a similar study in 1990, Green et al identified eight categories for significant predictor of PTSD, drug abuse, domestic abuse, and violent behavior, a statistically significant component being “causing severe harm to another,” or “killing another.” I explore these and other studies to draw out the Socratic claim that being a moral wrongdoer is worse than being a victim of moral wrongdoing. I explore the concept of *psuche*, *soul injury*, to use Socrates’ characterization, and connect it with *moral injury*, a contemporary term used in psychology today and in these studies.

These studies are tracking something important for moral psychology: that perpetrators of significant moral wrongdoing have distinct psychological experiences and a particular set of moral beliefs which are separate from the psychological beliefs, and corresponding distress, in having witnessed moral trauma. My interest is in the content of these beliefs, which regard the moral wrongdoer's actions as they bear upon their own character and moral belonging in their community. My interest is in the phenomenology of being a moral wrongdoer, and the sorts of beliefs and experiences which are held by this category of moral agent.

I clarify the contemporary term moral injury so that it refers to the judgments that certain moral wrongdoers make about their character in light of the wrongdoing, pace the empirical studies and pace the Socratic concept of soul injury in the *Gorgias*. I am characterizing the phenomenology of a special type of moral agent, that of having perpetrated a serious moral transgression and the new moral beliefs which are formed. I propose that having a characterization, or category, for moral injury provides moral philosophers with richer language about the phenomenology of moral wrongdoing. Usually, moral philosophers analyze the wrongness of moral wrongs in relation to victims, or look at the importance of following moral rules in themselves. For example, the contemporary moral theory does not focus on beliefs about the rightness or wrongness of killing to the killer, who is eligible to develop certain moral injury, post-killing.

By looking at the moral beliefs the killers hold about their own moral character, moral philosophy then has a unique concept, moral injury, to refer to regarding the

rightness and wrongness of killing.

My work places the moral wrongdoer's own moral judgments about her own moral character at the center of ethical analysis. I am clear to propose that not all moral wrongdoers have the capacity for moral self-reflection, and so not all moral wrongdoers would hold these sorts of self-critical, and self-referential, moral beliefs. Thus, only certain moral wrongdoers who have a threshold of moral sensitivity are eligible for the phenomenon of moral injury.

In Chapter Three I check the moral intuitions that the classic Ticking Time-Bomb hypothetical elicits against moral constraints through the lens of Jeff McMahan's revisionist just war theory. The TTB case acts as a serviceable framework for our moral intuitions, providing us with a clear, concrete thought experiment from which we come to a moral conclusion. But many have challenged the TTB hypothetical's moral and practical legitimacy as wildly unrealistic, that it fails to provide the mechanisms from which we would form an absolute moral prohibition against torture, or that it gives us (the wrong reasons) for bending the torture victim's will against itself, a Kantian criticism. These criticisms largely conclude that both in practical reality and in sincere philosophical analysis the case cannot do any real moral work. I accept that these criticisms have merit: I accept that the TTB hypothetical is a wild fantasy disconnected from real cases, that it fails at eliciting absolutist intuitions, that it does not allow for our sentiments to fully consider the infringement of the detainee's will (although it does give us the opportunity to weigh the harms done to the tortured against the harms to the many innocents). I accept the criticisms that the TTB hypothetical is a morally pernicious

thought experiment which has been radically misused in our contemporary intellectual, legal, and military debate. I examine our moral intuitions generated by the TTB hypothetical against permissions and constraints in Jeff McMahan's Revisionist Just War Theory. My aim is to take the TTB hypothetical—an important and relevant thought experiment which, while being a wild fantasy, nevertheless elicits moral intuitions which satisfy RJWT theory constraints—and incorporate into it appropriate moral concern for the moral injury of the torturer, which I will do in my future scholarship.

Chapter One:
What did Socrates mean when he claimed that moral wrongdoing is bad for the moral wrongdoer?

§1.1. Preface

In this Chapter, I examine Socrates' conversation on the nature of rhetoric with the character of Gorgias, as Plato depicts him, in the dialogue of the same name. I take Plato's *Gorgias* to be a moral dialogue exemplifying a Socratic moral philosophy: that the moral wrongdoer is not free from the bad effects of moral wrongdoing, and moral wrongdoing is sufficient for fairing badly. I articulate how unjust rhetoric bears upon the (unjust) rhetorician, such that his own *eudaimonia* is implicated. I explore the possibility of Socrates' viewing rhetoric as a kind of moral *techne*. Ultimately the question of *why should I be moral?* is answered by Socrates with, individual *eudaimonia* depends on it.

Both Gorgias and Socrates offer an account of what rhetoric is. For the character of Gorgias in Plato's *Gorgias*, rhetoric is, for the most part, morally neutral until, perhaps begrudgingly, Gorgias admits that unjust rhetoric ought be the moral concern of the rhetorician, and yet the rhetorician is only nominally blameworthy if his rhetoric leads to bad ends. In the dialogue, Gorgias himself takes a moderate view of the moral

implications of rhetoric while his students take a more strident perspective, and so I look more closely at what the character of Gorgias, the namesake of the dialogue, proposed regarding the moral color of rhetoric. The historical Gorgias who wrote the *Encomium of Helen* is much more explicit about the moral neutrality of rhetoric, and I will examine what I take to be the historical Gorgias' views on unjust rhetoric in his own work and compare and contrast that with the character of Gorgias in Plato's dialogue.

Towards the end of Plato's *Republic* Book I, Socrates offers Thrasymachus an elaborate argument with the conclusion that "injustice is never more profitable than justice." *Republic* I 354a9. The dialogue is as follows, starting at 354a1:

"And the man who lives well is blessed and happy, and the man who does not is the opposite."

"Of course."

"Then the just man is happy and the unjust man is wretched."

"Let it be so," he said.

"But it is not profitable to be wretched; rather it is profitable to be happy."

"Of course."

"Then, my blessed Thrasymachus, injustice is never more profitable than justice"¹

The argument's structure is what I am interested in, and this is my interpretation:

The excellence of the *psuche* is justice.

The just man with a just *psuche* lives well/happily, the unjust man badly.

The unjust man with his unjust *psuche* can never have *eudaimonia*.

Conclusion: Moral wrongdoing is of non-trivial harm to the moral wrongdoer.

I interpret this structure operating as a fulcrum within Socrates' overall moral ontology, in that the argument anchors moral premises that Socrates accepts as true, like his proposition that everyone aims at *eudaimonia*, and that *eudaimonia* is dependent upon

¹ Plato, *The Republic*, Second Edition, *With Notes and an Interpretive Essay*, trans. Allan Bloom (New York: Perseus Books, 1968), 33.

acting rightly in an *ethical* sense. The fulcrum argument's conclusion, that it never benefits a person to be unjust, is echoed, paraphrased, or referred to explicitly throughout the *Republic*, as in Book II, where the interlocutors re-state the fulcrum argument and repeat what they take Socrates to have said, that "it is in every way better to be just than unjust" *Republic* II 357b1. In this chapter I will clarify and explicate the fulcrum argument and analyze its importance within the context of Plato's *Gorgias*, the dialogue where the fulcrum argument is made explicit. I take much of the dialogue in the *Gorgias* to be a long explication of the fulcrum argument, with rhetoric and rhetoricians being central examples. If the fulcrum argument holds, then the interlocutors themselves will fail to have *eudaimonia*, and so the stakes are very high for the *Gorgias*' interlocutors, as it was for the *Republic*'s sophist Thrasymachus.

§1.1.1. Rhetoric in the Classical Context

In the *Gorgias*, Socrates compares the rhetorician with a housebuilder,

[The rhetorician] is just like any other craftsman, who having his own particular work in view selects the things he applies to that work of his, not at random, but with the purpose of giving a certain form to whatever he is working upon. You have only to look, for example, at the painters, the builders, the shipwrights, or any of the other craftsmen, whichever you like, to see how each of them arranges everything according to a certain order, and forces one part to suit and fit with another, until he has combined the whole into a regular and well-ordered production; and so of course with all the other craftsmen, and the people we mentioned just now, who have to do with the body—trainers and doctors; they too, I suppose, bring order and system into the body. Do we admit this to be the case, or not? *Gor.* 503e1-504a6.²

² Plato, *Plato in Twelve Volumes*, Vol. 3, trans. W.R.M. Lamb (London: Harvard University Press, 1967), Perseus Digital Library, Accessed June 20, 2016, <http://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0059.tlg023.perseus-engl>.

Socrates says that craftsmen fit together what is appropriate to a certain order, according to each component's own function. There is a standard to which these things can be evaluated, at *Gor.* 504d1-504e5. If the analogy of rhetoricians with craftsmen holds, then rhetoric is, despite Callicles' objection, eligible for our scrutiny, and so Socrates is setting the stage to talk about *how* rhetoric ought to be as well as *how the good rhetorician ought to be*.

Rachel Barney, in her examination of Plato's defense of rhetoric in the *Gorgias*, analyzes Socrates' and Gorgias' back-and-forth discussion as a kind of *elenctic* dance, where each offers a theory of what rhetoric is, how unjust use of rhetoric is or is not the concern of the teacher of rhetoric, and how Gorgias' presumption of rhetoric's neutrality is countermanded by Socrates' investigation of rhetoric as a *social* kind.³ Barney is right to remind us that the *Gorgias* announces itself as a dialogue of definition, that it is an "attempt to identify Gorgias by the profession he practices."⁴ What flows naturally from this sort of framework is that if we are to ask what rhetoric is, then we are, naturally to Socrates' view, asking what sort of person Gorgias is, and here is where I part ways with Barney and read the dialogue not as an examination of rhetoric as such, but an examination of unjust rhetoricians in particular, and the connection to (their, and our) *eudaimonia*. While all interlocutors comment and define what the craft of rhetoric is—a deception, an imposter, a knack, or a habit—what is at stake, on my reading, is what sort

³ Rachel Barney, "Gorgias' Defense: Plato And His Opponents On Rhetoric And The Good," *The Southern Journal 5 of Philosophy* 48, no.1 (2010):119.

⁴ Barney, 107.

of person practices this sort of thing, which may in fact be a social kind, and in what way, viciously or virtuously, and to what impact on their own well-being.⁵

Rhetoricians in the Classical period were, in George A. Kennedy's analysis, speech-makers of a certain kind, they "incorporated style and eloquence in order to persuade the listener toward a certain goal."⁶ In "How Good Should an Orator Be?" Ølivind Anderson writes that the Greeks of the Classic period had an ambivalent relationship to clever speaking and clever speakers, that discussions in Thucydides and Antiphon criticize unchecked, manipulative rhetoric, but that many prominent Athenians, like Demosthenes, praised eloquent, persuasive rhetoric.⁷ Josiah Ober, in *Mass and Elite in Democratic Athens*, characterizes the Athenian perspective on rhetoricians as a tension between admiration and distrust of rhetoric.⁸

⁵ Ibid., 107. Barney sees Socrates' project as a normative analysis of social kinds, and so the critique Socrates offers is not, to Barney's view and *contra* Aristotle, a critique about the bad effects of abusive rhetoric, but is instead a project to identify rhetoric as a social kind, eligible for a normative, depersonalized meta-critique. Scientific disciplines divide particulars into kinds; a natural kind is a particular which corresponds to a group belonging to the natural world. A social kind is a particular which corresponds to a group belonging to the interests and actions of, in this case, human beings, but there can also be bovine social kinds, or amphibian social kinds.

⁶ George A. Kennedy, in *A New History of Classical Rhetoric* (New Jersey: Princeton University Press, 1994) deconstructs and analyzes the *art* of rhetoric, in contrast to the non-artistic speech acts like those made when giving testimony in a trial. Kennedy says that "the artistic method of persuasive rhetoric has three distinct characteristics: the character or *ethos* of the speaker as trustworthy based on what he says in the speech; the arousal of emotion, *pathos*, in the audience; and the use of argument, *logos*, that appears to show something." Kennedy notes that this characterization "reflects Aristotle's discussion in *Rhetoric*, 1.3.1, where Aristotle says that there are only three 'species' of rhetoric: judicial, deliberative, and epideictic, depending not on the function of the speaker but on the function of the audience, and also 2.18.1, according to the success of the persuasion, determined in part by the emotions and states of mind awakened in the listener," Kennedy 57-59.

⁷ Demosthenes, from his oration "On The Crown," characterizes himself both as a good leader and a good rhetorician and that he is to be acclaimed for his speaking ability, as transcribed in Anderson, "How Good Should an Orator Be?" and in C. W. Wooten (ed.): *The Orator in Action and Theory in Greece and Rome. Essays in Honor of G. A. Kennedy*. (Leiden, Boston, and Cologne: Brill, 2001).

⁸ Josiah Ober, *Mass and Elite in Democratic Athens: Rhetoric, Ideology, and the Power of the People* (Princeton: Princeton University Press, 1989), 170, 178, 187, 189.

On balance, Athenians viewed skilled speech as something that could be virtuous and artful, as when it enables a powerful leader to succeed, but on the other hand it was problematic, as when a persuasive speech manipulates the audience to uninformed conclusions, threatening and undermining the stability of democratic society. When Socrates refers, in Plato's *Apology*, to his own speechmaking as both powerful and weak, this may be an example of Socratic irony, but it may also be a nod to the Athenian ambivalence, the tension Josiah Ober described as an Athenian *wariness*, despite the Athenian tradition to "listen willingly, even eagerly, to the speeches of trained orators both in the Assembly and in the courts."⁹

Aristotle, in his *Rhetoric*, views Socrates' critique of rhetoric in Plato's *Gorgias* as through the lens of the *harm criterion*. Aristotle offers an account of things that are presumptively good or presumptively neutral—justice, fairness, even neutral physical elements like earth and water. This is what we have in mind when we speak of things being good, generally, and reasonably, such that these things do not invariably have bad effects and thus they are presumptively good, *Rhetoric* I.I.12, or neutral (perhaps not neutral, exactly, but more an "abusable good," *Nichomachean Ethics* I.3, 1094b18).¹⁰ Aristotle offers a positive defense of the presumptive goodness of rhetoric, contra the

⁹ Ober, *Mass and Elite in Democratic Athens*, 185.

¹⁰ In E.M. Cope's analysis of the way in which Aristotle characterizes rhetoric (the way rhetoric relates to dialectics, 2 as branches, parts, or species of probable reasoning, lower in level and subordinate to the art of communication), Cope 2. Cope characterizes the sophist Callicles as, in Aristotle's eyes, performing persuasion at a morally low level, and that this is Socrates' primary criticism, of the (perhaps intentional) harm Callicles' rhetoric does to his audience, E.M. Cope, *Aristotle's Rhetoric With Analysis Notes and Appendices* (London and Cambridge: MacMillan and Co, 1867), 136, 257.

Socrates of the *Gorgias*, saying that rhetoric bends toward what is just, or better, or desirable, *Rhetoric* I.I.12.

Rachel Barney takes up the task of interpreting what the Socrates of the *Gorgias* is trying to do in the dialogue by noticing Aristotle’s interpretation (but ultimately disagreeing with Aristotle’s characterization):

But if Plato is willing to embrace that [harm] criterion elsewhere [in other dialogues], his attack on rhetoric in the *Gorgias* does not seem to depend on it (but with the weak objection that rhetoric can be unjustly used).¹¹

Barney, in contrast with Aristotle, views Socrates as undertaking to “analyze and critique rhetoric as such—as a social institution, a well-defined kind (like a natural kind) with intrinsic features of its own, evaluable independently of the intentions and qualities of its practitioners.”¹²

§1.1.2. My Examination of the Character of Gorgias on Rhetoric in Plato’s *Gorgias*

Gorgias says that rhetoric is no small power, that it is in fact “The greatest of human affairs, Socrates, and the best,” *Gor.* 451d8. Gorgias says that all powers—political, legal, social—flow from rhetoric, *Gor.* 451e1-6, and that powerful and successful rhetoric enables rulers to rule and freedoms to be exercised. Gorgias says that his point is self-evident, because we all see that political leaders advise the city, not

¹¹ Barney, 118.

¹² Barney, 118. Rhetoric may indeed be eligible for analysis, as Barney concludes, as a socially constructed practice, and that may be a good enough view of Socrates’ aim in Plato’s *Gorgias*, but it is certainly not the most interesting thing to say about that which can be misused such that the mis-user cannot be *eudaimon*. I engage with Barney later in my analysis.

craftsmen *Gor.* 455d1-456a7, and political leaders are themselves the most powerful rhetoricians who direct and guide the city's affairs.

Gorgias, the leader of a school on rhetoric and teacher to the interlocutors we see in the dialogue, thus argues for a philosophy of rhetoric's centrality in political affairs, and he makes four attempts to specify what rhetoric is. For the first attempt, Gorgias offers that rhetoric is concerned with the *logos*, that is, the reasoning, the words and speech which convey meaning, to a discourse, *Gor.* 449e1. This is unsatisfactory, because surely the *logos* is not exclusive to the practice of rhetoric. For his second attempt, Gorgias' says that rhetoric is about the subject it conveys, *Gor.* 451a3, but that is also an insufficient characterization of the particular thing that rhetoric is. Gorgias' third attempt regards what rhetoric produces, that is, the persuaded listener, and as such rhetoric is the greatest good, *Gor.* 452d5, a good which is evident in what it produces in the law courts and in the city broadly.

Gorgias' fourth and final attempt is that rhetoric, in conjunction with the *logos*, persuades the multitude, *Gor.* 452e1-8. Gorgias says that the power of rhetoric can make the doctor a slave, a trainer a slave. Later, Gorgias agrees that the good of rhetoric is in persuading the audience to what is just and unjust, *Gor.* 454b5-7. When the conversation shifts to whether the teacher of rhetoric ought then to know what is right and wrong, and whether he ought to be held responsible for when his pupil's persuade people against the good, Gorgias portrays rhetoric and rhetoricians as analogous to boxers and their trainers, that the trainers have taught their art assuming that the trainees will use it justly, *Gor.* 456d-457a.

Gorgias finally frames rhetoric as good when it is successful at achieving its aim of persuasion to what is right, and bad when it is successful at achieving its aim in persuading what is wrong. This view means that rhetoric takes on a morally neutral color, neither good nor bad *in itself*. In Dodds' commentary on the *Gorgias*, Dodds interprets Gorgias as framing rhetoric not as anti-moral but morally neutral, like other technical skills.¹³

Here is what I take Gorgias' fourth attempt, and his central argument in the dialogue, to be:

Premise 1: The rhetorician teaches his students to be powerful orators. Gorgias makes the analogy to a boxer who is trained in the ring and then later assaults his father.

Premise 2: We would not blame the trainer for the boxer's mis-use of the craft.

Conclusion: The rhetorician, the teacher, are not to blame for rhetoric's misuse, and so rhetoric is a morally neutral tool and the teacher is blameless.

Gorgias' defense of rhetoric as analogous to boxing seems plausible: if rhetors are like boxers, then their success in the ring is what matters, and they are not to blame if their students abuse it, at *Gor.* 455a1-457c5, the training and assaults someone outside the boxing ring.

Success in the boxing ring is analogous to success in the law court or political stage, and success in those environments

depended entirely on the ability to persuade the public. Gorgias, Polus, and Callicles all assume that the point of rhetoric is the acquisition of power, so that the status of tyrant—that is, command of all—is the natural terminus of successful rhetorical activity; they differ merely in their squeamishness about making the assumption explicit to themselves as well as to Socrates.¹⁴

¹³ Dodds, *Commentary*, 212.

¹⁴ James Doyle, "Socrates and Gorgias," *Phronesis* 55, no. 1 (2010): 1.

Rhetoric is ethically charged, on Gorgias' view, only in the way in which the legal and political successes *take form*, and so asking whether the rhetoric is ethical is not the right question; the right question is, is rhetoric persuasive to the good? To Gorgias' view, rhetorical skill is good when it is useful for legal and political success, successful rhetoric can determine the legal and political culture of the *polis*, the goodness of rhetoric is determined after its success had been determined.

I interpret Gorgias as viewing the teachers of rhetoric to be considered separately from the pupil who misuses rhetoric, and that blame for unsuccessful or wrongful rhetorical persuasion belongs to the wayward student orator, *Gor.* 457b1-457c4.¹⁵ My criticism of the analogy is that boxers are targeting other trained boxers, unlike rhetoricians persuading their untrained audience. In the boxing ring there is a meeting of minds, skill, and rules. In the ring there is a clarity of target, a defined scope and goal that two equally trained fighters are mutually aiming to achieve. The example of an unjust rhetor better fits the boxer who assaults the father: the untrained father suffers all the damage.

The unskilled audience is the father in this analogy, and the trained rhetorician is the boxer. Gorgias admits that he is aware that a boxer can use his skills unjustly akin to how a wayward student of rhetoric may use his skills unjustly. What could Gorgias be implying to his own students who are present about his own responsibility in teaching his students right from wrong? He seems to imply that Gorgias himself is not to be held

¹⁵ Dodds writes that "Plato was "always careful to distinguish the Socratic dialectic, which aims only at the attainment of truth, from its vulgar counterfeit the *εριστικ* or 'anti-λογη', which aims at personal victor and is a mark of *α παιδευσία*," as characterized in the *Phaedo*, 91a. Dodds, *Commentary*, 213.

responsible, nor is rhetoric itself to be considered a moral art or technique, but that both the teacher and the tool are morally free from blame.

The boxing analogy could also signify that rhetoric is no weak tool but one that can cause clear and serious damage, as rhetoric, through superior power of speech and persuasion, could significantly persuade an assembly. For Aristotle, Socrates of the *Gorgias* is arguing for a critical review of the practice, or craft, of rhetoric on the grounds of the harm that unjust rhetoric can cause an audience. In this way Aristotle seems to interpret the Socrates of the *Gorgias* as believing rhetoric to be presumptively bad. It is important to remember that Gorgias is speaking in front of a largely anonymous audience who occasionally applaud and who occasionally seem uncomfortably silent, as well as speaking in front of his own (paying) pupils.

Gorgias gives no explicit characterization of what justice is, or what he takes it to be, only that whatever justice is, it is not his purview. The conversation between Socrates and Gorgias concludes with Socrates asking whether it is the duty of a teacher of rhetoric to educate his pupil on what good and bad is, and to that Gorgias agrees that one should, and that he himself does, *Gor.* 460a2-3. When Socrates presses Gorgias about knowing or teaching justice, Gorgias says that he supposes that if the student truly does not know what justice is he would eventually learn it, *Gor.* 460b1-5.

Gorgias has been consistent through the fourth attempt at defining rhetoric, that rhetoric and knowledge of justice are distinct, that good rhetoric is persuasive, and that rhetoric ought to persuade toward the good, but that what the good actually is is indeterminate. There is a slight waver at the fourth attempt, when Gorgias concedes that

the rhetorician who has been taught what is good ought to persuade listeners to the good, *Gor.* 460d7, but that is in contrast with his earlier position that the teacher ought not to be blamed if the pupil acts wrongly, *Gor.* 457c1-3. Socrates' proceeds at length that the teacher of rhetoric must know what justice is, and be responsible for his students when they persuade listeners contrary to justice, *Gor.* 461a1-461b1. Gorgias becomes silent, and Polus intervenes to say that Socrates has tricked Gorgias into agreement, *Gor.* 461b2.¹⁶

The conversation between Gorgias and Socrates is “a pretty sedate affair, the more so by the comparison with the upheavals that follow,” writes James Doyle in “Socrates and Gorgias.”¹⁷ The *sedate affair* of examining what rhetoric is does seem like a trivial investigation, as James L. Kastely characterizes the dominant interpretation of the dialogue as something which “exemplifies bad philosophy or, at best, makes the trivial point that the practice of rhetoric can be abused.”¹⁸ Kastely's project is to deny the trivial reading of the dialogue and examine the possibility of moral complexity, and the philosophical significance, of unjust rhetoric. I view the dialogue as foundational to Socratic moral philosophy, and to the Socratic proposition that moral wrongdoing is bad

¹⁶Dodds' *commentary* says that the argument depends on the assumption that as one who knows to build is a builder, and one who knows music is a musician, he who has learnt what is just is just, *Gor.* 460b8, which is an assumption that Gorgias accepts, though Dodds notes that “to the modern reader it may appear a mere verbal quibble.” Gorgias accepts the view without resistance because, according to Dodds, from Homer onward moral conduct had been explained in terms of knowledge, and from knowledge of the good come good actions. Dodds writes that the *agathos* was the man “who did things well, and doing things well involved knowing how to do them.”

¹⁷ Doyle, “Socrates and Gorgias,” 1.

¹⁸ James L. Kastely, *Rethinking the Rhetorical Tradition: From Plato to Postmodernism* (New Haven: Yale University Press, 1997), 30.

for the moral wrongdoer in no trivial way, and that unjust rhetoric, and unjust rhetoricians, are an important illustration of this moral philosophy.

Barney sees Socrates' project as a normative analysis of social kinds, and so the critique Socrates offers is not, to Barney's view and contra Aristotle, a critique about the bad effects of abusive rhetoric, but is instead a project to identify rhetoric as a social kind, eligible for a normative, depersonalized meta-critique. Scientific disciplines divide particulars into kinds; a natural kind is a particular which corresponds to a group belonging to the natural world. A social kind is a particular which corresponds to a group belonging to the interests and actions of, in this case, human beings, but there can also be bovine social kinds, or amphibian social kinds.¹⁹

At the very beginning of the dialogue, *Gor.* 447c9, Socrates says to Chaerephon, "ask him," and Chaerephon replies, "ask him what?" Socrates says, ask him "what he is." This enigmatic statement and question hangs over the entire dialogue, for when Socrates asks Gorgias to explain what rhetoric is he is also implicitly asking Gorgias to explain what he himself is. I accept Barney's view of Socrates' project, and I take it further: I argue that the kind of practice to which rhetoric belongs is a moral *techne*, a particular particular, of a social kind, which is best categorized as belonging to the interests and actions of human beings, but which has moral import, and thus, the most meaningful kind. I imagine that Socrates would have no problem granting that rhetoric is a social kind, and my project is to clarify further, that rhetoric, to both to the audience and to the

¹⁹ Barney, 117.

practitioner such that it ought not be viewed simply through the harm criterion, pace Aristotle and the dialogues interlocutors.²⁰

Once Socrates asks Gorgias what rhetoric is, and once we realize that this is synonymous with asking Gorgias what he is, as a rhetorician and teacher of rhetoric, we see what Socrates is up to in terms of questions of ethics. Continuing to ask what rhetoric is confronts Gorgias with the possibility that what Gorgias has chosen for himself has deeper meaning. It seems to me that a dialogue about the harm that rhetoric causes to rhetoricians, in a dialogue with rhetoricians themselves, is deeply if not necessarily personal. In my view Socrates' investigation extends past the question of what rhetoric is and to the question of who the rhetoricians are, and whether if they are unjust in their practice can they be *eudaimon*. So long as the rhetoricians do not pay careful attention to the effects (the harm criterion) of their manipulative rhetoric then rhetoric is (a social kind) of deeply significant (moral, for the effects are to *eudaimonia*) practice.

Socrates' deeper questions imply a commitment to some kind of moral epistemology, which Dodds says was known to the Greeks of Gorgias' time, but which Socrates pushed further, "making explicit the unconscious presuppositions of traditional

²⁰ I take Barney to be locating her view within the purview of social constructivism, whose project is to reveal that some of our classificatory categories and practices, though they may appear natural, are actually socially constructed, contingent, and relative to the culture or social institution. Barney views rhetoric as belonging to this group, rhetoric as a constructed category, and further she sees the *Gorgias* as Socrates' attempt to reveal rhetoric to appropriately belong to this category. I do not see Socrates making this sort of classificatory argument, but my project, which views Socrates as arguing that rhetoric is a moral social kind, is not in tension with Barney's overall view. I diverge from Barney in that I see Socrates' investigation as to what rhetoric is, and ultimately who his interlocutors, all rhetoricians, are, as a very personalized investigation. If the unjust rhetorician cannot do well in life, and doing well on a deep and meaningful level is what Socrates means, then it is very personal indeed if what he is saying is that Polus and Callicles cannot, if they are unjust rhetoricians, be *eudaimon*.

Greek thinking about conduct: hence Gorgias accepts his view without a qualm.”²¹

Gorgias, in his reluctance or inability to explicitly state what he takes justice to be, indicates a reluctance to explain what the just teacher of rhetoric looks like, and what sort of person he himself is, and, “ There may not, in fact, be any instance when rhetoric has made the citizenry better,” *Gor.* 503e, also at *Gor.* 454c7 13 and *Gor.* 466a1. Perhaps he thinks of justice as a conventional arrangement between two autonomous, educated, self-directing parties, as when two boxers meet in the ring. If this is truly his moral epistemology, then moral responsibilities belong to the pupil and to the audience, analogous to a boxer in the ring. If the audience has failed to investigate the truth on their own, it is hardly the fault of the boxer, or rhetorician, who prevails in winning the battle. The pursuit of truth, then, is a political question to which all parties are personally responsible. The character of *Gorgias* finishes his dialogue with Socrates by exclaiming his own frustration, *Gor.* 463d9, and with a few short exclamations of agreement.

Perhaps one would disagree with my view that Socrates is interested in characterizing rhetoric as a moral *techne*, but it is clear that in his view justice or injustice implicates, or alters, the souls of the rhetorician’s fellow citizens, and never benefits the unjust rhetorician, a deeply troubling, personal scenario. To act unjustly is the greatest of evils, *Gor.* 474b4, and the *Republic* I.I 353c-354a, *Republic* I.I 354a9, and acting unjustly is worse than being wronged, *Gor.* 474b4, 473c1, 473e4-474b8, and the unjust person cannot be *eudaimon*, *Gor.* 473c2. If those claims hold true in the world broadly, within the category of rhetoric as a social kind generally, then they also hold for the unjust

²¹ Dodds, *Commentary*, 218.

rhetoricians in particular. Thus I accept this part of Barney’s view and accept that it applies to the rhetoricians who are investigating the benefits or harms, to themselves, of their own (social and moral) practice. Whether there is something non-socially constructed about the moral, that is, if Socrates believed that the *psuche* is involved in the moral in some way apart from the social construction of morality, I explore in the later section.

§1.1.3. My Examination of Gorgias on Rhetoric in the *Encomium of Helen*

In Ruby Blondell’s *Helen of Troy: Beauty, Myth, Devastation*, she describes the mythical figure of Helen as “the most beautiful woman in the world, and the most destructive, she is both the most in need of control and the least uncontrollable.”²² When Helen elopes to Troy with her lover, Paris, she triggers the great Trojan War. Gorgias purports to explain the truth about Helen, with four explanations for her actions. An explanation, or exoneration, has merit, because Helen can never be as Blondell notes “just a scapegoat, since the behavior for which she is to blame is inseparable from her infinitely desirable beauty.”²³

Gorgias’ explores the different powers at play in Helen’s decision, and offers three hypotheses for investigation:

²² Ruby Blondell, *Helen of Troy: Beauty, Myth, Devastation* (Oxford: Oxford University Press, 2013), ix.

²³ Blondell, *Helen of Troy*, x.

Helen was by force reduced
 Helen was by words seduced
 Helen was by love possessed. *Helen* 6.²⁴

Gorgias spends the most time in his analyzing the power of *by words seduced*, the rhetoric and the *logos* of the persuasion, and thus gives short-shrift to analyzing persuasion by rape (for who can blame the rape victim?) and short-shrift to analyzing falling in love (for who has not been powerless in love)?

In my overall reading I see this as Gorgias' primary argument:

Premise 1: The rhetorician uses *doxa* as a primary instrument. The *doxa* is the presupposed psychic state, a pre-existing state of the audience/listener/reader's opinion, which the rhetorician uses to his own purposes.

Premise 2: The rhetorician is concerned with communication and persuasion. It makes sense that he uses the *doxa* of the one listening to aid in the persuasion.

Conclusion: Truth is a combination of persuasion of the rhetor and the persuadable capacities of the listener.

Gloss: For example, great literature rests on a deception, which is not moral or immoral but the means to communicate. Rhetoric is also not moral or immoral, in part because belief is co-created by the listener's own *doxa*.

Although Gorgias writes that his objective is to exonerate Helen, it seems to me that his true aim is to exonerate. Charles Segal agrees, writing that the *Helen* may have served as a declaration of the professional rhetorician and the methods of the art of rhetoric.²⁵

Insofar as Gorgias develops an exoneration and analysis of rhetoric he is also developing a theory of rhetoric's power on the *psuche*, and offers many analogies. One particular

²⁴ Gorgias, *Encomium of Helen* in *The Greek Sophists*, trans. J. Dillon and T. Gergel (London: Penguin Classics, 2003), 76-85.

²⁵ Charles Segal, "Gorgias and the Psychology of the Logos," *Harvard Studies in Classical Philology* 66 (1962): 109.

physical analogy which he spends the most time on is the analogy that speech on the hearer is as drugs are to the body, *Helen* 13.

Just like a drug can be abused and the body can suffer, rhetorical tools can be used for good or bad ends. Just as a drug abuser can exploit drugs the rhetorician can exploit the reason and *doxa* of the listener.²⁶ Like a doctor or an artist, the rhetorician is aiming at something, and that thing which he aims at is good or bad, but the tool of rhetoric is morally neutral, and thus it is an “equal effort and mistake to blame the praise-able and to praise the blame-able,” *Helen* 1. Rhetoric, to Gorgias’ view, is an emotional tool on the *psuche* of the listener, but Gorgias says that the tools themselves have no moral content.

The listener is offered proof by hearing the speechmakers opinion, *Helen* 9, but the listener’s own conviction, *doxa*, in his own opinion is malleable, *Helen* 11. So, of course there is no reason to blame the rhetorician, for the rhetorician wields a tool in a neutral fashion that would not have the effect it has if it were not for the listener’s own emotional contagion, his own uncertain, impressionable *doxa*, *Helen* 10. For Gorgias the complicity between the *psuche* of the listener and the morally neutral tool of rhetoric means that one ought not blame the rhetorician.

Gorgias’ view depends on a few things being true, including the impressionability of the listener’s *psuche*, and that *doxa* is co-created by the influence of the rhetoric and the impressionability, and then malleability, of the listener. This is a plausible reading of how our opinions are formed, and so as a psychological characterization of the development of belief, it has merit. Gorgias’ view also depends on the *psuche* being some

²⁶ Segal writes that the biological allusion may be a reference to Empedocles’ Physical Theory, 99.

kind of tangible reality, something which coordinates with its surroundings. The rhetorician, aware of the complicity of the listener's *psuche* and the coercive power (morally neutral) of rhetoric can distort, via the medium of the *logos*, and transform *doxa*.

Rhetoric is thus the master of *doxa*, not by overpowering but by directing, toward the rhetorician's own preconceived ends, what Dillon and Gergel refer to as the "all conquering power of persuasive speech."²⁷ As Porter says, the rhetorician, through the *logos*, changes the listener's *psuche*. This characterizes rhetoric as a very powerful tool which acts, in some physical way, on the psychological entity of the *psuche*. The emotions and *doxa* also take on a physicality, as they too change. Porter says,

the *tropoi* probably refer more immediately to the ordinary ethical values upon which the stability of the *psuche* in society rests, but these values are forgotten under the impact of a powerful *opsis*, just as *phobos* drives out *nomos*. *Opsis* thus serves as the intermediary which transmits the purely physical stimulus to the emotional life of the *psuche*. Thus, the *psuche* can yield, irrationally, to the emotional, non-rational response.²⁸

This, says Porter, is what Gorgias believes about language and rhetoric, that it is being and nonbeing, it is communicated and impossible to communicate. In this way the *Helen* is not really about Helen, but it is about rhetoric, language, the *logos*, and ultimately blameworthiness: if some language bears upon Helen then all of it does, and she is not to be blamed for any of it. If some of Gorgias' language bears upon the reader then all of it does, and Gorgias is not to be blamed for any of it.

Thus the *Helen* offers a picture of the psychological mechanisms whereby rhetoric and *logos* effect the *psuche*, a force which bears upon all human beings including on the

²⁷ Dillon and Gergel, *The Greek Sophists*, 76.

²⁸ J. Porter, "The Seduction of Gorgias," *Classical Antiquity* 12, no. 2 (1993): 275.

listener in ways akin to Helen. In this way Gorgias succeeds in demonstrating the power of rhetoric, in the form of emotional and analytic poetry, illustrating the compelling power of the *logos* and its ability to manipulate emotions and alter *doxa*. Those who listen may be moved to pity Helen, as Gorgias writes, *tearful pity, Helen* 9. The reader may feel shuddering, *fearful shuddering, Helen* 10.

Of course, the listener is not enslaved or so entangled in forces beyond their control as Helen was, so how much less blameworthy is Helen, then, if the listener is likewise transformed? This perspective cultivates sympathy for Helen. The power of Gorgias the rhetorician-poet is a power over the listener's *doxa*, and in that way we the listener are co-conspirators, complicit, with Gorgias in our newly formed beliefs, our opinion, *doxa*, "being slippery and insecure, casts those employing it into slippery and insecure successes," *Helen* 11.

On Gorgias' account he is not to be blamed for our newly formed opinions, for his *Encomium* only brought out the latent tendencies of our *psuche*. The rhetoric is a force upon our *psuche* such that we have an excuse, for force is an excusing condition, but we have no cause to hold Gorgias more blameworthy than we hold ourselves. Helen, like the audience, is not blameworthy in any robust sense, because such is the nature of *logos* on the *psuche*. All the less blameworthy is the rhetoric itself, which is the morally neutral art or tool in between the rhetorician and the listener.

It is interesting to note that Gorgias is being very gender-equal in his encomium, in that his argument is about the nature of *logos* and the nature of the *psuche* is not gender specific. Thus the *Helen* is an account of how someone, anyone, can be persuaded to act

in a certain way, and is a demonstration, to the audience, of what that experience is like. A criticism may be that Gorgias' view entails that Helen, as well as the audience, have very weak agency. One wonders if Gorgias' audience, the men of Athens, would be insulted at being compared to the psychic strength, or frailty, of the likes of Helen. Perhaps this is alluded to at the end, where Gorgias gives a nod to the speciousness of his argument as an *amusement*, *Helen* 21. Gorgias' *Helen* is not the sort of thing an Athenian courtroom speechmaker would say, and so perhaps we are not only to be judging Helen but judging our own reaction, judging our own amusement, interest, enjoyment, and the transformation of our own opinion, *doxa*, about Helen's guilt.

After setting out these explanations Gorgias exonerates Helen, and he did this both by pitting one kind of language, defamatory, against another, respectful, such that language seems to collapse in on itself, and by drawing the audience into the effect. What I take away from this examination is that Gorgias sees language as driven by many forces, and effecting the *doxa* and *psuche* of the audience in ways that are akin to art, poetry, drugs, fire, music, but none of these things are themselves blameworthy for the transformation. Gorgias accepts that rhetoric has productive powers, that it can produce emotions like pity and anger, desire and resentment, and so it is a powerful art, but we are complicit, and so the art of rhetoric is morally neutral. Blondell writes that "Gorgias has not only defended the most notorious of mythic adulterers, but done so using arguments that undermine all moral judgement."²⁹ Porter says that one can only conclude, and Gorgias can only mean, that "his speech is no more than an imaginary itinerary, one

²⁹ Blondell, 175.

drawn (and erased) in the mind of its producer and consumer: going through the motions of going somewhere the speech literally goes nowhere.”³⁰

From a Socratic point of view, Gorgias’ perspective is probably absurd. Socrates contrasts *techne* with *alethia*, arguing that the tool itself is either geared toward or away from truth in such a way that the rhetorician is intimately implicated, being the director of the rhetoric, *Gor.* 454d, 459e. What is probably troubling for Socrates is the argument in the *Helen* enables the rhetorician to abdicate all responsibility.

§1.1.4. The Character of Socrates on Rhetoric in Plato’s *Gorgias*

For Socrates in the *Gorgias*, rhetoric has no subject of its own but is a practical skill in persuasion or flattery, not an art in any rich sense of the word, but a habit,

And I say that this sort of thing is a disgrace, Polus—for here I address you—because it aims at the pleasant and ignores the best; and I say it is not an art, but a habitude, since it has no account to give of the real nature of the things it applies, and so cannot tell the cause of any of them. *Gor.* 465a1-7.

Socrates proceeds to liken rhetoricians to shoemakers, physicians, and artists but who are less than what they practice. This devaluation and pejorative characterization are clear condemnations of rhetoric and rhetoricians. Characterizing rhetoric as flattery and a habitude invokes a quality of deception, misdirection, or unreflective response, as when a cook proclaims his food is nutritious because it is tasty, *Gor.* 501a1-501c5. Cooking, on Socrates’ account, is no more able to heal the ills of the body than unwise or untruthful rhetoric is able to address the moral questions which arise in, for example, a court of law.

³⁰ Porter, “The Seduction of Gorgias,” 280.

Here Socrates is offering that a true art must benefit humanity, and rhetoric clearly does not satisfy this requirement. When Callicles responds that there are some rhetoricians, public leaders, who care about the Athenian people, Socrates replies that it is not *necessary*, that is, it is not *intrinsic* to the practice of rhetoric, that a rhetorician care about the people, *Gor.* 503a-d. Perhaps there are two different kinds of rhetoricians, one who improves the soul of the Athenian people, but when Socrates asks for an example of such an orator, Callicles says that he has none to offer, *Gor.* 503b5-6. Socrates' rebuke is a moral condemnation, for if rhetoric can be used to make the people worse off, *Gor.* 502e8, then rhetoric has moral content.

The *Gorgias* shows the application of Socrates' moral lens on rhetoric. Socrates frames rhetoric as the art of influencing the *psuche* through words in all types of speaking, *Gor.* 453d-454, and since influencing can be done more or less badly, then the rhetorician can be a more or less good person. This is clear in Socrates' discussion with Polus about the technique of rhetoric in the law courts, where Socrates says to Polus that the best use of rhetoric is not that which enables the guilty man to be acquitted, but that which persuades the jury that the guilty man needs punishment. This is a paradox, because the defense ought to seek for an acquittal, and as Polus notes, the defendant seeks to avoid avoid punishment. Socrates counters that the Athenian polity wants to reinstate the integrity of law, order, *and* punishment, and as such the defense and the defendant must adjust their aims, *Gor.* 480c.

Socrates is unambiguously distinguishing himself from those who speak for the defense in a court of law, and from rhetoricians who charge fees for their services. This

recalls Socrates' statements in the *Apology*, 31d, that he purposefully separated himself from professional rhetoricians and purposefully set his own role as distinct from their schools, followers, teachers, and views. Socrates proposes that he has offered a valuable service to Athenians under the guidance of Apollo, which is a rebuttal to the accusation of corrupting the youth, *Apology* 24b-28a. Both Socrates and Gorgias (and Callicles and Polus) are referring to the good they do for Athenian society, and the important differences in what they each value. Gorgias does not have a method for discriminating between what is truly right or wrong, false or true, *Gor.* 459c, something that emerges clearly from Gorgias' reply to a series of questions put forth by Socrates about what role truth and falsehood play in persuasion, and rhetoric, *Gor.* 459c-460a. Gorgias is vulnerable on this point, for he supports the idea that rhetoric aims at creating opinion, *doxa*, in the listener as opposed to truth about the matter through a slow method of rational instruction.

The dialogue presents us with two "teachers" who have different opinions on their roles as teachers and on the meaning of what they practice. I see Socrates alluding to the interlocutors who are unjust agents wielding a moral *techné* without recognizing their own complicity in a wrongdoing, and that they should accept the punishment, in the form of the truth of the matter, which Socrates is doling out. This is a system that is alluded to at the end of the dialogue, *Gor.* 517a, leading the discussion back to the (moral) nature of rhetoric, which is where the discussion began.

§1.1.5. My Examination of Rhetoric as a *Moral Techne*

In what follows I characterize what I think Socrates is arguing for, ethically, in Plato's *Gorgias*, and how his argument fits into his overall moral ontology. My focus is on Socrates and Gorgias' debate about the nature of rhetoric, *Gor.* 448e6-461b2. I characterize what I take to be Socrates' overall moral proposition, what I call his *fulcrum argument*, and how that bears upon the questions of what rhetoric is, but more to the ethical point, who an unjust rhetorician is.

The fulcrum argument:

The excellence of the *psuche* is justice.

The just man with a just *psuche* lives well/happily, the unjust man badly.

The unjust man with his unjust *psuche* can never have *eudaimonia*.

Conclusion: Moral wrongdoing is of non-trivial harm to the moral wrongdoer.

Rhetoric is a *moral techne*:

Unjust rhetoric can harm the audience.

If the harm has moral content, then the agent of harm, and the technique of harm, have moral implications to the rhetorician.

Conclusion: rhetoric is a moral *techne* with moral implications to rhetoricians.

Contemporary moral philosophy aims to characterize moral reasons as *categorical*, meaning that moral reasons have a claim on us independent from our preferences, interests, and desires. The categoriality of moral reasons is achieved through careful reasoning, and so moral reasons are special reasons which confer special meaning. It is prudent for us to pay attention to the moral reasons which bear upon us, and we have prudential reasons to pay attention to them. I interpret Socratic moral philosophy in the *Gorgias* as an early attempt to introduce the categoriality of ethics and the reasons they bear upon the agent, in particular how the moral wrongdoer has special reasons to be

concerned for his own moral condition. Here I am working in the vein of Gregory Vlastos in his *Socrates: Ironist and Moral Philosopher*, to examine in what ways Socrates thought that moral wrongdoing is bad for the moral wrongdoer, an example of which is the harm of unjust rhetoric to the *eudaimonia* of the rhetorician.³¹

Towards the end of the *Republic I*, Socrates offers Thrasymachus an elaborate argument to the conclusion that "injustice is never more profitable than justice," *Republic I* 354a. We can connect this passage with that posed later in *Republic*, Book II, where Socrates says that "it is better in every way to be just than to be unjust" *Republic II* 357a, and connect them back to the *Gorgias*, where Socrates states explicitly that being a moral wrongdoer is worse than being a victim of moral wrongdoing, *Gor.* 474b4, that "doing wrong is worse than suffering it," and that "to do [a moral] wrong is the greatest of evils," *Gor.* 469b8-9. My overarching investigation into rhetoric being a moral *techne* is an inquiry into the Socratic proposition that unjust rhetoric, his main investigation in the *Gorgias*, is an example of injustice which is of no trivial harm to the unjust rhetorician: one cannot be *eudaimonia* if one is unjust. This proposition is noteworthy in a couple of ways, including in that it is a positive assertion on Socrates' part rather than a neutral investigation into the nature of rhetoric in which he takes no stand. It is also noteworthy in that Socrates is making a proposition about the harm of injustice rather than of the good of virtues like truth, wisdom, and justice, and so I see his investigation into the question of *why not be immoral* as opposed to just *why be moral*.

³¹ Gregory Vlastos, *Socrates, Ironist and Moral Philosopher* (Ithaca: Cornell University Press, 1991).

Vlastos writes that Socrates' true place in Greek thought was that he was "the first to establish a *eudaimonist* foundation for ethical theory ... a non-instrumentalist theory ... but something else ...," which is to say that Vlastos reads Socrates' as a *eudaimonist* moral theorist who accepted that happiness desired by all human beings is the ultimate *telos* of their rational acts.³² Vlastos goes on to argue that the *arete* of the man is virtue and that moral wrongs like injustice go against the *arete* of man.

In a section of Plato's *Apology* which follows Socrates' gloss on his own life and personal story, Socrates raises a point that echoes the fulcrum argument in the *Gorgias*:

But perhaps someone might say: "Are you then not ashamed, Socrates, of having followed such a pursuit, that you are now in danger of being put to death as a result?" But I should make to him a just reply: "You do not speak well, Sir, if you think a man in whom there is even a little merit ought to consider danger of life or death, and not rather regard this only, when he does things, whether the things he does are right or wrong and the acts of a good or a bad man, *Apol.* 28b

And this, from one of Socrates' longer monologues toward the end of the *Gorgias*, 508c3-7, "that to do wrong is worse, in the same degree as it is baser, than to suffer it, and that whoever means to be the right sort of rhetorician must really be just and well-informed of the ways of justice," Socrates argues in both dialogues, although his pursuit in the *Gorgias* is his entire pursuit in the *Gorgias*: considerations of reputation, personal safety, power and success are subordinate considerations to the pursuit of virtues like justice, and that there can be no happiness if there is no virtue.

In Plato's *Gorgias*, Socrates explicitly confirms that an inquiry about the nature of rhetoric is an *ethical* inquiry, "For I consider that a man cannot suffer any evil so great as

³² Vlastos, *Socrates*, 10. Vlastos writes that his reading of Socratic *eudaimonism* is in contrast with that of Irwin, who reads Socrates as arguing that the virtues are instrumental in happiness but not that virtues are intrinsic, and necessary, for happiness, see Vlastos 203.

a false opinion on the subjects of our actual argument,” *Gor.* 458a8-b1. Socrates’ initial questions to Gorgias, later taken up by Gorgias’ students, are an inquiry into the nature of rhetoric with, I pose here, Socrates’ deeper investigation is into the moral dimensions of injustice, that is, moral wrongdoing, of which rhetoric is the central example under examination. The nature of rhetoric has no particular intrinsic moral significance, but it has relevant practical significance for the characters in the dialogue, who are themselves rhetoricians. Socrates’ ethical inquiry is therefore a different approach from his addressing the merits or problems of, for example, superior speaking as a virtue in the city.

As is typical in Socratic dialogues, Socrates investigates a proposition by asking, *what is ____?* Where the responses and answers are subjected to an *elenchus* of definition, refinement, re-defining, and re-examining. The same fundamental framework is in play in the *Gorgias*, except that the fundamental question is, *What is rhetoric?* Socrates’ explicit claim that rhetoric is not a reputable *techne* but a kind of habitude or knack is to make a positive claim about what rhetoric is as well as to diminish what Gorgias claims is the “greatest good... whose persuasion aims at mankind’s freedom, to the domination of cities, the speeches of judges in courts, and ultimately deals with what is just and unjust,” *Gor.* 452a1, 454d8-9.

Socrates says that he is not trying to defeat Gorgias in debate but, as he emphasizes, trying to discover the truth about Gorgias’ craft, *Gor.* 457c-458d. I propose that Socrates’ investigation into rhetoric is an example of his investigation into his overall moral ontology and the fulcrum argument, which is that moral wrongdoing, of which

rhetoric is one example, is of no trivial harm to the moral wrongdoer. As such, the interlocutors are not morally neutral practitioners of a craft like pottery or cookery but practice a *techne* with actual moral import, as it persuades the audience to the *doxa* of opinion rather than to the truth of knowledge.

Gorgias agrees that when a *techne* is aimed at persuading the average man into thinking about good, bad, truth and justice, then the persuasion takes on a moral color and content, which is powerful and distinct from other non-persuasive, non-moral *techne*, *Gor.* 452e4-8. Gorgias offers the example of servitude to illustrate both the power of rhetoric and its ethical characteristics, and Gorgias says that successful rhetoric can make the doctor his slave, the trainer his slave, and the banker his puppet, *Gor.* 452e1-9.³³

I see Socrates, with Gorgias' help, drawing out the ethical characteristics of rhetoric, and thus the kind of ethical analysis the rhetorician is eligible for. Characterizing rhetoric a kind of *moral techne* by virtue of it being a *techne* with ethical import is problematic, and I will highlight the problems as this analysis progresses, but what is distinctive is how the dialogue frames rhetoric as something that can belong to the category of ethics. My overall view of Socrates' project is that he is arguing that if rhetoric is the kind of *techne* that can change someone's beliefs, perhaps in ways that are contrary to the truth, then the rhetorician is engaging in an act that is eligible for ethical scrutiny. To frame rhetoric as belonging to the category of ethics, and to frame

³³ Dodds, in his commentary on this section of the *Gorgias*, notes that making an analogy between the orator's power over doctors is consistent with the orator's influence over his fellow citizens in general, which was significant, 208.

rhetoricians as either achieving or failing ethically, is to make the dialogue an ethical investigation rather than a linguistic examination.

Rhetoricians, in giving an account of what rhetoric is, are also giving an account of who they are, an idea that Gorgias' pupils resist when the implication becomes that they are unjust in their careers. Insofar as an audience can be persuaded against the truth, rhetoric is an ethical tool such that the unwise, or unknowing rhetorician fails ethically without realizing his own failing, *Gor.* 454d, 459e. To say that rhetoric has moral content and that the rhetorician is engaging perhaps unwisely in an ethical art is anathema to the interlocutors who view rhetoric as, to harken back to Gorgias' *Helen*, morally neutral, or to Callicles ultimate point, something that the powerful rhetorician gets to wield regardless of the costs to the audience.

Socrates and Gorgias were both famed rhetoricians, but only Socrates was a famed moralist, and as such it is important to notice his ethical critique of rhetoric, and to better understand the interlocutor's resistance. Perhaps as a way for the *Gorgias'* interlocutors, who are technical and good, *Gor.* 504d5-6, to fully comprehend the fulcrum argument, Socrates inserts himself into the Athenian wariness and pushes the interlocutors further by characterizing their craft as one which has ethical implications. This is a surprising move, as rhetoric would not have been widely considered an ethical

technique—in fact, as discussed, the historical Gorgias argues that rhetoric is not, nor should it be considered, something eligible for ethical judgment.³⁴

David Roochnik, in his reading of the historical Gorgias, says that

[w]hile it is not possible to be specific, [the treatise attributed to Gorgias] *Encomium of Helen* suggests clues as to what a Gorgian conception of rhetorical *techné* would be...like many contemporary thinkers and orators, the value of Truth was not, to Gorgias' thinking, the purview of the rhetorician.³⁵

For the historical Gorgias that emerges in the work attributed to him, the truth of an argument is at best an adornment of *logos* and of the *doxa* which guides human behavior. Roochnik writes that Gorgias believed that the good rhetorician was able to “manipulate opinions in [Truth’s] absence.”³⁶ Dodds, in his commentary on Plato’s *Gorgias*, echoes the view that Gorgias’ *Helen* analyzed the power of persuasion and its efficacy, without a deep interest in the power of rhetoric in light of ethics.³⁷ On Roochnik and Dodds’ analysis, the Gorgian rhetorical school aimed at mastery of verbal *logos/praxis*, without an undue interest in the truth, as it were, of the argument.

³⁴ From the discussion of Gorgias’ *Encomium of Helen* in David Roochnik, *Of Art and Wisdom: Plato’s Understanding of Techne* (University Park: Pennsylvania State University Press, 1996), 74. On the historical figure of Polus, a follower of Gorgias and a teacher of rhetoric, see Dodds 1959, 11-12, and Ruby Blondell, *The Play of Character in Plato’s Dialogues* (Cambridge: Cambridge University Press, 2002), 31-37.

³⁵ Roochnik, *Of Art and Wisdom*, 74.

³⁶ Roochnik, 71.

³⁷ From Dodds’ commentary on the *Gorgias*, at 452e1-9, p202, Dodds refers to Gorgias’ *Encomium of Helen* as relevant to our contemporary understudying of what the interlocutors would have thought about the right aim of rhetoric. In *Helen*, Gorgias writes that the effect of rhetoric on the *psuche* is comparable to the effect of drugs on bodies, but as de Romilly says, “the effects on bodies is different: some bring disease, some bring life. In the case of speeches, some bring distress, others delight, some cause fear, others make the hearers bold, and some drug and bewitch the soul with a kind of evil persuasion.” de Romilly footnote 14 in Roochnik, 75.

Early in the *Gorgias* at 456c1-457c9, Plato seems to represent the historical Gorgias well, when Gorgias says rhetoric is neither good nor bad in itself but a *techne* to which the good was determined in the manner in which it was used, as when a boxer either uses his skills against a trained opponent or an innocent bystander.³⁸ This view is echoed in Plato's *Meno*, 95c, where Meno discusses the proper product of rhetoric:

That is a point, Socrates, for which I admire Gorgias: you will never hear him promising this [that orators can teach virtue], and he ridicules the others when he hears them promise it. Skill in speaking is what he takes it to be their business to produce.³⁹

In Socrates' account, if the rhetorician is not interested in truth he had better be interested in justice, and how rhetoric is either justly or unjustly used to change the beliefs of the audience, *Gor.* 454B1-458e9.

Roochnik's analysis of the historical Gorgias' rhetorical ideology is relevant to my calling attention to Socrates' first move in the dialogue, which was to draw out, perhaps in contrast with the historical Gorgias' own view, the concept of rhetoric as a kind of moral *techne*. Socrates' underlying premise is that belief ought to have some connection to what is true, so that if we are manipulated into believing something that is not true, we would have an ethical complaint against the speaker. For Socrates, if rhetoric can change

³⁸ The interlocutors are "Gorgianic figures," to use the description by Dodds, 192. Levitt refers to the Gorgian school as generally "rationalistic" in nature, that Gorgias and his pupils were familiar with Pre-Socratic philosophy generally, Levitt footnote 32 : 221, in Brad Levitt, "Platonic Parody in the *Gorgias*," *Phoenix* 59, no. 3/4, : 210-227.

³⁹ Plato, and W.R.M. Lamb, *Meno* in *The Loeb Classical Library. No. 165, : Laches Protagoras Meno Euthydemus*. (Cambridge: Harvard University Press, 1924).

an audience's beliefs despite the truth of the matter, then rhetoric starts to take on a distinctly ethical color.⁴⁰

Toward the end of his conversation with Socrates, Gorgias admits, in contrast with the other speaking interlocutors and in contrast with his own *Helen*, that rhetoric is a moral *techne*, that is, a craft with ethical characteristics, *Gor.* 454a1-454b9:

Socrates: ...and we shall be able to demonstrate that all the other arts which we mentioned just now are producers of persuasion, and what kind it is, and what it deals with, shall we not?

Gorgias: Yes.

Socrates: Hence rhetoric is not the only producer of persuasion.

Gorgias: You are right.

Socrates: Since then it is not the only one that achieves this effect, but others can also, we should be justified in putting this further question to the speaker, as we did concerning the painter: Then of what kind of persuasion, and of persuasion dealing with what, is rhetoric the art? [454b] Or do you not consider that such a further question would be justified?

Gorgias: Yes, I do.

Socrates: Then answer me, Gorgias, since you agree with me on that.

Gorgias: Well then, I mean that kind of persuasion, Socrates, which you find in the law-courts and in any public gatherings, as in fact I said just now; and it deals with what is just and unjust.

⁴⁰ From Roochnik: "...the goal of the (putative) Gorgian *techne* was to master *logos/praxis* and to do so in the absence of Truth," because a) knowledge is inaccessible, b) *doxa* (opinion) guides human behavior, and so c) Truth is an adornment of good *logos*. According to Gorgias, there is no accessible Truth, and so rhetoric is something which can mold opinion which may itself constitute the moral life," 73. The historical Gorgias is a certain kind of skeptic, the kind of person that deduces that if there is no ability to access Truth, then the best a person can do is form logical opinions and create a variety of *doxa*, creating opinion, reasoning, rules and logic.

Socrates gets Gorgias to agree that knowledge and truth are not the same, *Gor.* 454d9, and so the rhetorician is playing at a craft with ethical significance.

Because the rhetor can dangerously manipulate the listener *unjustly*, altering the audience's beliefs about what is true, Socrates gets Gorgias to agree that a rhetorician "must know what is just and unjust either already, or else must learn justice later," *Gor.* 460a1-461b5. If the rhetorician does not care about truth, then his persuasion can be unjust, and so the first Socratic move is to get Gorgias to admit, contrary to his own rhetorical ideology in *Helen*, that there is a moral characteristic to rhetoric, *Gor.* 461a5.

In light of these interpretations of the Gorgian views on rhetoric, Socrates' attempt at framing rhetoric as a *moral techne* is a surprising and notable move. Brad Levitt in "Platonic Parody in the Gorgias," believes that Gorgias acquiesces not by virtue of the sound logical argumentation Socrates provides, but by virtue of Socrates' emotional persuasion.⁴¹ Thus there is a bit of irony within this part of the dialogue, as Socrates is highlighting the moral power (and moral wrong) of unjust persuasion by employing it.⁴² The orators in the *Gorgias*, members of the Gorgian school, were perhaps hesitant to put their social standing on even more precarious footing, eligible for even more Athenian wariness, by agreeing that theirs was a moral *techne*.⁴³

⁴¹ Levitt, "Platonic Parody in the Gorgias," 214.

⁴² Levitt, 215.

⁴³ Kahn, 93. Of course the dialogue is not historical record, and the character of Gorgias is not identical to the Gorgias the orator and author of *Treatise on Helen*. It seems overly reductive to interpret the character of Gorgias as a complete fictional, perhaps parodic, representation of the historical Gorgias: likely there is something in between parody, fiction, and history at work in Plato's dialogue.

Athenians, despite their cautious attitude toward the power of rhetoricians, did not pass laws against training in rhetoric nor against the making of public speeches, and so the Athenian ambivalence at the time was perhaps precariously tipped in favor of the orators.⁴⁴ Doyle writes that “[rhetoric] fascinated the Athenian elites because success in court or on political stage depended entirely on the ability to persuade the *demos*.”⁴⁵ Gorgias, Polus, and Callicles all claim that rhetoric is determined to be successful if it succeeds in persuading the audience, and Doyle says “they differ merely in their squeamishness about making the assumption explicit—to themselves as well as to Socrates.”⁴⁶ Socrates’ overall argument is that if the rhetoricians accept that the *techne* of doctors and carpenters is such that they can be criticized for doing their job unjustly, then this is analogous to the *techne* of rhetoricians. If the rhetorician accepts that there is such a thing as true knowledge, and that they can persuade someone into believing the truth or not, then they can be criticized for the (un)justness of their persuasion.

Socrates appears to interpret Gorgias as having agreed with him, that rhetoric is about those things which are just and unjust, although Gorgias did not in fact cede that point (Gorgias says that rhetoric is about persuasion to what is just or unjust “in the law-courts and in any public gatherings” *Gor.* 454b7).⁴⁷ Later, when Socrates calls Callicles a

⁴⁴ Ober, 186.

⁴⁵ Doyle, “Socrates and Gorgias,” 7.

⁴⁶ Doyle, 8.

⁴⁷ Kahn argues that Gorgias was insincere in ceding the point to Socrates, in part Kahn says, because Gorgias, a foreigner, may have been concerned that he would be run out of Athens for not teaching students about justice. Levitt has a different interpretation, that the ethical use of rhetoric was not the foremost aim for Gorgias, but that effective persuasion was and ought be a rhetor’s goal, Levitt 212 and Kahn 79-84.

moral technician and calls rhetoric a moral art or *techne*, Callicles explicitly, unlike the other interlocutors, refutes that claim, *Gor.* 490e4. Perhaps the disconnect is semantic: Gorgias, and later Callicles, are talking about what Levitt calls *common-sense morality*, whereas Socrates, perhaps disingenuously, takes them to agree with his higher moral standard for rhetoric.⁴⁸

From arguing that rhetoric is a kind of moral *techne* the orators are then faced with a completely new perspective of their own careers. The argument would follow that the rhetoricians would have to, by virtue of their profession as rhetorician, inform themselves such that they are *sophos*, persuading audiences correctly:

Socrates: in fact, for my own part, I always regarded public speakers and sophists as the only people who have no call to complain of the thing that they themselves educate, for its wickedness towards them; as otherwise they must in the same words be also charging themselves with having been of no use to those whom they say they benefit. Is it not so? [520c]

Callicles: Certainly.

From this first move of framing rhetoric as a moral *techne*, Socrates progresses to his conclusion, that there is such a thing as an unjust rhetorician and that rhetoric done unjustly results in consequences to *eudaimonia*. While Athenians held deep criticism of the use, style, import, value, and harm of rhetoric, there was not much, according to Anderson, Kennedy, Kahn, and Ober, evidence that Athenians delved into the concept of

⁴⁸ Levitt interprets Gorgias and Callicles as talking about a “common-sense or consensual view of right and wrong, rather than [their] making a claim about the true nature of justice,” 212.

rhetoric as an ethical practice or moral *techne*, and so Socrates and the interlocutors are delving into unsettled territory.⁴⁹

Socrates and the hedonist Callicles discuss whether rhetoric has ever made anything better, that is, has rhetoric ever improved the condition of the city or any citizen, to which the answer is, perhaps not: “a change from the worse state in which he originally found them? For my part, I have no idea who the man is?” *Gor.* 503d7-9.

Socrates then revisits the initial claim agreed to by Gorgias, that rhetoric is a matter of persuading someone toward Truth, and the good man will keep his eye on the Truth of the matter before he persuades someone to a falsehood, *Gor.* 503e1.

Rhetoric, to Socrates’ view, is a form of, perhaps, indiscriminate social manipulation and control, wantonly used by rhetoricians who aim at persuasion and not the truth. Socrates pursues the definition of rhetoric unrelentingly but he never explicitly proposes that unjust rhetoric indicates an unjust rhetorician, but that seems to be what is looming over the entire pursuit. Who is the rhetorician if he unjustly manipulates an audience? The proposition at play in the fulcrum argument is that unjust rhetoric indicates an unjust rhetorician, and the unjust rhetorician is not free from the consequences to his own *eudaimonia*.

⁴⁹ Kahn explores the concept of whether rhetoric can be a *techne* at all. He begins with Plato’s *Ion*, and the discussion of whether poetry is a *techne*, at the passages where Socrates argues that poetry cannot be a *techne*, 105. There are three supporting arguments: One, that a poet cannot know *all the poetry*, he cannot be a broad, substantive expert on poetry. Second, that poets are inspired by muses who transmit power, enabling poets to entrance audiences, and as such they are in part conduits of the poetry and not the sole author (this point is echoed in the *Apology*, where Socrates says that poets have natural gifts, *Apology* 22b, artistic skill, and *sophia*) 3), that a *techne* is about a specific task, distinct, specialized, with a hierarchy of experts which the poets do not, and cannot, obtain (*Ion* 536d-542). Kahn notices that *Ion* never admits defeat, despite Socrates’ logical moves, 99-120.

Dissatisfaction with Rhetoric as a *Moral Techné*

Accepting that rhetoric is some kind of *moral techné* would require that the rhetorician contribute to his own moral education, but then the question arises, *What method of moral education ought he use?* A *moral techné* onto himself, perhaps? Or the moral instruction provided by a teacher like Gorgias? If a *moral techné* is something altogether different from flattery, manipulation, and empty technique that requires certain knowledge about morality, then a *moral techné* would require that moral knowledge is knowable and teachable, and it would need to be motivated as a thing higher in goodness than craftsmanship and other labor techniques.

If rhetoric is a special kind of speech that persuades the many to believe what the rhetor convinces them to believe, from *Gor.* 466c1-466d1, then it has a moral quality, which requires moral knowledge. This sets rhetoric apart from *techné* like cookery or pottery, and sets the rhetor apart from the cook and potter. Socrates, in getting Gorgias to accept that there is such a thing as unjust rhetoric, and getting Callicles to agree that the best *techné* of rhetoric is not distinguishable from the pursuit of human good, is making a moral claim, not just about the characteristics of rhetoric, but also about the knowability of truth and who ought to pursue it. It requires that the rhetorician ought then to agree that moral knowledge is desirable and attainable by those most skilled in the technique—this may be requiring too much of the average person. It would also require that moral knowledge ought to be something politicians and rhetors aim to master in the political/ethical life. Thinking of rhetoric as a *moral techné* is, for all these reasons, problematic.

Roochnik is not sure that *techne* is the right model for something like moral knowledge, but says that if we consider the dialogue to be a medium where conversations go back and forth, we can conceive of moral knowledge as something that is discursive and creative, and that the dialogue is a way in which interlocutors practice a *techne* about moral knowledge. Roochnik says that “The dialogue, I conclude, points to or suggests the possibility of some sort of nontechnical knowledge ... which does not have a determinate and analyzable subject matter, suffers no gap between *logos* and *ergon*, cannot be readily mechanically taught, and is thoroughly precarious.”⁵⁰ Precariously aiming at moral knowledge in a non-technical way still practicing something which has a moral quality.

In the absence of a positive moral framework, it is understandable that R. Hall and others see the concept of a moral *techne*, from which the fulcrum argument emerges, as unsatisfactory.⁵¹ Hall says that the concept of a moral *techne* fails because of the “impossibility of patterning a theory of the moral education of the individual along the lines of a *techne*,” in part because moral knowledge is not like carpentry, and moral technique, whatever it may be, is not a good analogy to woodworking.⁵² Hall, in his discussion of the *Gorgias*, says that despite Socrates’ offer to Callicles that orators are moral technicians, Socrates “probably does not consider it a realistic possibility.”⁵³

⁵⁰ Roochnik, 71.

⁵¹ R. Hall, *Platonic Readings, Platonic Writings* (New York: Routledge, 1988), 186.

⁵² Hall, *Platonic Readings, Platonic Writings*, 187.

⁵³ *Ibid.*, 203.

Later in the *Gorgias*, when Socrates says to Callicles that a rhetorician is a *moral technician*, Callicles is unconvinced, perhaps because he is skeptical that there is such a thing as moral truth to which a *techne* can be applied. At a time when philosophers were still debating whether moral knowledge is a thing that can clearly be known, Callicles' overall hesitation and incredulity seem prudent. None of this is at all, per the ensuing discussion with Polus and especially Callicles, what these rhetoricians envision for themselves, nor do they agree that it ought to be what they should envision. These hurdles come up later in the dialogue, when Socrates is faced with Callicles' explicit challenges as well as the implied disapproval via Callicles' silence.

This section of my analysis shows that Socrates was addressing justice and how rhetoric can be practiced justly or not, framing rhetoric as a moral *techne*. The larger philosophical question of whether one can know what is just, and whether, once one knows, one ought to then act justly, is a larger investigation than my point here. My point is that Socrates is explicating a moral ontology wherein the fulcrum argument is central, beginning with a proposition about rhetoric. As Doyle writes:

One may see all this, however, and still miss the full significance of the confrontation. It is clear enough that the question, What is rhetoric? so understood, assumes for orators all the importance of Socrates' great question: it is their way of life that is in the dock. Socrates is addressing orators...and orator was not just any old ethically-charged category: unlike, for example, beautician, it was fundamental to legal and political power in Athens.⁵⁴

Socrates' use of rhetorical flourishes, of employing emotion, of making a point relevant to the interlocutors in particular, are unsatisfactory from the viewpoint of logical analysis,

⁵⁴ Doyle, "Socrates and Gorgias," 3.

but are less unsatisfactory when viewed as a method of engagement with at least some historical figures who practiced the very art under examination.

Socrates was able to get Gorgias to admit that rhetoric changes the audience's beliefs without imparting knowledge, *Gor.* 454c7. Socrates got Gorgias to admit that the rhetorician ought to know what justice is, *Gor.* 466a1. Socrates got Gorgias to confirm that, if a pupil of rhetoric does not know what justice is, he will eventually learn it from Gorgias himself, *Gor.* 460a3-4. This is in contrast with Gorgias in the *Helen*, who claimed that rhetoric is morally neutral. Perhaps the contradiction, and the problem, have to do with what rhetoric produces and not what it is, or what impacts unjust rhetoric have on the justice of the rhetorician, which Gorgias never quite admits to, *Gor.* 459c8. But this is exactly what is at the crux of what I interpret to be the Socratic fulcrum argument, that "injustice never benefits the wrongdoer." Socrates and Gorgias are arguing past each other: Socrates has asked what rhetoric is, Gorgias has answered, but the looming implications, of what unjust rhetoric means for the unjust rhetor, are left, to Kastely's point, unsatisfied.

Socrates appears to be running two aspects of his inquiry together, the question of what is rhetoric, and the question of whether the rhetorician deeply understands what he is up to in terms of consequences to an audience as well as to himself. Running these two aspects of an inquiry together make sense if one accepts, as I think Socrates does, that the aim of a life is to pursue the good, and insofar as rhetoric is one example in which a life can go awry from the good, the rhetorician ought to care deeply about such a process. But the rhetoricians, most clearly Callicles, do not accept that a life is meant for pursuing the

good, as Callicles is famously the sort of character that pursues power, prestige, and acclaim regardless of what good, truth, or justice are.

Once we understand this aspect of conflating two concepts of inquiry into one, we understand the ethical import of the question *What is rhetoric?* and better understand what Socrates is *truly* asking the rhetoricians to examine, which is, *What sort of person are you?* If rhetoric is dis-analogous to pottery or cooking in these ethically charged ways, if rhetoric is a kind of moral *techne*, then coming to understand what the rhetoric is engaging in, morally, is coming to evaluate what constitutes the ethical life, of which a rhetorician is intimately involved in.

As a whole, the *Gorgias* is considered to be bad philosophy by critics, a dialogue, which, according to James Kastely, makes “the trivial point that the practice of rhetoric can be abused.”⁵⁵ Kastely, in his critique (and defense) of the *Gorgias*, refers to the overarching theme of the dialogue as one that demonizes rhetoric in a dissatisfying way.⁵⁶ My overall analysis of the fulcrum argument does not hinge on rhetoric being a *moral techne*, but I do hold that if the aim of rhetoric is to change the beliefs of another, then the act of persuasion of belief has moral content in ways that are very relevant to the interlocutors.

⁵⁵ James L. Kastely, “In Defense of Plato's *Gorgias*,” *Publications of the Modern Language Association of America* 106, no. 1 (1991): 96. Critics who have found the dialogue to be underwhelming, badly argued, and dissatisfying on the whole are notably Terrence Irwin, *Plato's Moral Theory*, 311; Dodds, 30. Brad Levitt, in his analysis of the *Gorgias*, argues that the dialogue functioned as a parody of the Gorgian school, and of Gorgias and his followers, and as such the dialogue is a Platonic critique of a mode of argumentation. Levitt cites Cicero's reflection on the dialogue, Cicero, *De Organum*, 1.11.47, where Cicero marvels at Plato, saying that Plato seems to be a consummate orator even as he was mocking them. Levitt sees the dialogue as a joke at Gorgias' expense, referring to how Gorgias the famous orator occupies a small portion of the dialogue and seems to be silenced rather quickly by Socrates' own verbal strategy.

⁵⁶ Kastely, “In Defense of Plato's *Gorgias*,” 96.

Socrates is addressing rhetoricians whose very lives, career, and power are bound-up with their success at oratory, and so the question, “What is rhetoric?” is no trivial matter to them. Socrates says to Polus, “Were you not this moment saying something like this: Is it not the case that the orators put to death anyone they wish, like the despots, and deprive people of property and expel them from their cities as they may think fit?” *Gor.* 466d1-6. This is yet another explicit reference to the Socratic view that rhetoric and rhetoricians are what J. Doyle calls “an ethically charged category,” such that the question of “What is rhetoric?” bears upon not just rhetoric as a moral *techné*, but what sort of human being the rhetorician is when he employs it unjustly.⁵⁷ Investigating what rhetoric is through an ethical lens is to investigate who the rhetorician is through an ethical lens, which allows the investigation of whether the rhetorician is aligned toward what is good, right, true and just.

Early in the *Apology*, at 17a, Socrates says: How you, men of Athens, have been affected by my accusers, I do not know; but I, for my part, almost forgot my own identity, so persuasively did they talk; and yet there is hardly a word of truth in what they have said. The fulcrum argument, again:

The excellence of the *psuche* is justice.

The just man with a just *psuche* lives well/happily, the unjust man badly.

The unjust man with his unjust *psuche* can never have *eudaimonia*.

Conclusion: Moral wrongdoing is of non-trivial harm to the moral wrongdoer.

The fulcrum argument proposes that unjust rhetoricians, of which accusers in the law courts count, are doing themselves ethical harm when they, for example, persuade an

⁵⁷ Doyle, 6.

audience or jury against the truth. This shifts the focus from harms to the audience to the harms to the particular rhetorician, accuser, and interlocutor, making Socrates' moral philosophy personal. By proposing that rhetoric might be an example of a moral *techné* which can be unjustly employed, then the unjust rhetorician ought to care about their unjust (moral) act, for the sake of his own *psuche* and his own *eudaimonia*.

The questions of *What is rhetoric?* and *Does it have moral import?* are questions intended to evaluate and understand the fundamental orientation of the rhetorician in his moral world. Does rhetoric have a *logos*, or is it a habitude, rather empty, such that Gorgias and his pupils have chosen a very empty way of life. I pose that confronting the interlocutors with the possibility that rhetoric is a moral *techné* confronts them with their own justness, and with their own jeopardizing of their *eudaimonia*. Because all the interlocutors are rhetoricians, the question of how they ought to live is intimately bound up in the perspective that rhetoric can be used more or less justly in ways that are significant to the interlocutor's own *eudaimonia*.

§1.1.6. My Examination of the Ethical *Psuche*

Ultimately the question of *why should I be moral?* is answered by Socrates with, *your eudaimonia depends on it*. But how is that connection made? The next section explores the foundational assumptions Socrates makes about ethics and the *psuche* in order that ethics is connected, via the *psuche* to *eudaimonia*.

Socratic *Eudaimonism*: Psychic Happiness and Unhappiness

My overarching aim is to take Socrates at his word when he claims, in various ways and at various times in the *Gorgias*, that doing injustice is worse than suffering it, and that to do a moral wrong is the greatest of evils, and to view this as a positive claim about how moral wrongs necessarily equate to unhappiness. I take seriously his explicit view on the virtues and *eudaimonia* as something that he explicitly knows, in contrast with his famous claim in the *Apology* 21b-d that he himself knows, and claims, nothing. I follow Vlastos' interpretation of Socrates' epistemic stance, that Socrates was sincere when he "avowed a positive conviction" that moral wrongdoing carries serious moral import for the moral wrongdoer, despite interlocutors like Callicles laughing at the idea.⁵⁸

I follow Vlastos in that Socrates, in the early dialogues, has a consistent moral ontology, that virtue and happiness are not distinct entities. This means that the Socratic view is that virtue is not simply instrumental for happiness but is intrinsic to it, and that injustice is indistinct from unhappiness, as explicated in Socrates' claim that the moral wrongdoer is worse off than his victim. Because Socrates holds that *eudaimonia* is the ultimate end for which everyone aims, the virtues like true belief, of which the rhetoricians in the *Gorgias* are accused of ignoring, is to be pursued because the virtues are intrinsic to *eudaimonia*.

⁵⁸ Vlastos, 5. Vlastos is especially keen on the affirmative moral philosophy that Socrates lays out in the *Gorgias*, and the affirmative position Socrates takes. Vlastos says that he cannot imagine that Plato would have put affirmative moral philosophy, that virtues like justice are internal to happiness, into the mouth of a character who maintained that he is suspending judgement about morality and the virtues, and that instead of this being a paradox on the part of Socrates, what is paradoxical are the way in which we interpret him, Vlastos 6-11. For Vlastos' view on Socrates, contra Irwin, virtue and happiness are not distinct, and that is the Socratic perspective I am untangling in the *Gorgias*, that injustice and unhappiness are indistinct.

The tension in the *Gorgias* is largely with the interlocutors argument with Socratic *eudaimonism*, as they argue and claim that they themselves are very happy, the tyrant on the hill is happy, and so happiness is disconnected from virtues like justice and truth. Callicles the hedonist is happy with power and the success he has in persuading the audiences toward the ends he desires, despite the truth or the justness of his persuasions. The question then, for Callicles, is not at all *Why should I be moral?* but instead, *How can I be powerful?* The Socratic position is that the most depraved rhetorician is still aiming at his own *eudaimonia*, except that he is wrong in thinking that he can successfully manipulate the audience toward false beliefs and maintain his own *eudaimonia*. This is an epistemic failing on the part of the rhetorician which must hold despite the orator Callicles stating clearly that he and other “moral wrongdoers” are in fact fairing quite well, *Gor.* 491e4-c6.

The *Psuche* Aims at What is Ethical

Socratic moral ontology rests on the *ethical* attunement of the *psuche*, that the human *psuche* or soul has a moral attunement.⁵⁹ By the 5th century, “pleasure, desire, reasoning, ethical reflection and ethical desires are all attributed to desires of the *psuche*,” as in Euripides’ *Ion* 1170 and Plato’s *Laches* 192c, and so Socrates’ conception of a moral is not new.⁶⁰ What is new and what seems most difficult for the dialogue’s interlocutors is that one cannot fare well if one is unethical, that is, one cannot be

⁵⁹ In the Homeric poems, only human beings are said to have (and to lose) their souls.

⁶⁰ David B. Claus, *Toward the Soul: An inquiry into the meaning of psuche before Plato* (New Haven: Yale University Press, 1981) 70-85.

εὐδαιμόν if one's *psuche* is not fulfilled by that which is ethically good. I will investigate Socrates' first underlying assumption, that the *psuche* aims at the ethical, first.

David B. Claus, in his exploration of what the *psuche*, the soul, meant to Greeks before Plato, offers several examples: as reported by Thucydides, the *psuche* is described as “the bearer of moral qualities such as justice;” in Herodotus, the semantic expansion of *psuche* as “that which animates the living” to “that which denotes an agent's moral character;” a Erupidean fragment refers to the *psuche* as “that which desires the just, temperate, and good;” Pindar's second *Olympian*, salvation is for “those who keep their *psuche* free from unjust acts;” Pythagoreans refer to the moralization of the *psuche*.⁶¹

On these accounts, the human being is a *self* with an ethically-oriented *psuche*, that is, by the end of the fifth century the human *psuche* was credited with qualities beyond life-giving essence and with moral sensitivities, which are developed in relation to the community's codes and practices. Socrates is not offering a radical view of an ethically attuned *psuche* as, by the 5th Century, in both philosophical and non-philosophical treatments, the *psuche* is a source of and has an affinity for the ethical.⁶²

⁶¹ All of these examples and references are given by Claus, *Toward the Soul*, 73-85.

⁶² Claus, 73, 75, 84.

All animated living things, in the fifth-century Ancient conception, have a *psuche*.⁶³ The concept of a *psuche* is what distinguishes that which is alive from that which is not:

having a soul is simply being alive; hence the emergence, at about this time, of the adjective ‘ensouled,’ *empsuchos*, as the standard word meaning “alive”, which was applied not just to human beings, but to other living things as well.⁶⁴

Given that the *psuche* is the distinguishing component of all living things, and given that it was commonly understood that the *psuche* is both the bearer of and is aligned with moral qualities, then we who are alive have a biological affinity for moral virtues.

The arguments about the value of the just life in the *Gorgias* and their presentation in the *Republic Books I and II* are made plausible against this background of the kind of rich conception of the self with an ethical *psuche*, which underlies Socrates’ ontology and gives clarity to the part of the *fulcrum argument* which proposes that the excellence of the *psuche* is justice. We see this in Herodotus, in Thucydides regarding Pericles’ funeral oration, and in Euripides *Hippolytus*, as Lorenz writes, “To educated

⁶³ “The meaning of *psuche* transforms in the 6th and 5th century Greek. By the end of the the 5th century, the *psuche* refers to the distinguishing mark of living things, to that which is the subject of emotional states, to that which is responsible for planning and practical thinking, and to that which is the bearer of moral virtues. Soul, *psuche*, is on the one hand what every living thing has, and on the other hand that which departs the body at the time of death. Thought of as one in the same thing, it was further developed in the 5th and 6th centuries to refer to the source of ways of acting and being acted on.” Lorenz, Hendrik, "Ancient Theories of Soul", The Stanford Encyclopedia of Philosophy (Summer 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>>.

⁶⁴ In ordinary fifth century Greek, the concept of a soul is what distinguishes that which is alive from that which is not, “having a soul is simply being alive; hence the emergence, at about this time, of the adjective ‘ensouled’ [*empsuchos*] as the standard word meaning “alive”, which was applied not just to human beings, but to other living things as well.” Lorenz, Hendrik, "Ancient Theories of Soul", The Stanford Encyclopedia of Philosophy (Summer 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>>.

fifth century speakers of Greek, it would have been natural to think of qualities of *psuche* as accounting for and being manifested in a person's morally significant behavior.”⁶⁵

While it is true that the alignment of the *psuche* towards the ethical can be distracted, confused, and overwhelmed, “it will eventually regain its clear conception of its own misalignment.”⁶⁶

The Ethical *Psuche* and *Eudaimonia*

Vlastos on Socrates’ moral ontology:

Plato never uses the “parts/whole” terminology for the relation of intrinsically valuable goods to happiness. Nor does he speak of them as being desired both for their own sake and for the sake of happiness. But what he does say can be put together in a pattern substantially like the one in Aristotle.⁶⁷

Vlastos proceeds to articulate how he sees the Socratic moral ontology as that which is best understood as Classical Virtue Ethics, that all non-moral goods are matters of indifference, and moral goods, the virtues, are the supreme principle of *eudaimonia*, the sovereign choice of all rational men. For this Socratic moral ontology to hold, the virtues cannot be merely instrumental to *eudaimonia* but they must be intrinsic to *eudaimonia*; *eudaimonia* must consist of goods like justice and just action, and we must be arranged in such a way that our *eudaimonia* depends on the virtues. Again, the question is, why is this the case? What are we that Socrates holds this moral ontology?

⁶⁵ Lorenz, Hendrik, "Ancient Theories of Soul", The Stanford Encyclopedia of Philosophy (Summer 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>>.

⁶⁶ Lorenz, Hendrik, "Ancient Theories of Soul", The Stanford Encyclopedia of Philosophy (Summer 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>>.

⁶⁷ Vlastos, 208. Vlastos is pulling dialogue from the *Apology*, the *Crito*, and the *Gorgias* to characterize what he sees as Socrates confronting competing values and arguing for the *eudaimonist* position.

My interpretation of Socrates' moral ontology is that he has a certain perspective on the human *psuche*, that the *psuche* is attuned to moral goodness such that moral wrongdoing affects psychic *eudaimonia*. In Socrates' characterization, significant moral wrongdoing is particularly impactful to the agent's *eudaimonia* because:

(A) The world is arranged in such a way that moral wrongs generate their own punishments: unhappiness and happiness.

(B) We, the *psuche*, are arranged in such a way that we can only fare well, happiness, when we act morally, virtuously.

The virtues are only one of the parts to *eudaimonia*, but the virtue-less *psuche* cannot be *eudaimon* if it is devoid of the virtues, and nothing can compensate us for the loss of the virtues.⁶⁸ Moral wrongdoing must be bad for the moral wrongdoer at some sort of psychic level, for otherwise Callicles would be right that he could pursue injustice with impunity, to which Socrates argues, toward the end of the *Gorgias*, that Callicles is quite mistaken.

The difference between people who are able to achieve *eudaimonia* and those who are not has to do with the condition of their *psuche*. That the *psuche* is the bearer of wisdom and excellence of character was well understood and embedded in ordinary Greek thought and language was likely accepted by the interlocutors, and so what the interlocutors are missing is the inextricable, ontological connection between the ethically aligned *psuche* and *eudaimonia*.⁶⁹ The interlocutors were not confused that the *psuche* desires the ethical. What they were skeptical about was that the unethical impedes the

⁶⁸ Vlastos, 211.

⁶⁹ Hendrik Lorenz, *The Brute Within: Appetitive desire in Plato and Aristotle* (Oxford: Oxford University Press, 2006), 38.

wrongdoer's flourishing. Socrates is connecting the ethical *psuche* (a familiar concept for the Greek thinkers of the time) to an objective *eudaimonia*, such that if the *psuche* is not achieving the ethical then the *psuche* is not faring well and the agent cannot be *eudaimon*.

The interlocutors would have been familiar with the concept of an ethical *psuche*, but what they are resistant to is that the ethical *psuche* is necessary for *eudaimonia*. When Socrates admonishes the interlocutors that ethical wrongdoers cannot be *eudaimon*, he is making both a semantic and an ontological connection whereby the ethically-aligned *psuche* is necessary for *eudaimonia*.

Socrates' interlocutors focus on the narrow sense of *eudaimonia*, highlighting how wrongdoers seem to be doing very well in their communities; they appear happy, and Calicles himself claims that he is very happy. Socrates is challenging this characterization, arguing that moral wrongdoers like Calicles, despite their own opinion to the contrary, are not *eudaimon*. Thus there must be something in the human *psuche* which suffers at moral wrongdoing even when the moral wrongdoer denies that this is the case.

The Unethical *Psuche* and Unhappiness

Socrates' central theses is that wrongdoing always harms the wrongdoer, and furthermore, doing wrong is worse for the wrongdoer than it is for the victim. Socrates offers Thrasymachus an elaborate description of why injustice is never more profitable than justice, *Republic* I 353c-354a, because only the just person's *psuche* is free from the

misalignment of having desired what is unjust. To phrase it using Platonic language, immoral acts result in unhappiness, because:

(A) *The world* is such that wrongdoers cannot be *eudaimonia*.

(B) *We* are such that if we are a wrongdoer we cannot be *eudaimonia*, thus virtue is necessary for well-being.

For Socrates, (A) and (B) are necessary truths about the way we and our world are.

Because the *psuche* is the subject of emotional states, it is implicated in the agent's wrong acts: it is the subject of the predicate "X is a wrongdoer." A person's morally significant behavior is an extension of the morally sensitive *psuche* and it is the qualities of the *psuche* from which moral behavior flows. Socrates, being a special kind of moral *eudaimonist*, believes that having an immoral character is bad because the perpetrator will be unhappy/fare poorly, and it is no trivial matter to fare poorly, because *eudaimonia* is the *telos* of all our actions.⁷⁰ *Eudaimonia* is desired by all human beings, it is the ultimate aim of all rational acts, and insofar as a moral wrongdoer is endangering his own *eudaimonia*, he ought to care very much about his wrongdoing.

I am working here in contrast but also in concert with Barney, who writes that Plato is working toward the invention of what we might call *philosophical social critique*,

[Plato's] aim is to offer something different in kind from the complaint that rhetoricians are bad people with bad intentions, who use rhetoric to do bad things. Instead, he undertakes to analyze and critique rhetoric as such—as a social institution, a well-defined kind (like a natural kind) with intrinsic features

⁷⁰ From Plato's *Symposium*, 205a2-3: "Of one who wants to be happy there is no longer any point in asking, for what reason does he want to be happy? This answer is already final." From Vlastos, "the final reason why anything is desired is happiness, and virtue is the only constituent of happiness—that virtue is happiness, the whole of it..." 208. From Johnathan Barnes, ed, *The Complete Works of Aristotle: Volume Two*, (New Jersey: Princeton University Press, 1995), For the Ancient Greeks, each living thing has a *telos*, a purpose to its life, and a function, which it ought perform in light of its *telos*. *Eudaimonia* is the activity of the *psuche*, in conformity with *ergon*, in creating a satisfactory *telos*.

of its own, evaluable independently of the intentions and qualities of its practitioners.⁷¹

Barney sees the *Gorgias* as a kind of satire, where the defense and the refutation of the moral import of rhetoric include ad hominem attacks, and a general analysis of rhetoric broadly. I have to contrast my investigation to that of Barney's at this level: I see the *Gorgias* as a deeply personal investigation of the interlocutors themselves, as ethical or unethical people, and not as an investigation into rhetoric as such. The investigation into the *psuche*, *eudaimonia*, truth and justice as they relate to interlocutors who are ignorant about such concepts is, to me, a personal, ethical exploration of the individual interlocutors and their own well-being. And so, my interpretation is that Plato's aim is to, in contrast with Barney, investigate whether these particular rhetoricians are acting unethically, and so I argued here that the connection between rhetoric, unjust rhetors, the *psuche*, and *eudaimonia* are the moral fabric of the *Gorgias*, and Socrates' main investigation.

We ought, I think, take Plato to be talking about people and the state of their souls, rather than about a social institution. My view is grounded in Socrates' overall moral ontology, where he characterizes the highest good as that which has an ethical quality, and then questions specific interlocutors about their own understanding of what they themselves are doing, ethically. He also characterizes the person who acts in accordance with the highest ethical good as *eudaimon*, something for which we all strive,

⁷¹ Barney, 118.

teleologically, as ethical beings, and which the particular interlocutors may be failing to understand.

What is unique in this dialogue is that, for Socrates, aiming at and achieving the highest ethical goods is sufficient for *eudaimonia*, while Aristotle acknowledges that the un-virtuous can be happy; in this way Socrates is an ethical *eudaimonist*, but Aristotle is not. In my view, Socrates' central project in the *Gorgias* is to analyze his claim that unethical rhetoric is an example of moral wrongdoing, centrally important to rhetoricians who are ignorant of this truth, who may be acting un-virtuously, and therefore must not be *eudaimon*.

Socrates says that a moral wrongdoer will fare worse than his victim, for the victim with a virtuous *psuche* can still be *eudaimon*. These radical claims, that it is worse to be a wrongdoer than a victim, and worse to act wrongly than be victimized, make sense against the rich conception of the ethically aligned *psuche* and the necessity of the virtues for *eudaimonia* such that the wrongdoer cannot diminish the power of his own unhappiness any more than he can divest himself of his ethically-oriented *psuche*. This makes sense in a way which is different from Barney's view regarding the social characterization of a *techne* like rhetoric, and moves rhetoric into an ethical, and very personal, category. It may be a social kind after all, but this is not my interpretation of Socrates' central investigation in the *Gorgias*.

My drawing-out of the moral impact of unjust rhetoric on the *psuche* comes from the semantic expansion of *psuche* in the fifth and sixth centuries, when it became common to use *psuche* to distinguish the animate from the inanimate (not just restricted

to humans), as well as the *psuche* as a locus of moral sentiments.⁷² My reference to *psuche* as that thing which an ethical alignment or misalignment to the ethical, such that the alignment determines unhappiness, connects Socratic moral *eudaimonism* to just or unjust actions, of which rhetoric is one (moral) example.

In contemporary thinking we have the tendency to understand virtuous actions as those actions that are good for all people concerned, in terms of consequences, results, respect, rights, and so on. But Socrates was interested in what makes a good life possible for the individual, and so virtuous and un-virtuous actions are analyzed at the level of the individual's own *psuche* and *eudaimonia*, a unique concern in and of itself, separate-able from good or bad consequences to others. The individualist focus and the connection between the individual's ethical *psuche* and the individual's own *eudaimonia* are, for Socrates, necessary: every individual ethical person is fairing well, every individual unethical person is not.⁷³

Aristotelean moral philosophers commonly see morality as necessary, but insufficient, for *eudaimonia*; Socratic moral philosophers thus have the responsibility of mounting a rebuttal, the position that morality is at minimum necessary for *eudaimonia*, and perhaps, under certain conceptions, sufficient. For Aristotle, *what a person should do*

⁷² For the historical discussion of the ancient concept of an ethical soul I refer to: Lorenz, *The Brute Within*, 25; J. Burnett, "The Socratic doctrine of the Soul", *Proceedings of the British Academy*, 7 (1915-16): 235–59; D. Furley, "The early history of the concept of soul," *Bulletin of the Institute of Classical Studies*, University of London, no. 3 (1956): 1–18.

⁷³ This sort of individualistic Socratic ontology works in conjunction with Socrates' intellectualism. Socrates is an intellectualist—he believes that one does not act wrong knowingly: if you know what is the highest good, then you behave accordingly. If you behave against the highest good, it must be that you have made a cognitive mistake about what the highest good is. For Socrates, knowing what is good, and knowing facts about the *psuche* and its affinity for the good, motivates the agent to behave in accordance with the good. No one would behave immorally if they properly understood the ethical attunement of their own *psuche*, and properly understood the tragedy of abdicating *eudaimonia*.

in order to be a happy, eudaimon, person, is not strictly speaking, a moral question, but for Socrates it is, and he is able to characterize his moral ontology by making a necessary, ontological connection between the ethically-atuned *psuche* and *eudaimonia*.

Chapter Two

Is the Socratic claim that moral wrongdoing is bad for the moral wrongdoer evident in empirical studies? In what way? How can the concept of bad for the moral wrongdoer, dubbed moral injury, be a useful concept for moral philosophy?

Preface

In this Chapter I discuss two different concepts, *moral trauma* and *moral injury*. The terms are technical terms used mainly in behavioral science contexts, and I analyze these concepts and introduce modified definitions. By refining the definitions of these concepts I endeavor to increase their linguistic precision and understanding, such that they can be useful for moral philosophy. I discuss the conceptual difference between *holding accountable*, *moral responsibility*, and *blame*. I argue that for certain moral wrongdoers, the morally injured, there is a legitimate opportunity for *moral repair*.

Psychologist Jonathan Shay, in *Achilles in Vietnam: Combat Trauma and the Undoing of Character*, coined the term *moral injury* to describe, categorize, and characterize the moral significance of the experience of being the perpetrator, witness, or victim of moral trauma, trauma which has a significant moral quality.⁷⁴ Shay's term

⁷⁴ Jonathan Shay, *Achilles in Vietnam: Combat Trauma and Undoing of Character* (New York: Atheneum, 1994).

captures the phenomenon of the moral soldier who, as a result of witnessing or perpetrating serious moral transgressions, forms new moral beliefs about his own moral character and/or the moral character of his world.⁷⁵ These new moral beliefs can be significant, such that Shay used the term *injury*: an injury which manifests as a deeply altered sense of the moral goodness of the self, of others, and of the world.

Shay's term moral injury is too broad, I will show, and will be more useful when it is narrowed and given precision. I retain the concept of an injury to the soul, or character, and from there craft a definition which retains this idea. The injury component of the term becomes more precise in my analysis, as it will refer to the wrongdoer's own character, on her own judgment as well as by the judgments of others. I will preserve the concept for perpetrators and culpable bystanders, but I remove witnesses and innocent bystanders, because this term must track moral responsibility, of which bystanders are excluded. Shay's concept captures a phenomenon under-examined in moral philosophy, that of the moral significance of being a perpetrator of a serious moral transgression, and the new moral beliefs about one's own character which are then formed. In this way I am preserving Shay's original insight about moral injury being a *soul injury* or *character injury*.

What Shay, Bret Litz, Shira Maguen, Nancy Sherman and other specialists in moral injury have thus far offered is a symptoms-based analysis, and what I am offering is a mental-mechanism analysis which refers to one's own moral responsibility for a serious moral transgression. The two together, symptom-based and mental-mechanism-

⁷⁵ Shay, *Achilles in Vietnam*, 21.

based, help to articulate what sort of unique moral judgments certain moral wrongdoers (and I will articulate how certain agents, commonly referred to as sociopaths, are excluded) form new moral beliefs about their own moral goodness. These newly formed beliefs include new judgments which are not wholly psychological but are pragmatic, like beliefs about diminished standing in the moral community, captured by the term moral injury. This refining of the term is conceptually important to moral philosophy. This clarity of scope enables the opportunity for new normative discussion about how we ought think about the goodness or badness of certain moral wrongdoers.

In Section One of this chapter, I will talk about:

- Defining a *moral community*,
- Defining *common morality*, and the *narrative view* of moral experience
- Defining *significant moral transgression*
- Characterizing what I mean by a *morally sensitive agent*
- Stipulating to the requirement of *moral responsibility/holding responsible*.

In Section Two of this chapter I will articulate:

- The differences between Post Traumatic Stress Disorder and moral injury
- The usefulness of the concept of moral injury for moral philosophy

In Section Three of this chapter, I will characterize:

- *Moral repair*. for certain moral transgressors, building on the work of Margaret Urban Walker's *Moral Repair*,
- Apply Walker's research as applicable to agents with moral injury

In the Conclusion I will address the normative questions that emerge, like: does it make sense that a moral transgression alters one's conception of oneself as good? Is it right that perpetrating a moral transgression is a reason for others to downgrade the agent's moral standing? What are the morally significant features of perpetrating a moral

wrong such that wellbeing and happiness are implicated? These questions pertain to an investigation like this, which connects moral beliefs, character, and wellbeing to moral transgressions and to the transgressor.

§2.1.1. An introduction, overview, and preparatory discussion of the subsequent sections

The Journal of Traumatic Stress has published several articles linking killing in combat with high instances of shame, regret, remorse, and debilitating psychosocial disorders, as compared to soldiers who did not kill.⁷⁶ What is emerging in these studies are a unique set of moral beliefs, which these moral wrongdoers formed about their own moral goodness, character, and belonging after recognizing their moral responsibility for a serious moral transgression. What is emerging from these studies is that among people who have all witnessed serious moral transgressions and trauma, like killing in war and the other moral horrors of the battlefield, soldiers who perpetrated significant moral wrongs held distinctly different moral beliefs about themselves, their actions, and the

⁷⁶ R. Strayer and L. Ellenhorn, "Vietnam Veterans: A Study Exploring Adjustment Patterns and Attitudes," *Journal of Social Issues* 31 (1975):81-93, killing another results in more symptoms of withdrawal, hostility, and life-outcome maladjustment vs. veterans who didn't; N. Breslau, G.C. Davis, & L.R. Schultz, "Posttraumatic Stress Disorder and the Incidence of Nicotine, Alcohol, and Other Drug Disorders in Persons who have Experienced Trauma," *Archives of General Psychiatry* 60 (2003): 289-294, killing poses a significant risk for PTSD; B.L.Green, "Defining Trauma: Terminology and Generic Stressor Dimensions," *Journal of Applied Social Psychology* 20 (1990):1632-1642, Eight categories for stressors that can cause PTSD, including "causing death or severe harm to another.; *On Killing: The Psychological Cost of Learning to Kill in War and Society*, ed. D. Grossman (Boston: Little, Brown & Co, 1995); H. Hendin and H.P. Haas, *Wounds of War: The Psychological Aftermath of Combat in Vietnam* (New York: Basic Books, 1984), The association between killing and severity of PTSD, noting that those who did not participate in violence did not [in their sample] develop PTSD.

moral world, in contrast with soldiers who witnessed moral horrors but did not perpetrate them.⁷⁷

In the studies and articles cited here, what I am calling attention to is that psychologists are differentiating between how perpetrating a serious moral offense is results in different and notable predictors for post-traumatic effect like alcohol and substance abuse, and PTSD, when compared to those who witnessed, but did not perpetrate, moral trauma. One such study, the *Maguen* study of US soldiers in 2009, shows that killing was a significant predictor of PTSD, alcohol, and substance abuse, as compared to soldiers who had witnessed killing but who had not killed anyone.⁷⁸ A statistically significant component in these studies include “causing severe harm to another,” or “killing another.” These studies conclude that soldiers who directly killed had unique, particular, and elevated psychological effects, as compared to soldiers who did not kill.

These studies are tracking something important for moral psychology: that perpetrators of significant moral wrongdoing have distinct psychological experiences as compared to witnesses of moral trauma. My investigation is into this phenomenon, and to articulate and characterize the concept of moral injury as a concept relevant to certain perpetrators of moral wrongdoing. I am building on Shay’s overarching thesis, that there is a unique and particular sort of moral experiences for some moral wrongdoers, a concept which will be useful for moral philosophy.

⁷⁷ As the discussion continues I refer to the studies cited in footnote two as well as the ongoing work by Bret Litz and Shira Maguen.

⁷⁸ Shira Maguen *et al*, “The Impact of Killing in War on Mental Health Symptoms and Related Functioning,” *Journal of Traumatic Stress* 22, no. 5 (2009): 435-443.

My work characterizes, articulates, and defines the scope of the moral wrongness of moral wrongdoing for certain moral wrongdoers, apparent in these psychological studies and relevant to moral philosophy. I posit that this concept is under-examined in contemporary moral philosophy and that it ought be seriously considered, for a variety of reasons. For example, the contemporary war narrative does not focus on the rightness or wrongness of killing *to the killer*, who by most accounts is a moral agent who may, predictably according to the studies cited above, develop certain moral judgments about their own moral goodness, their moral standing, their own moral self-worth. These beliefs connect to post-traumatic experiences of shame, grief, remorse, substance and alcohol abuse, and in some studies, elevated suicidal ideation in such a way that moral injury is a rich concept which involves moral judgements as well as psycho-social experiences. The rich concept of moral injury is significant to the behavioral sciences but also to moral philosophy, as it becomes clear that there are behavioral and psychological aspects to moral wrongdoing for certain morally sensitive moral wrongdoers, that have thus far gone largely under-articulated. When moral philosophers analyze the costs and benefits of moral wrongdoing, or the rights, duties, and transgressions of moral wrongdoing they ought account for moral injury and the psycho-social experience of being a certain kind of moral wrongdoer.

A theory about moral wrongdoing and its implications to the wrongdoer was offered by ancient Greek thinkers, Chapter One of this Dissertation. The Socratic claim is that moral wrongdoing is bad for the moral wrongdoer on a variety of levels, including prudential aspects like being seen as morally upstanding in the community,

and also psychological aspects like the negative judgments the wrongdoer holds about her own moral character. I characterize the Socratic concept reviewed in Chapter One in the contemporary language of *moral injury*, because my interpretation of the Socratic concept appears very much like what is being discovered in contemporary psychological studies.

Socratic ethics emphasizes *eudaimonia*, well-being, and looks primarily at the positive aspects: how virtues are connected to happiness, flourishing, or goodness. What is under-examined in contemporary ethics is how the vices, or moral wrongdoing, are connected to unhappiness, failing badly, and badness. It is my project here to recognize Socrates' emphasis on moral wrongdoing harming the wrongdoer by examining such a phenomenon, that of being a remorseful wrongdoer, and noticing the moral beliefs that are expressed.

This concept, of expressing regret, shame, remorse and in various ways *suffering* from perpetrating a serious moral transgression is what I claim Socrates was exploring in Plato's *Gorgias*, and what I argue that today can refer to as *moral injury*. If it is the case that moral injury is a concept which characterizes a particular moral agent, *the morally sensitive moral wrongdoer*, and a particular moral culpability, that of acknowledging responsibility (in a later section I specify *moral accountability*) for a *significant moral transgression*, then the concept can be integral to investigations of what moral philosophers determine a moral agent ought or ought not do.

§2.1.2. Moral Wrongdoers with Moral Injury: Narratives

In this section I incorporate actual first-person narratives by a certain subset of moral wrongdoers: soldiers who have perpetrated serious moral transgressions, hold themselves morally accountable, and express a series of new moral beliefs about themselves as a result of their actions. In her autobiography, *I Like My Rifle More Than You*, Army specialist Kayla Williams discusses her abuse of detainees, which she attributes to her now judging herself to be a morally bad person.⁷⁹ Williams says that her character has changed as a consequence of having followed orders to participate in interrogation torture, which quickly devolved into punitive torture. Williams says that her post-war psychological experiences are unique and separate from her fellow veterans. Williams writes,

What are we as humans, that we do this to each other? It made me question my humanity and the humanity of all Americans. It was difficult, and to this day I can no longer think I am a really good person and will do the right thing in the right situation.⁸⁰

She describes her beliefs about her character as being permanently, morally inferior, believing she does not belong with the moral community, and believing herself to be evil.⁸¹

Army Corporal Damien Corsetti says that he is experiencing moral injury as a result of his failure to prevent torture. Corsetti says that his experience at Bagram Military base, specifically as a military interrogator, was emotionally traumatic. He

⁷⁹ Kayla Williams, *I Love My Rifle More Than You* (New York: W.W. Norton & Company, 2005).

⁸⁰ Williams, 252.

⁸¹ *Ibid.*, 252.

describes torturing detainees as traumatic to him, that the emotional experience was so pronounced that he, too, suffers from PTSD and other emotional harms. He writes, “The cries, the smells, the sounds are with me. They are things that stay with you forever... I almost went crazy myself.”⁸²

Corporal Martin Webster was a British army soldier who filmed his fellow soldiers beating Iraqi civilians. The video was released publicly and it depicts severe, violent, spontaneous abuse of civilians. Webster is shown encouraging his fellow soldiers to escalate the violence. The video was widely seen on social media, and Webster was criticized publicly. He says that, “Perhaps my true character came out there. Perhaps I am an evil person ... It’s the ultimate humiliation of yourself ... When you’re carrying around effectively the soul of a murderer, you need to learn how to cleanse your soul, how to come to terms with that.”⁸³

I interpret these wrongdoers to be describing and exemplifying a different perspective of moral understanding. The narratives of these wrongdoers offers a model for thinking about wrongdoing that is contextual, subjective, and sensitive, rather than abstract and impartial. They offer concrete information about their particular case, not a rational schema to which there are no outliers.

I highlight these three narratives because they have written publicly about their experience being perpetrators of significant moral wrongdoing.⁸⁴ For all three, they now

⁸² Damien Corsetti, interview for “El Mundo” *El Mundo* 2007 trans. Juan Cole for *Informed Comment*, <http://www.juancole.com/2007/12/former-us-interrogator-recounts-torture.html>

⁸³ Martin Webster, interview for “BBC World,” *British Broadcasting System* November 12 2009 http://news.bbc.co.uk/local/cornwall/hi/people_and_places/religion_and_ethics/newsid_8356000/8356454.stm

⁸⁴ In the case of interrogation torture, for both Williams and Corsetti it was sometimes authorized and sometimes off the books; for Webster, the abuse was spontaneous.

question the aspects of moral character that moral theorists are interested in: they question their own moral goodness, they question their moral belonging in their world, they question if these actions indicate that they are, at their core, *bad people*. These experiences and these questions are represented in their medical diagnosis of PTSD and other anxiety disorders, as well as represented in their behaviors like substance abuse, domestic abuse, and difficulty reintegrating in the moral community.

These narratives open up new questions, questions about whether the particular voice of moral wrongdoers amounts to a unified perspective on (im)morality, and whether this viewpoint offers a useful contribution to ethics. I aim to tease-out moral injury as a concept that ought refer to be informed by these sorts of narratives, narratives by moral wrongdoers who have formed new beliefs about their own moral goodness. I began this chapter with these sections in order to motivate the idea that moral injury is a concept well-known in psychological literature that can and ought have a place in moral philosophy, and so I will move next to providing definitions, limiting scope, and further creating a framework, or typology, of *moral injury*.

Chapter Two, Section Two

§2.2.2. Typology of Moral Injury: Moral Trauma

Moral trauma is the term I use for a psychologically traumatic *event*, which has moral weight and content. The moral content regards the moral codes that the traumatic event violates, like the codes a community has created about sexual assault, murder, rape, child abuse. The *moral* refers to the rich moral concepts that are violated in the

event, like concepts of autonomy, fairness, justice, liberty. The *trauma* refers to serious transgression that is destructive in scope and depth, a harm to something which matters to us like our ideals, the value of persons, the value of self-direction, and consent.

Rape, for example, is a significant *moral trauma* in that it violates the victim's perspective of how she's valued in her world, it transgresses against her and our moral reverence for bodily autonomy, and it can transform the victim's judgments about her own moral value, worth, and standing. Rape, then, is morally traumatic on many levels in that it make us, and the victim, question, transform, and perhaps completely abandon certain moral values. This is what I mean by *moral trauma* being a term for a transgressive event that has significant moral content, and I will expand and refine this definition as this Chapter progresses.

The concept of moral trauma as I am using it has been used by moral theorists, feminist theorists, and psychologists, but the descriptions are varied. My working definition of moral trauma, then, is: a significant transgression against the moral beliefs, moral codes, moral expectations, held by most people. My characterization and definition of the term moral trauma is born from the conjunction of a traumatic experience and the significant breach of strong moral code. Moral trauma is thus the event, an event which is interpreted as having violated the codes of the moral community, and/or the moral sense of the agent.

Trauma, by definition, is extraordinary circumstances and extraordinary responses. The term *psychological trauma* is applied in a variety of context and by may different researchers. *The Diagnostic and Statistical Manual of Mental Disorders, 5th*

Edition, is useful, but it has been criticized as limiting; specifically, it limits the definition of psychological trauma to an event which threatens death or serious injury, something which critics say is too high a bar to capture what is actually experienced, psychologically, as significant and highly upsetting. This criticism is perhaps more about treatment, and so I will use the DSM-5 definition in order to lay the groundwork for describing the distinctive aspects of moral injury as a kind of traumatic effect.

The DSM-5 specifically defines trauma as:

Criterion A1: The direct personal experience of an event that involves actual or threatened death or serious injury, or other threat to one's physical integrity; or witnessing an event that involves death, injury, or a threat to the physical integrity of another person' or learning about unexpected or violent death, serious harm, or threat of death or injury experienced by a family member or other close associate.

Criterion A2: The person's response to the event must involve intense fear, helplessness, or horror (in children, disorganized or agitated behavior).⁸⁵

The term *psychological trauma* refers to the negative events that produce distress (Criterion A1), and it also refers to the distress itself (Criterion A2). The DSM-5 lists events that could count as traumatic, which includes physical and sexual threat, assault, disaster, attack, illness.

Psychological trauma, as the terms suggests, affects the psyche. If one follows a psychoanalytic model of the psyche, one conceives of it as made out of "parts" that interact with each other. Trauma disrupts the interaction between the various parts of the psyche in a way that causes extreme pain to the person. There is ongoing debate about

⁸⁵ American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. (Washington D.C., 2013).

whether a certain event has to satisfy diagnostic definitions of trauma in order to be, in fact, trauma. Trauma theorists reach their own conclusions as to whether a broader conception of trauma would better capture traumatic experience; I offer a conception of moral trauma, below, in the spirit of theorists whose work is broader than the current DSM-5 classifications. The remainder of this section comes from *The Principles of Trauma Therapy*, Second Edition, DSM-5. Because moral trauma does not have a settled definition I will pull from Jonathon Shay, Bret Litz, Nancy Sherman and Margaret Urban Walker (whose work I use here, but who do not use that particular term) to formulate a definition, so that I can talk about these authors and the concepts they refer to, like *moral repair* and *moral injury*.⁸⁶

Jonathon Shay's *Achilles in Vietnam* is a close exploration of the psychological experience of veterans, with a special emphasis on the significance of trauma in war. Shay looked primarily at combat trauma, which is rife with instances of rape, executions of innocents, torture of innocents, child molestation, and murder.⁸⁷ What Shay noticed was that there are nuances in experiencing the trauma of war, and that some trauma has a moral quality which is significant to the soldier in a variety of ways. What Shay noticed, both in his psychological practice and in his own lived experience as a soldier (Shay is a clinical psychologist and Vietnam veteran), is that particular moral emotions arise in

⁸⁶ The term moral trauma is a term I am using to refer to trauma which gives rise to moral emotions and moral injury, and although Shay does not use this term he does describe certain kinds of trauma as "psychological trauma which betrays deep moral codes," and "trauma which betrays our sense of what's right." These descriptions highlight the moral quality of certain kinds of trauma. I interpret Shay and others as characterizing a unique concept, moral trauma, part of that larger event we call psychological trauma: significant events which with components of serious moral transgressions. I use the term in the broad sense to apply to everyone who is traumatized by an event which has moral content: this would be victims, perpetrators, witnesses to the morally significant traumatic transgression.

⁸⁷ Johnathan Shay, "Moral Injury," *Intertexts* 16, no. 1 (2012): 58.

people who have experienced, caused, failed to prevent, or witnessed significant morally transgressive events. Shay uses a variety of terms like “the moral dimension of combat trauma,” “moral transgression,” “the moral dimension of severe trauma,” and the “moral construction of trauma.” On Shay’s view trauma of a moral kind challenges moral beliefs, like faith in the moral goodness of others and belief in the moral goodness of oneself. Even though Shay never uses the exact phrase *moral trauma*, I am going to discuss his work as if this phrase were part of his vocabulary.

There are lots of moral wrongs; stealing, downloading movies, not giving a dollar to the homeless person, urinating in the swimming pool: these acts will not, for most of us, be traumatizing in the ways articulated in the DSM-5 and contemporary trauma literature, nor would they rise to the level of moral trauma.⁸⁸ Moral trauma can refer to the violent act which destroys trust, like when a rape victim’s trust in the moral goodness of her peers is seriously undermined.⁸⁹ Moral trauma can refer to the morally transgressive act which raises the question about the other’s moral sensitivities, raised the question of whether codes of morality have any real meaning, raises the question of whether moral wrongdoings which go unpunished or un-repaired have any meaning.

⁸⁸ Moral trauma is not a term I have coined, and so I borrow it from theorists in other areas as a useful term for discussions in moral theory. In *The Geo-Politics of Hunger and Power*, Sylvie Brunel calls politically-driven unjust food distribution a moral trauma, because on every conception of morality refusing food aid, and therefore allowing death, is a serious moral transgression, Sylvie Brunel, “Lessons from the Kosovo Tragedy,” in *Action Against Hunger, The Geopolitics of Hunger and Power* (USA: Lynn Ripener Publications, 2001).

⁸⁹ Dostoevsky referred to the economic strain to Russia’s lowest economic classes as an economic and political wrongdoing which the peasants were acutely aware of, an injustice which they called morally traumatic. Joseph Frank, *Dostoevsky: the Mantle of the Prophet, 1871-1881* (New Jersey: Princeton University Press, 2002), 104.

Moral trauma is a central concept for feminist philosopher Margaret Urban Walker, although she does not formulate a clear definition of the term as such. For Walker, serious trauma which has a moral quality, like serious wrongdoing which raises the questions of moral codes and standing, serious moral transgressions which call us to ask, “What do we stand for?” are all components of what Walker calls “significant moral wrongdoing” and what I am calling moral trauma.⁹⁰

Walker includes the less acute and visible but still morally disrespectful and undermining actions and conditions in which some groups or people have been “persistently denigrated, subordinated, conquered, colonized, terrorized, exploited, or excluded by other groups of peoples.”⁹¹ These would be historic situations like colonization, slavery, genocide, etc. Philosopher Jean Hampton also discusses the concept of moral trauma in ways that I am using it, although she also does not use that specific term. For example, Hampton says that rape is *morally injurious* because the victim has had their autonomy fractured, their self-mastery annihilated, their agency denied. These are aspects of the experience which have a moral quality.

Jean Hampton’s characterization of the moral components of rape make it a fit for my definition. She says that

“insofar as [rape] is part of a pattern of response of many men toward many women that aims to establish their mastery qua male over a woman qua female

⁹⁰ Margaret Urban Walker, *Moral Repair: Reconstructing Moral Relations after Wrongdoing* (Cambridge: Cambridge University Press, 2006).

⁹¹ Walker, *Moral Repair*, 35.

.... Rape confirms that women are ‘for’ men: to be used, dominated, treated as objects.”⁹²

These moral criticisms may be internalized by the rape victim such that she shames and blames herself, holding herself in some way culpable for the assault, and morally diminished by it. The *moral trauma* of rape thus has and needs its own vocabulary which is beyond the scope of my investigation here, but I refer to rape as an example of something which I would apply the term of moral trauma to.

§2.2.3. Typology of Moral Injury: Moral Injury

This section is not a clinical analysis but a discussion of how *moral injury* as I define it can be useful in reference to the affect to the moral identity of agents who are morally responsible for moral trauma. I combine insights from the empirical psychological research with the insights of moral philosophy, especially as it relates to the psyche and character of the moral transgressor. My aim is to enrich and expand our moral vocabulary about the concept of moral injury, separate and distinct from concepts like moral dissonance and from PTSD.

Moral injury as it is commonly used is a term for the psychological effect of having been traumatized by *moral trauma*.⁹³ This is also a term which is used loosely, in a variety of ways depending on the theorist, but which I argued in Chapter One is what the ancient Greek thinkers meant about moral wrongdoing harming the moral

⁹² Jean Hampton, “Defining the wrong and defining rape,” *A Most Detestable Crime* (New York: Oxford University Press, 1999), 134.

⁹³ Note that from the Greek, trauma can mean an injury or wound. I explore the idea of trauma referring both to an event and to a wound later in this chapter.

wrongdoer. I define *moral injury* as something which is more specific than experiencing moral trauma: that it ought more narrowly refer to a moral agent's assessment of her own moral responsibility in a moral trauma, which includes new post-traumatic beliefs about her own morally diminished standing and inferior moral character. Thus, witnesses to moral trauma ought be excluded from the concept's scope, because their character is not implicated, being as they are not morally responsible for the moral trauma. I will expand upon this point, below.

Shay coined the term *moral injury* to describe, categorize, and characterize the psychological response to being the perpetrator, witness, or culpable bystander to moral trauma. Shay defines moral injury as,

“ [1] Betrayal of what's right—that's squarely in the culture; [2] by someone who holds legitimate authority—that's squarely in the social system; [3] in a high stakes situation—that's inevitably in the mind of the service member being injured, such as the love he has for his buddy. The whole human critter is in play here: body, mind, social system, culture.”⁹⁴

Shay is creating a new concept, for use in the trauma literature and diagnostic community, for the what-it's-like-ness of experiencing *moral trauma* and having a certain deep reaction to it.

William Nash, a senior Navy psychiatrist, psychologists Bret Litz and Shira Maguen, and other clinician-researchers, have diverged from Shay's concept in ways that I will be following and then refining.⁹⁵ Their research provides a comprehensive review of moral injury, including working definitions, the earliest research and

⁹⁴ Shay, *Achilles in Vietnam*, 178.

⁹⁵ Bret T. Litz, Nathan Stein, Eilenn Delaney *et. al.*, “Moral Injury and Moral Repair in War Veterans,” *Clinical Psychology Review* 29 (2009): 695-700.

characterization, and intervention modalities. They consolidate Vietnam-era research on soldiers who had perpetrated a serious moral transgression.

What is especially notable in this research is their findings that perpetrators have a significant increase in suicidal ideation and suicide rates. Litz and his fellow researchers track the emotions that are involved in suicidal ideation as a result of the soldiers having been a causal factor in moral trauma. The perpetrators deeply question their own ideals, their own ethics, and their moral standing in the community, something which traditional PTSD diagnoses does not capture. These researchers are tracking the moral components of post-traumatic experience in ways unusual to traditional trauma research, but important in my own work here.

It is in following Litz and Maguen's research, and in reflecting on the ancient Greek characterization of wrongdoers harming themselves, that I narrow the term *moral injury* even further, and offer my own definition:

Moral injury is the fracture of the basic moral assumptions, attitudes, judgments and emotions someone has about their own moral character. The moral emotions involve guilt, shame, and remorse in accepting moral responsibility for perpetrating/failing to prevent moral trauma.

I narrow the scope of the term from how Shay, Litz, Maguen, Sherman and others use it. These researchers use the concept of moral injury to include, for example, the disappointment in the moral goodness of the world that a witness to moral trauma can feel.⁹⁶ My narrowing preserves the concept as an *injury* to the character of a morally

⁹⁶ For Shay, the scope of his term moral injury regards the phenomenon of downgrading the moral goodness of the world, as a result of the moral horrors one has seen and experienced. Shay calls this a "betrayal of what's right," which deeply changed the strongly held moral beliefs that the agent has about the goodness of others," Shay, *Achilles in Vietnam*, 59. Leading moral injury researchers also include non-culpable witnesses and bystanders within the scope of the concept, see Litz *et al.*, "Moral Injury and Moral Repair in War Veterans," 700.

responsible agent agent, the injury being the result of having been a causal factor in some way (omission or commission) of a moral trauma.

This narrowing of the scope allows for the moral judgments that refer to the moral failure of an moral agent to make sense. My narrowing puts the focus back on what those early psychological studies have shown, that there is something unique in the psychological experience of being an agent, that is, being morally responsible for a serious moral trauma, like killing.

Shay says that the injury is at the level of character, the “undoing of character,” and “character brought to ruin,” and I retain this part of Shay’s analysis. My narrowing of the scope thus preserves this component of his concept, because if the injury refers to character, and if character is only implicated in a transgression insofar as there is moral culpability, then moral injury tracks moral culpability. Thus, my definition, scope, and use of *moral injury* elucidates:

- Scope of the act: *moral trauma*
- Scope of the agent: The kind of agent who committed/failed to prevent moral trauma
- Scope of the moral responsibility: An agent, as well as a moral community who hold beliefs about the transgressor’s moral standing, moral worth, and moral character such that these judgments are validated by the agent

What Shay, Litz and other specialists in moral injury have thus far offered is a symptoms-based analysis, and what I am offering is a mental-mechanism-based analysis, which outlines the moral quality of moral trauma and moral injury, and what sorts of beliefs the morally injured may hold. The two together, symptom-based and mental-mechanism-based, help to articulate how moral wrongdoing is involved in

wellbeing and its converse, suffering. My endeavor is to articulate the moral mechanisms involved in moral injury and the ethical implications to character and wellbeing, augmenting Shay's focus on psychological mechanisms and symptoms of moral injury.

Thus, my focus is to untangle some of Shay's language and propose that moral injury—pace the psychological literature which shows how the wrongdoer faces unique, particular moral challenges to her own character—is a distinct, useful concepts for moral philosophy.

Shay, Litz and others conflate *moral injury* with *moral disappointment*, as when a bystander loses all faith in the moral goodness of the world. Agents who witness moral trauma may understandably downgrade their moral evaluation of the moral nature of the world or the goodness of society, or the moral quality of a cause they formerly believed in, but these are not experiences which regard character and moral responsibility for moral trauma. These moral disappointments which refer to the world around are not, I argue here, symptomatic of moral injury. Moral judgments about the moral character of the world fit within the scope of the experience of PTSD, and can result in altered, maladaptive feelings of frustration, betrayal, contradiction, and extreme inner conflict, but moral injury on my account is a concept best preserved as a character injury, that is to say, a new set of beliefs formed about the moral defectiveness of one's own character as a result of perpetrating or failing to prevent a serious moral wrong.

Here I am diverging from Shay, Litz and others, to say that there needs to be a careful accounting of the attribution of responsibility for the moral trauma. If I

reasonably attribute the cause of a significant moral transgression to something that I did or failed to do, then I reasonably feel blameworthy such that a host of moral emotions and judgments arise. In contrast, my moral judgments about the world around me—and here Shay talks about how Vietnam veterans experienced a deep and severe loss of faith in the moral worth of humanity—is a different sort of moral judgment.

Thus is not right to say that post-traumatic moral emotions like questioning the moral world ought be subsumed under the same concept which captures the moral criticism I have of myself. Moral injury is at its most robust and useful when it refers to the judgments one has about one's own character in light of perpetrating and holding oneself morally responsible for a serious moral wrong. Someone who blames themselves for causing or failing to prevent a moral trauma is experiencing moral emotions like shame that have verifiable connections to the actual judgments held by the moral community, and this is a clear contrast to the sufferer of, for example PTSD, who may have witnessed serious moral trauma but does not and ought not blame herself for causing it.

If Shay, Litz and other researchers want to retain the diagnosis of moral injury for witnesses or bystanders of moral trauma who lose faith in their moral community, then the term is diluted and is interchangeable with other terms and concepts like disillusioned, let-down, disenchanted, embittered, soured, disappointed, and other terms for a lost of faith in someone or something outside ourselves. What is insightful, and I think that Shay, Litz and others recognize this insight, is that moral injury is particularly

useful and insightful as a term which captures the self-referential moral criticism born from actions which are morally criticizable.

Distinction Between Moral Injury and Moral Dissonance

Nancy Sherman has written extensively about the psychological dissonance of soldiers who believed that they were good, decent, and moral people but who acted in ways that later challenged those deeply held, and deeply valued, beliefs.⁹⁷ The psychological tension of believing yourself to be a good person yet recognizing that you've made a serious moral transgression is called *moral dissonance*, the state of experiencing oneself as a morally good person yet acting/having acted against deeply held moral beliefs.

The phenomenon of *moral dissonance* is different from *moral injury*. Moral injury is the recognition of having perpetrated or failed to prevent moral trauma, and then believing yourself to be significantly impacted in such a way that, among other post-traumatic effects and newly formed beliefs, downgrade your own moral character. The sufferer of moral injury is not dissonant; she believes, perhaps rightly (to be discussed below) that the moral wrongdoing impacts and implicates her moral character.

§2.2.4. Common Morality

In this section make an appeal to Bernard Gert's characterization of the common

⁹⁷ Nancy Sherman, *The Untold War: Inside the Hearts, Minds, and Souls of Our Soldiers* (New York: W.W. Norton & Company, 2011).

moral systems in order to contextually locate the moral wrongdoer who experiences moral injury.⁹⁸ The wrongdoer has transgressed against this public, other-regarding system, and her claims about breaking moral codes are primarily claims about codes of behavior which others hold her to. This moral wrongdoer experiences the post-traumatic regret, shame, remorse of her actions, holds herself morally responsible for moral trauma, and is downgrading her own moral character. This sort of agent has some sense of the moral rules of her community, evidenced by her moral sensitivity to her actions.

As Bernard Gert coined and used it, *common morality* functions as an umbrella term for a wide range of empirical observations about how people act and what they believe in relationship with customary moral codes. The *morality* component of *common morality* refers to the moral codes which govern resources and protections, to those duties shared in common as well as to those judgments held by the community. To be a member of the *moral community* is to recognize and participate in the *common morality* of said community, the moral codes a community generates for itself.

The concept of a *moral community* with a *common morality* does not exclude dissenters, nor does it imply that all members of a community abide by the same moral codes. The term *common morality* thus refers to the existence of a world of moral resolutions and its participants, known or could be known, chosen or could be chosen, by rational persons.⁹⁹ How the individual behaves and what she thinks about these moral resolutions is a different investigation. It is too general to assume that the term *common*

⁹⁸ Bernard Gert, *Morality: Its Nature and Justification* (New York: Oxford University Press, 1998).

⁹⁹ Gert, *Morality*, 6.

morality must refer to a unitary concept, one where there is a homogeneity of acceptable moral norms and behaviors.

Being a member of the *moral community* is to be expected to recognize and participate in the *common morality* of said community.¹⁰⁰ Members are entitled to the commonly shared protections, and are expected to adhere to commonly shared duties. Members of the *moral community* engage with each other regarding desert, obligations, duties, codes and mores in ways that are tested, judged, and interpreted. As such there is a broad allegiance among a given *moral community* to the moral aspects of the community's rules, laws and practices. The moral community's *common morality* regards deontological norms and consequentialist calculations, in the co-creation of shared moral beliefs and expectations.¹⁰¹

The codes of common moral systems are perhaps not justifiable, but they are *believed* to be so. This does not entail that they are unsophisticated, as it is the system that all rational persons, "given a desire for agreement and appropriate limitations on beliefs, favor adopting."¹⁰² *Common morality* is that system which *refers* to the codes of

¹⁰⁰ Gert says that theories on offer, like the Utilitarianism of John Stuart Mill, lost it's original task. After developing a set of principles, Gert says that Mill forgets that the point of moral principles is to "provide support for common morality." Because the aim of a moral theory is to set out and provide support for first principles initiated by contextual people looking for guides to actions, then it's no surprise that their principles connect to common morality if not providing justification for it, *Morality*, 18.

¹⁰¹ Gert is talking about a moral system and not a moral theory, a system which is contextual and narrative, bottom-up as well as top-down. Gert is talking about, pace Hobbes, a conception of commonly-used morality which is primarily concerned with the behavior of people insofar as that behavior affects others (whether they themselves plan to follow this guide is a different question). Gert is also talking about the way moral principles work and are applied in community, in contrast with what are correctly regarded as life principles like hedonism, egoism, and stoicism, *Morality*, 18.

¹⁰² *Ibid.*, *Morality*, 8.

conduct that applies to all persons, whose behavior is guided and judged by the system, that everyone is subject to.

Common morality is informal yet with boundaries, generally agreed-upon yet constantly refining itself. *Common morality* is more than just a code of conduct adopted by a group. For example, there is no such thing as *nazi morality*, but instead there are *nazi codes of conduct*; to call the code a moral code is to talk paradoxically. The members of the moral community generally want others to follow common morality, and know that others want them to follow it as well. It is not, as Gert is clear to emphasize, a closed system a person adopts for her own conduct: it is moral codes writ large, its largeness including one's immediate family, peers, and society, and applies to those persons who are co-creating it rationally.¹⁰³

My investigation is not into whether the moral community's moral codes are justified, or whether the individual is justified in fitting her own moral self-criticism to those codes, but instead I take it that these sorts of moral emotions arise in certain moral agents, even some agents who perpetrate serious moral wrongdoing. The agent who expresses moral injury indicates that she holds beliefs about the fittingness of the moral community's judgments on her moral standing in relation to her morally transgressive acts, and so this kind of agent is a particular kind, distinct from the morally dissonant agent, and distinct from the agent who has few to no moral sensitivities.

¹⁰³ "Morality is an informal public system applying to all rational persons, governing behavior that affects others, and includes what are commonly known as the moral rules, ideals, and virtues and has the lessening of evil or harm as its goal," Ibid., 12-13.

Thus moral injury is expressed only by a particular sub-set of moral wrongdoers, those sufficiently morally sensitive people who connect to their community's moral codes and recognize themselves as moral codebreakers. Corporal Martin Webster, the British army soldier who filmed his fellow soldiers beating Iraqi civilians, expresses his shame at having participated a morally transgressive act of recording unwarranted violence. Webster expresses a recognition of the moral transgression, a recognition and an agreement with the his community's criticisms. Despite being an agent of serious moral wrongdoing, he is able to reflect back on his actions and recognize his transgression against common morality. Webster is thus the particular kind of agent in a particular environment, the moral transgressor who, for whatever reasons at the time, in hindsight recognizes and respects the moral codes he has broken.

It is not a foregone conclusion that the moral community's judgments, which can include blaming the wrongdoer for her significant moral transgressions, are correct. The literature on *holding responsible* and *blame* asks questions like, what ought the moral community believe, and it is no way clear that the moral community ought blame a particular moral transgressor. It is also not a foregone conclusion that the agent is not free to define for herself to what degree she bears moral responsibility. The moral wrongdoer has the freedom to walk away from the blaming attitudes of her community; I notice that on my account, moral injury is experienced by those moral wrongdoers who do not not shirk moral criticism. I remind the reader that the concept of moral injury is richest and most useful when it refers to the condition of believing oneself to be morally diminished in some way as a result of moral responsibility for a serious moral transgression. This

belief-state excludes those agents who are insensitive to the richness of a moral community and moral codes. The morally injured are sensitive to moral codes and values, and so those who are insensitive are not best categorized as morally injured.

Moral injury thus refers to sort of agent who has some sort of moral sensitivity, and who recognizes common morality. This requires that the moral wrongdoer meets a minimal threshold of sensitivity for the moral codes of her community, whatever community that may be. This sensitivity is stipulated in my discussion of *moral injury*: the moral wrongdoer who notices that she has transgressed against a moral code has a threshold of awareness of that code, as well as some acceptance that those moral codes and moral standards have legitimacy, authority, and purchase on her. The *injury* of moral injury refers to the downgrading of one's own moral worth in light of the moral wrongdoing, and so with moral injury the wrongdoer is embedded in and has validated common morality, to some degree.

Moral Reactive Attitudes

What moral injury indicates is that the moral wrongdoer is contemplating the moral reactive attitudes of her peers. By establishing that there exists such a concept as common morality which takes seriously social moral criticism, there is then a moral concept from which to begin talking about what the moral wrongdoer with moral injury believes to be true about herself.

When I judge myself for transgressing against common morality, I am expressing my sensitivity to some conception of my own moral responsibility for moral trauma, and

my sensitivity to the moral judgments of my peers.¹⁰⁴ If I validate the moral codes of my community in such a way that I take seriously certain moral judgments like the shaming attitudes of my community, then I have met some threshold of sensitivity to the codes of common morality. In his essay “Freedom and Resentment,” P.F. Strawson reminds us that we in fact care about the actions of other people, especially those people who we deem important and whose actions bear upon our own lives.¹⁰⁵ We also in fact care about the reactive attitudes others have about us, in particular strong moral attitudes like resentment, forgiveness, and blame. I am accepting Strawson’s argument that the very same considerations of our more general, personal reactive attitudes are scaleable and apply to what he calls *moral reactive attitudes*. These moral reactive attitudes matter to us because we are intimate, inter-personal, psychological beings in community, embedded agents, to use the term from the previous section.

Those of us who are members of the moral community have certain *moral reactive attitudes*, about our and other’s judgments about a significant moral

¹⁰⁴ *Moral responsibility*, as conceived of as in the everyday conception of being a member of the moral community, and bearing moral responsibility for a moral wrongdoing. The characterization of morally responsible is to distinguish it from non-moral responsibility, for example, there is a (non moral) causal connection in having committed or failed to prevent a morally significant action, and there are (moral) judgments, like praise and blame, are appropriate. Moral responsibility is thus some sense of a causal connection to a morally significant act, inuring moral judgments in the actor, from Fischer, John Martin and Ravizza, Mark, *Responsibility and Control: A Theory of Moral Responsibility* (New York: Cambridge University Press) 1998.

¹⁰⁵ P.F. Strawson’s “Freedom and Resentment,” first published in 1962, is an influential contribution to the “free will” debate as it was generally understood around the middle of the twentieth century. As Strawson frames it, the classical free-will debate can be described in terms of the dilemma of determinism, the thesis that everything which occurs in the world is subject to causal laws and antecedent causal necessitation of actions. The classical dilemma is that free will and moral responsibility seem impossible if our actions are causally precipitated, but also, they seem impossible if our actions are not causally precipitated, often called the *skeptical conclusion*, in *P.F. Strawson: Freedom and Resentment and Other Essays* (London: Harper & Row, 1974): 34.

wrongdoing.¹⁰⁶ These various sorts of reactive attitudes, says Strawson, are all “humanly connected,” and eligible for the same rational and psychological constraints, and we the members of the moral community have a practical commitment to moral reactive attitudes.¹⁰⁷ In this way they are a fundamental thread of human nature. People are products of communities, lived spaces where moral codes are constructed, tested, deconstructed, and redrawn.

I am narrowing the scope of the concept of moral injury onto to certain particular wrongdoers, those who form judgments about their own character and moral worth in light of a) holding themselves morally accountable (to be discussed and differentiated, below) for perpetrating/failing to prevent moral trauma and b) being sensitive to common morality, the moral community, and to the moral reactive attitudes their peers have. In this way, the moral community, in conjunction with the moral wrongdoer, co-create a normative presupposition of all moral action in the social sphere, formally and informally, as well as socially and individually. Ultimately, then, *moral injury*, on my account, is limited to those moral wrongdoers who have a certain threshold of sensitivity to the moral reactive attitudes of their moral community.¹⁰⁸

¹⁰⁶ It is Strawson’s view that the main parties in the compatibilist/incompatibilist dispute have failed to identify what really matters to us as people in the moral sphere. Strawson’s attempt at re-framing the debate accepts that the incompatibilist is correct in holding that something vital is left out of the account of moral responsibility in a determinist world. Strawson shifts away from talking about moral responsibility as something which hinges on “freedom” and “responsibility,” and instead turns to what holds a person responsible. It is less about conceptual analysis and more about a descriptive account about actual human moral psychology.

¹⁰⁷ Strawson, 23.

¹⁰⁸ Excluded from the scope of this discussion are agents who do not meet a minimal threshold of sensitivity toward the moral reactive attitudes of their peers, and/or who do not claim moral responsibility for their part in a serious moral transgression.

There are three conceptions of morality that dominate philosophical analysis. The first conceives of morality as a social concept, accepting the claim that there is no deep concept of morality for the individual when they are alone in a cabin in the woods or stranded on a desert island. This view, which is the dominant approach to moral theory, privileges the analysis of moral agents and their decisions in relation to others as well as the decision procedures they ought to use that engages with justice, equity, inclusivity, and consequences. With this view, morality is a social concept such that some costs are outweighed by the benefits for many.

The second conception of morality that dominates philosophical analysis regards morality as a dogmatic concept where moral rules apply because they are *sui generis*, meaning that they are applicable regardless of the agent's participation in society. This entails that there is a deep concept of morality whether the agent is in community or alone on an island because the rules apply and always bear on the agent's actions.

This is not to say that these two conceptions of morality—the social and the individual—are distinct. In fact, they are importantly intertwined. The intersection of these two concepts is found in a third conception of morality, Ancient Greek moral theory, where the development of the question *what kind of person am I?* is central. For the Ancient moral philosophers, this question is relevant to the cabin-dweller or island castaway, because rightdoing and wrongdoing relates to the agent's own well-being.

These theories attempt to articulate and justify themselves within their own moral theory. References to theory and principles are invoked in an attempt to ascertain where moral judgments come from, an attempt to create or discern a neutral moral code,

impersonal, applicable to anyone. These theories are agent-neutral action-guiding theories, such that one can develop a kind of moral expertise. The questions that I am interested do not regard whether an agent has a developed moral expertise, only that she believes she is engaged in moral ideas in some way, in light of the moral codes she accepts.

The on-the-ground view of the embedded agent engaging with common morality is in contrast with looking at moral injury through a lens of a strictly code-like theory of moral systems. Moral judgments in the lived experience of moral agents who experience *moral injury* are, on the view I reflect on here, not strictly reasoned and comprehensive in ways articulated in, for example, Aristotle's *Nicomachean Ethics*, but invoked as they are needed and derived from individual experience, from laws of the community, from the particular and general picture of the moral world. Instead I appeal to the diverse way that moral systems are experienced by the individual. In reflecting on Bernard Gert's emphasis on the subject within a community of moral code-makers, and in reflecting on the narratives the moral wrongdoer tells about her own moral wrongdoing, I am appealing to way of looking at *moral injury* as a socially created concept. As Margaret Urban Walker notes, there is a way to describe us as socially embedded subjects whose moral judgments are not best represented by code-like theories of a Utilitarian, Kantian, or Contractarian kind, but are co-created as part of our lived experiences in community with each other.

Walker notes that this methodology is in good company, a company of moral philosophers sensitive to broadening moral analysis past a view of a neutral agent acting in code-like ways,

In the last few decades a remarkably diverse collection of moral philosophers—Aristotelians and Wittgenteinians, casuists and communitarians, pragmatists and feminists, Hegelians, postmodernists, and assorted others—have thought [to look at morality with an eye to lived experience].¹⁰⁹

To say that the moral judgments that the morally injured holds are *fitting* with the moral codes she has experienced is to recognize her contingency within her moral community. Her moral self-criticism is particular to her experiences, her relationships, her problems and habits, her remarkably diverse experience which are both unique and yet universal.

As Walker says, an individual's narrative about her specific values and commitments (personal, religious, professional or cultural) may matter crucially to her,

...Maintaining integrity and coherent moral self-understanding over time...To know what general values or norms mean in situations now requires appreciating how these have been applied and withheld, circumscribed, and reinterpreted before within individual, social, or institutional histories. So adequate moral consideration needs to follow these writers of identify, relationship, and value, to see how they can go on, and whether it is better or worse that they do.¹¹⁰

In this way the subject who experiences *moral injury* and her *moral community* are inextricably intertwined, relevant, and co-created.

This *narrative view* privileges the moral subject *in relationship* with common morality, in relationship with value systems she may not agree with, in relationship with a

¹⁰⁹ Walker, *Moral Repair*, 126.

¹¹⁰ *Ibid.*, 128.

perspective on her value system in light of her own actions. This narrative view takes seriously her perspective on her own moral identity, and the ways that her narrative connects to the principles and concepts defined by the broader community. I follow this view because the morally injured is a contextual being, involved in her personal drama in ways which have individual and universal moral features. On the narrative view the question is expanded to, what kind of person do I believe myself to be *in relation to* the moral codes, and the moral reactive attitudes, that I and others hold? The moral philosopher who investigates the experience of moral injury via an exploration of whether the moral codes that were broken are relevant, valid, and justified, or whether the costs of the code-breaking outweigh any benefits, is missing an exploration into the narrative of what these agents are experiencing.

A moral codes-focused philosopher may be interested in investigating whether an agent *ought* feel shame as a result of her having cheated on a test. Emphasis is on the codes of morality, the moral weight of certain outcomes, the undermining of trust in moral codes, and other rules-based or consequences-based analysis. For my investigation I follow Gert and Walker and the “diverse collection of moral philosophers,” to use Walker’s phrase, who are also interested in the agent’s narrative of herself in relation to her having perpetrated a significant moral wrongdoing within a moral community. This view takes into account the rules she broke and *how she feels* about breaking them, the consequences she’s wrought and *how she views* the subsequent judgements of others. My emphasis is on the narrative because my starting-point is the psychological studies which

show that there is something distinct and notable happening in certain moral wrongdoer's psycho-social experience.

The Fittingness of *Moral Reactive Attitudes & Moral Accountability*

I offer the description of *holding oneself accountable* to the moral reactive attitudes to ground much of what is interesting in moral injury, that this kind of moral wrongdoer holds herself accountable to the moral reactive attitudes to such a degree that her view of her own moral goodness is implicated. The designation of *holding oneself accountable* is a two-part relation, where the moral community holds the wrongdoer accountable (as expressed by their moral reactive attitudes), and the wrongdoer holds herself accountable to those attitudes (as expressed by her taking them seriously, including downgrading her own moral standing).

Moral injury and *moral dissonance* share the same component of acknowledging responsibility, where the agents agree that they are in fact morally responsible for the moral wrongdoing. This requires a certain threshold of sensitivity to common morality, where the agent is sensitive to the concept of *moral wrongdoing*. Moral injury is notable in that it is the belief that one ought downgrade one's own moral goodness, in light of the acknowledgement of moral accountability for the moral wrongdoing, whereas moral dissonance indicates a tension between the agent continuing to hold the belief that she is a morally good person, untouched by the moral wrong, despite having perpetrated a serious moral wrongdoing.

Thus with moral injury there is acknowledgement of responsibility *and* taking seriously the moral reactive attitudes like blame, resentment, disappointment, and moral criticism expressed by the community, what I call *holding morally accountable*. The designation of holding oneself morally accountable is a two-part relation, where the moral community holds the wrongdoer accountable (as expressed by the moral reactive attitudes), and the wrongdoer holds herself accountable to those attitudes (as expressed by taking them seriously, including downgrading her own moral standing). Moral injury entails holding oneself accountable to common morality such that accountability is a rich concept involving character, moral standing, and moral goodness. The morally injured hold themselves morally accountable such that they question their own moral character, whereas the morally dissonant acknowledge moral responsibility but do not question their moral character: they do not, in contrast to the morally injured, hold themselves morally accountable.

This characterization of the holding oneself morally accountable in moral injury corresponds to a main theme in Strawson--namely, that our practices of being sensitive to our own actions, acknowledging responsibility, holding ourselves morally accountable in such a way that we question our own moral character. It is Strawson's view that the main parties in the contemporary debate about moral responsibility have failed to identify what really matters to us as people in the moral sphere. Strawson's attempt at re-framing the debate accepts that something vital is left out of the account of moral responsibility, by all parties in the debate. Strawson shifts away from talking about moral responsibility as something which hinges on freedom and responsibility, and instead turns to what holds a

person responsible. It is less about conceptual analysis and more about a descriptive account about actual human moral psychology.

Moral philosophers tend to attribute responsibility to an agent by assessing the intellectual judgments of the agent, and asking whether those judgments satisfy some set of objective requirements. Holding oneself morally accountable, in contrast, is not so much about attributability of responsibility but about self-reflection about actions, codes, community membership, community attitudes, and one's own moral character in light of having perpetrated a significant moral transgression(s).

With moral injury there is the process of holding oneself morally accountable to the reactive attitudes and thus feeling, appropriately, shame, remorse, regret, and uncertainty about one's own moral character. What I am characterizing here is new in the sense that we do not commonly acknowledge that perpetrators of serious moral wrongdoing can hold themselves morally accountable to moral reactive attitudes, and yet on my definition this is intrinsic to the concept of moral injury. We have been negligent in understanding this, as Allan Young reminds us in "Suffering and the Origins of Traumatic Memory," all people have an equal eligibility to suffer, psychologically, after a traumatic event, including (certain morally sensitive) moral wrongdoers.¹¹¹

Strawson says that in our human relationships we can have participant attitudes. For example, we can believe that the moral wrongdoer is the sort of everyday agent who is free enough, rational enough, and thus eligible for our judgments. Or, we can take an objective attitude, where we believe that the wrongdoer is an inappropriate object, for

¹¹¹ Allan Young, "Suffering and the Origins of Traumatic Memory," *Daedalus* 125 (1996): 245-260.

some reason, for our moral judgments. Feeling shame is to feel *shame-worthy*, which connects both to an acknowledgement of responsibility and to agreement about the shaming attitudes of the moral community. In this way, a moral reactive attitude like shaming makes sense because we include a mechanism for discerning whether the agent is eligible for our shaming attitudes, and thus our moral reactive attitudes have validity.

The moral wrongdoer who is accountable to the shaming attitudes does not need to entertain questions about determinism, emotivism, or a sterile ethical theory: what she is experiencing is the natural human response to ordinary inter-personal attitudes like shame. This experience is psychologically predictable, normal, natural, with built-in mechanisms of appropriateness. Even if the moral wrongdoer were to explore why she behaved wrongly, explore her level of responsibility for the wrongdoing—all reasonable explorations—what I emphasize here is that she quite reasonably holds herself accountable to the moral reactive attitudes of her community when she feels shame, and feels shame-worthy, because these are the “natural,” to use Strawson’s description, psychological mechanisms at play. Thus packed in to moral injury is moral accountability to the responsibility for the transgression and to the moral reactive attitudes of the moral community.

Moral philosophers who determine, analytically, the components of responsibility talk about legitimacy, rightness and wrongness, desert, and other neutral properties in causal analysis. The components of accountability that I notice are the judgments that we the moral community hold, messy and socially contingent as they are, which helps to characterize and distinguish moral injury as a reasonable experience of a morally

sensitive agent. The morally injured is holding herself accountable to the moral reactive attitudes in a contextual, intuitive, and not clearly analytic but nonetheless descriptive way, which helps to differentiate and characterize it.

The function of our reactive attitudes is, according to Strawson, to “express how much we actually mind, how much it matters to us, whether the actions of other people—and particularly some people—reflect attitudes toward us of good will, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other.”¹¹² When people both recognize themselves as members of the moral community and they hold themselves accountable to the moral reactive attitudes, then they are standing in a position which is sensitive to the moral scrutiny of others.

Moral Injury as *Malady*

Walker and Gert emphasize the importance of the moral dimension of our social reality – how normative expectations shape both our self conception and our relations to others in fundamental ways. My work above connected how the wrongdoer with moral injury is notable in that they recognize the codes of common morality and hold themselves accountable to the moral reactive attitudes of their community. The narratives I referenced above indicate that there is a connection made between acknowledging and taking accountability for perpetrating a serious moral wrong, accepting the community’s criticisms, and forming new beliefs about the moral inferiority of one’s own character. My aim here is to acknowledge the contested term *disorder* in *Post Traumatic Stress Disorder*, to anticipate a similar resistance to the *injury* in *moral injury*, and to show that

¹¹² Strawson, *Freedom and Resentment*, 10

moral injury can be thought of as a malady, in part because of the mental distress connected with considering oneself to be morally bad.¹¹³

Mental maladies are conceptualized as mental processes and behavioral aspects associated with the mental mechanistic patterns. Mental maladies involve:

Something going wrong in the person's way of feeling or thinking, often involves mental pain or a volitional disability. [Mental maladies] often involve an increased risk of death, pain, disability, loss of freedom, or loss of pleasure.¹¹⁴

Gert *et al.*, admit that the term *malady* comes value-laden, but they appeal to the view of common-language morality which can accept that there are judgments which come included with labels and terms, and yet this does not mean that the term cannot be useful and objective.

Many causes are harms, and not all external harms indicate a malady. The conceptual move Gert and his co-authors make is to introduce the concept of a *sustaining cause*, which for a mental malady is located within the individual himself. Getting hit by a train is not a mental malady but an external trauma, but when the suffering from the train's impact persists, significantly, for an extended period of time, the suffering is a malady. The sustaining cause, in the train case, is within the individual himself, it is part of the individual and cannot be removed simply by changing his physical or social environment. Moral injury is a condition such that there is significant suffering which persists after the trauma has past.

¹¹³ "In this sense, to have a malady is to have an abnormal condition, other than her rational beliefs or desires, such that (she) is suffering, or is at significantly increased risk of suffering, nontrivial harm (death, pain, disability, loss of freedom, or loss of pleasure...)" Bernard Gert, Charles M. Culver, K. Danner Clouser, *Bioethics: A Systematic Approach, Second Edition* (Oxford: Oxford University Press, 2006)143.

¹¹⁴ Gert *et al.*, *Bioethics*, 166.

The narratives of the serious mental suffering in the morally injured, as described in the sections above, are consistent with Gert et al's scope of a *distinct sustaining cause*.¹¹⁵ The mental suffering is part of the individual, in conversation with her own moral codes and her holding herself accountable to the moral criticisms of the moral community. With the malady of moral injury, we have the connection between having moral accountability for a moral transgression, of holding oneself accountable for the transgression, of recognizing the transgression as seriously infringing on common morality, and being sensitive to the moral reactive attitudes—one's own, and on the part of the moral community—that then result. The morally injured have a unique and distinct mental experience, of considering themselves morally bad, in light of these connections.

This mental disturbance of moral injury sounds like the medical definition of disease: disease can be defined medically as disturbance in the body, or a disturbance in the environment. Disease can refer to a lack of health, or a lack of order. In this way, moral injury seems adequately thought of as a mental disorder or disease, except that disease does not include the vital aspect of moral judgments, those beliefs which form the connective tissue between judging oneself in relation to the judgments of others.

When the morally injured fractures her moral relationships she can develop her own moral reactive attitudes like disgust, disappointment, anger, fear, shame. Her reactive attitudes react to the norms her community has co-created, with her as both participant and consumer, developing a set of moral standards and expectations she holds for herself as well as others. When she broke a moral norm she damaged her relationship

¹¹⁵ Gert et al., *Bioethics*, 148.

to herself as well as to others, she damaged her conception of herself as a good person. Walker calls these transgressions and their consequences *moral damage*.¹¹⁶

Moral damage is evident to us in subtle and obvious ways. We acknowledge moral damage to our moral relationships when we express attitudes of disappointment, anger, resentment and other wrongful harms and losses for which human actors bear responsibilities, which develop when our normative expectations of others are challenged. For example, when one romantic partner betrays the trust of the other by violating the monogamous terms of their relationship, both perpetrator (cheater) and victim (cheat-ee) experience moral damage in their understanding of themselves and each other. The moral relationship is fractured by the damage, casting uncertainty and insecurity over both parties. Although these terms are academic, *moral damage* and *moral relationship* are phenomena that we are familiar with as laypeople, as members of the moral community.

I accept Strawson, Gert, and Walker's characterization that the moral expectations that we place on ourselves and others are important, and I take those characterizations and apply them to the narratives that moral wrongdoers with moral injury tell about their post-traumatic experiences. I find it psychologically understandable that from this set of experiences and beliefs serious psycho-social stress can result, as the trust-based relationships that the moral wrongdoer had with others and with herself are fractured. Resentment, indignation, insult, disappointment and other moral reactive attitudes are expressed by both wrongdoer and her community. These attitudes are

¹¹⁶ Walker, *Moral Repair*, 27.

understandably very stressful, and I expect that they are significant to the studies which show that perpetrators of moral wrongdoing have unique and significant post-traumatic suffering. These attitudes also demand an appropriate response.¹¹⁷ We commonly think that the demand is largely on the part of the victims and the moral community writ large, who “desire some forms of resolution and satisfaction, what Howard Zehr calls “an experience of justice.”¹¹⁸

Above, I characterized the concept of *holding oneself accountable* to the *moral reactive attitudes*, of one’s own and to one’s moral community, to ground much of what I define as *moral injury*. The characterization of holding oneself accountable is a two-part relation, where the moral community holds the wrongdoer accountable (as expressed by the moral reactive attitudes), and the wrongdoer holds herself accountable to those attitudes (as expressed by taking them seriously, including downgrading her own moral standing), such that serious, unique, and significant suffering is psychologically predictable. I posit next that the moral wrongdoer demands, of herself, a response for repair, in an attempt to remedy the *malady* of *moral injury*.

Chapter Two, Section Three

§2.3.1. Moral Repair for Moral Injury

This section is about expanding the scope of the concept of *moral repair* to a specific subset of moral wrongdoers, those with the capacity for moral reflection and

¹¹⁷ Walker, *Moral Repair*, 25.

¹¹⁸ Walker, *Moral Repair*, 21-32.

moral sensitivity such that they are eligible to suffer *moral injury*. Margaret Urban Walker has a notion of moral repair, but extends it primarily to victims of wrongdoing. Talking about *moral repair* in the vein of Walker's work is to talk primarily about two things: that *moral damage* is specific enough that it can be clearly categorized and characterized, and that *moral repair* is a process for addressing and repairing moral damage.¹¹⁹ My contribution is an argument that (in certain cases) perpetrators also ought to be eligible for moral repair. In order to be deserving of moral repair, the perpetrator has to be sensitive to reactive attitudes and feel appropriately accountable for past action, feel remorse, and show a willingness to change. Insofar as the perpetrator does exhibit these tendencies, she merits an avenues of moral repair, including critiquing the blaming attitudes of her moral community, and constructing new interpretations of her actions/character in a way that holds accountability but better protects her reputation and self-conception. Walker's project largely excludes moral wrongdoers from the scope of her argument and analysis of moral repair, and I will argue here for including this subset of moral wrongdoers.¹²⁰

The *repair* in Walker's work regards repair to the actual, tangible relationships that the victim has with moral wrongdoers as well as with moral relationships in society more broadly. Walker's focus is on victims of moral wrongdoing and my focus is on perpetrators, but I note that it is certain perpetrators, those with all the conditions which are components of moral injury. Walker characterizes the way that moral damage can be

¹¹⁹ Walker, *Moral Repair*, 22.

¹²⁰ Walker, 22, 172, 189.

addressed, amended, and perhaps made whole as *moral repair*, a repair to the moral relationships of our lives, and I apply her work to the subset of moral wrongdoers that I am interested in here.

For Walker's work, moral damage to our moral relationships *demand*s repair.¹²¹ What we the members of the moral community, and specifically the perpetrator of moral damage, demand is a normative confirmation that we have a right to expect that the damage be addressed, and that the damage is significant. We desire confirmation that the norms we hold are valuable to our moral community and also that we are valuable, such that we deserve an appropriate response.¹²² In this way, the perpetrator as well as the community can rightfully have a normative expectation for acknowledgement of the wrong and a serious response.

Walker has six specific points which can encompass repair, which include acknowledging and placing responsibility on the wrongdoer, acknowledging and redressing the moral damage, reinstating moral terms and standards, replenishing trust, nourishing hope, and the possibility of reconnecting moral relationships.¹²³ Victims and the community are the agents who naturally demand repair from wrongdoers: the community re-iterates the standards and the victim takes an active lead in the process of repair. The community must acknowledge the wrongdoing, and the victim must reclaim the right to have moral norms restored. The judicial process, if appealed to, is only a

¹²¹ Walker, 12-18, specifically the empirical studies she cites on 15.

¹²² *Ibid.*, 27.

¹²³ *Ibid.*, 28.

partial “and partially effective expression of such moral consensus as exists in communities,” and as such the avenues and agents of repair are multiple.¹²⁴

Repair can look like a a judicial process which addresses reckoning, criminality, restorative or reparative justice, but it can also look like the ways a social institution responds publicly, or when a romantic partner responds privately, recognizing and acknowledging their wrongdoing. Repair, at the very least, is an “acknowledgment of fault and responsibility (when justification and excuse do not apply), and, beyond that, some attempt at repair, including apology or amends.”

It is my project here to expand the scope of moral repair to include a particular, limited group of moral wrongdoers. The wrongdoer, after all, is the agent and perpetrator of wrongdoing. I take from Walker’s work and show that some aspects of moral damage are the responsibility of the wrongdoer to repair not just insofar as the victim has damage which demands repair, but because the wrongdoer also has injury, the malady of moral injury, which demands repair. It is for the wrongdoer’s own good that she pursue repair, and it is for reasons outlined in Walker’s work that she too merits such a process.

To recall an example from earlier in the chapter, Damien Corsetti, the torturer at Bagram, is publicly known to his moral community for having abused, tortured, and been implicated in the deaths of citizens in Iraq. Blame is placed upon him – both from the families of the people he abused and by his own family, friends, hometown, media reporters, as well as himself. He has transgressed against the codes of common morality which he takes to be important. Courtesy accepts the blaming attitude of his community,

¹²⁴ Ibid., 33.

judging it consistent with common values and as such, they are legitimate. He therefore concludes that he is a *bad person* – worse than others in his moral community.

What I have identified above is that sometimes with the malady of moral injury the agent forms beliefs that they are not deserving of repair, that is, they do not deserve to pursue moral re-integration into the moral community due to the egregious moral trauma they've perpetrated. At this juncture I identify an important point: that the moral wrongdoer is also a member of the moral community and continues to be the subject and object of normative expectations. The moral wrongdoer is akin to the moral victim in that both are morally damaged and have experienced moral damage: the wrongdoer as perpetrator, the victim as recipient.

Moral relationships do not simply obtain but they must be sustained, and the wrongdoer who is sensitive to the consequences and moral weight of their wrongdoing, who is experiencing the malady of moral injury, is as integral to the continued sustenance of moral relationships as is the victim of wrongdoing. It is incumbent upon the wrongdoer to address and nurture components of moral relationships like trust and mutual respect. My aim here is to acknowledge that the morally injured merits repair (to her own moral self-conception as well as to her moral standing in the moral community) and offer reasons why moral wrongdoers with moral injury can hold themselves accountable while at the same time re-assessing their and our blaming attitudes, such that they can re-assess their own moral goodness.

§2.3.2. The Morally Injured Can Legitimately Pursue Moral Repair

The Independence of Blameworthiness and Accountability

In this section I argue that because moral responsibility, moral accountability, and blame are conceptually distinct, the morally injured can be fully morally accountable while mediating the blaming attitudes of her community. Characterizing the two concepts as distinct from each other means that the moral wrongdoer can accept a full accounting of moral responsibility while yet re-dressing blame. This allows for the conceptual opportunity for the moral wrongdoer to merit a process, on their part and on ours, of moral repair.

In common morality we blame someone when we feel they are blameworthy, that is, we have formed the theoretical judgment an agent is morally responsible for a moral wrongdoing. In common understanding blame and moral wrongdoing tend to operate in tandem, and as such when the morally injured accept moral responsibility for their actions they also accept the blaming attitudes of their moral community. Strawson called this a *natural attitudinal reaction* to the perception of moral responsibility, expressed in our stance that the content of blame like shame, rebuke, or other moral evaluations which comprise blame are appropriate.¹²⁵

In R.J. Wallace's characterization of the force of blame, Wallace writes that blame that is not acknowledged is not blame in the fullest sense. The morally injured accept our attributions of blame and moral responsibility, in a way that Wallace might call a *morally accountable* response. The morally injured's acceptance of the legitimacy of blaming

¹²⁵ Strawson, 4-6.

attitudes seems reasonable and rational, to both the wrongdoer and to the moral community, because the wrongdoer is in fact morally responsible for a determinate moral wrong. The argument I put forth here is not that wrongdoers ought not accept blame, but that the morally injured merits critiquing the blaming attitudes of others. I will argue that the morally injured can have a robust conception of their own moral responsibility while still challenging some of the blaming attitudes of their community. This would allow a conceptual pathway for the morally injured to pursue moral repair, that is, a path to ameliorating blaming attitude while still accepting the full accounting of their moral responsibility.

Kayla Williams, the American soldier and interrogator in Iraq mentioned in above, says that blaming herself for her moral wrongdoing means that she believes she's a completely different, and now morally *bad*, person. Williams says that it is not possible for her to hope for a transformation of the blaming attitudes of her community which include actions like being ostracized.¹²⁶ She believes that her true character is morally very different from, and inferior to, the character of her peers.¹²⁷ She writes that she is justified in holding those beliefs because the people around her, whom she trusts and respects, hold strong blaming attitudes toward her. Williams accepts the blaming attitudes and blaming acts of her community and views them as a permanent and justified indictment of her character.

¹²⁶ Williams, epilogue.

¹²⁷ Williams, 19, 226, 251.

The moral community's blaming attitudes may appear to Williams to be intransigent, in part because the ones blaming feel *justified* in their blame. This justification tracks Williams' participation in unsanctioned torture and her brutal harassment of civilians. It is not irrational for the moral community to hold her to account, because our blaming attitudes correctly identify her personal responsibility in serious moral transgressions. When we the moral community blame someone, we are both *attributing responsibility* and *attributing blame*, in conjunction, and we are correctly identifying a determinate moral wrong that the wrongdoer perpetrated.¹²⁸ That is to say that in the everyday pattern of moral thinking, our blaming attitudes tend to mirror our responsibility attitudes, so that our forming blaming attitudes is a way in which we confer moral judgment while at the same time conferring and determining accountability.

In contrast, we say that the assault victim who blames herself is accepting irrational, misattributed, illegitimate blame. But both the assault victim and Williams share a common psycho-social space: both are accepting the blaming attitudes of the community, attitudes which shame and diminish their sense of moral worth. Both are members of the moral community who are involved in the continued renewal of moral relationships. Both victims and moral wrongdoers are part of the ongoing process of creating new moral relationships, and as such they are both co-creating and re-creating moral trust, moral accountability, moral worth, and the other components of their moral relationships.

¹²⁸ Decoupling *blame* from *attribution of responsibility* in G Williams, "Blame and Responsibility" *Ethical Theory and Moral Practice* (Springer, 2003) and K Shaver, *The Attribution of Blame: Causality, Responsibility, and Blameworthiness* (New York: Springer-Verlag, 1985).

Sergeant Williams admits to her wrongdoing and she holds herself *morally accountable* for her actions. She describes herself as someone I would consider to be morally injured, because she regards herself as morally deficient and diminished due to her actions, and her diminished moral self-regard is a serious stressor to her. If she can continue to hold herself morally accountable while at the same time believe that she can and ought redress the blaming attitudes, then she has a legitimate opportunity for what Walker calls *moral repair*. I will discuss Margaret Urban Walker's moral repair, for the morally injured like Williams, below.

I want to affirm that the morally injured is right to hold herself accountable for the moral wrong, she is right to consider herself morally blameworthy, and she is also right to consider whether certain blaming attitudes are, for example, hypocritical. This is because, *pace* Scanlon, Watson, Lamb and others, we can see that being blameworthy, blaming attitudes, moral responsibility and moral accountability are conceptually distinct, and yet these distinct concepts are conflated by the moral community.

§2.3.3. Mitigating Blame for the Morally Injured

In this section, taking my cue from work by Williams, Watson, and Lamb, I argue that the blameworthiness of a wrongdoer can be mitigated, re-characterized, and addressed while at the same time the wrongdoer fully embraces her *moral accountability* for the wrongness of her transgression. It seems right that moral wrongdoers are to be blamed while being held responsible. At first blush, we can feel that to be held accountable for a moral wrongdoing necessarily invites blaming attitudes, for what is

being responsible for an action if not being blamed for an action? In this section I will articulate how blame and holding accountable are conceptually distinct, and that holding accountable (something intrinsic to the malady of moral injury) can be distinguished from blame.

While *being to blame*, in our common language, can mean the same thing as *being causally responsible*, these two attitudes have different emotional content. For example, you could imagine a driver of a car hitting a soccer ball that children had kicked across the road. A legal analysis would notice components of *reasonable foresight* on the part of the driver, *contributing negligence* on the part of the children, and an assessment of damages. The conclusion may be that both the driver and the children contributed to the act of the car running over the ball and that all parties were mutually responsible. There are no robust moral attitudes to bring to this case because the process of attributing responsibility is a morally neutral analysis of cause and effect, contributory negligence and reasonable preparedness. If the driver could add emotional content to the situation, by accusing parents of outright negligence, saying that the parents ought be seriously reprimanded for leaving their children unattended. These new accusations have moral content, they are blaming attitudes.

Attributing blame, attributing responsibility, and attributing/holding oneself accountable are not always parallel judgments and are not, despite our assumptions, always expressed to the same degree.¹²⁹ The reactive attitudes of blame and maintaining accountability can combine or disentangle in our reactive attitudes. Attributing

¹²⁹ The soccer ball case is my own example, I am re-imagining Scanlon, 139-152.

responsibility can be an analytic, rational, morally neutral process; attributing blame is emotionally loaded, and perhaps irrational; holding oneself or another accountable is an integration of the two. I will here focus on the blame aspect of accountability, to separate it from responsibility, such that the morally injured can retain accountability while redressing blame. I do this by noticing that blame and accountability are conceptually distinct, such that the morally injured can retain accountability while mitigating blame.

Attributing responsibility is value-neutral (you can be responsible for good or bad acts); being accountable has a moral connotation itself be praiseworthy; blame is often used as punishment or emotional expression (and is always negative, no one 'blames' you for doing good things). Blaming attitudes can transform into actions, like when blame becomes a verbalized comment, a restricted freedom, a strained relationship.

Blame, being a negative moral assignation, can be hypocritical or illegitimate, to use Scanlon's characterization. Scanlon calls one kind of illegitimate blame *moral hypocrisy*, as when the blamer has committed the same, or similar wrongdoing.¹³⁰ For Scanlon, blame is illegitimate when it comes from a blamer with morally unclean hands, that is, a blamer who has committed a similar wrongdoing but who does not blame themselves.¹³¹ To blame someone else a thing which you do not blame yourself for is a performative contradiction which, to Scanlon's lights, is illogical and deceptive.

Feminist moral theory offers a critique of illegitimate blame, as when members of the moral community question an assault victim's choices and actions surrounding the

¹³⁰ Scanlon, 175-179.

¹³¹ *Ibid.*, 175-176.

assault, which is a process of blaming her, to some degree, for being assaulted. The moral community question the victim's liability and culpability, questions which are meant to both determine responsibility and determine blame, questions like what the victim was wearing, what she was drinking, where she was going. This is an attempt to implicate the victim's own decision-making in the causality of the assault, as well as an attempt to justify blaming attitudes toward the victim.

Scanlon argues that hypocritical blaming changes the moral relationship from one of moral legitimacy to an undermined, dishonest relationship. Scanlon's example is one where blame is selectively applied, and on his view disingenuously and thus illegitimately applied, for if blame and moral responsibility operate in tandem then the blamer ought blame herself as robustly. Scanlon's view of hypocritical attribution of blame shows that blame, responsibility, and accountability are not necessarily an analytical conjunction, and that a wrongdoer can accept moral accountability, accept moral responsibility, but challenge the kind and degree to which she is blamed for her actions.

When the moral community assigns blame they, in parallel, assign moral responsibility, which invites new (irrational, inappropriate, and illegitimate) reactive attitudes like guilt and shame. Blaming the victim is antithetical to most moral theory, which argues that assault victims are not blameworthy in any substantive way because blameworthiness requires some degree of agency and consent. Blaming the moral wrongdoer seems correct, but I point out here that we can hold the wrongdoer responsible, that she can hold herself accountable, and blame can still be mitigated.

What Walker and other moral feminists have brought to light is that accepting and validating blaming attitudes need not be the final word on the matter. The assault victim has the right to analyze and critique her own blameworthiness in light of the blaming attitudes of her community, and I posit here that so too does the moral wrongdoer. Moral wrongdoers with moral injury have the moral sensitivities and hindsight, they have the awareness of the codes of common morality, such that they have good reasons to address the (illegitimate) blaming attitudes, and to engage in a process of restoring their conception of their own blameworthiness and moral value while retaining moral accountability for their wrongdoing.

I noted above that blame is not a positive emotion but a negative one, because to blame is to scold, rebuke, shame, it is the emotional counterpart to praise. If the moral wrongdoer can critique blame on the grounds that it is hypocritical, or on the grounds that the blaming attitudes ought evolve with the evolution of the wrongdoer's own newly-formed beliefs and actions, then she merits the opportunity to re-assess these attitudes. My argument above made the conceptual point that blaming attitudes can be addressed without the morally injured having to discount her moral accountability, because the concepts are conceptually distinct. My argument above also made the conceptual point that the morally injured continue to be members of the moral community, continually renewing their moral relationships. These points together ground my argument below, that the morally injured merit a process to redress blaming attitudes while yet holding themselves robustly morally accountable for their wrongdoing.

Thus I am applying the concept of moral repair, largely focused on victims of moral wrongdoing, to certain perpetrators who have the moral resources to redress blame. Seeing that the morally injured can also recognize that *blame* is a separate concept from *moral responsibility and accountability* provides the opportunity to re-assess the attitudes of blame which implicates their moral character, and which contributes to the suffering, a component of the *injury*, of moral injury.

Assault victim and the morally injured are akin to each other in that neither believes that she ought challenge the blaming attitudes, but if the assault victim can address blaming attitudes then so too can the wrongdoer, as blame, responsibility, and accountability are conceptually distinct. For these reasons the morally injured can also retain holding herself morally accountable from accepting robust, permanent, intransigent blame. These concepts are not all or nothing, they can admit of degrees, and with this conceptual separation comes the justification for a wrongdoer to pursue maintaining moral accountability while addressing and ameliorating blaming attitudes.

Moral injury is indicated by those moral wrongdoers who hold themselves accountable for the determinate moral wrong, who recognize the moral reactive attitudes of their community, who consider their own moral character to be severely compromised, and who experience significant distress at those attitudes. The *malady* of moral injury is indicated by the suffering that comes from this experience.

The malady of moral injury is not insignificant, as the narratives I presented early indicate debilitating stress and suffering such that moral injury qualifies as a *malady*. I remind the reader of the disproportionate rate of suicide, drug abuse, domestic abuse and

criminality of perpetrators of who accept blame and accept these types of moral judgments upon their character. To re-stabilize attitudes and ideas of trust and confidence there needs to be active renewal and active repair, restoring or creating trust and hope in a shared sense of value and responsibility. I argue that this process is conceptually available to the morally injured because they accept responsibility, because they can redress blaming attitudes, and because moral relationships are continually sustained by the morally injured's own membership in the moral community.

Moral repair for the morally injured is likely a long and complex process, and I am only arguing here that the morally injured legitimately merit the opportunity for repair, due to the many factors of their experience including holding themselves accountable for their wrongdoing. Hope for repair for the morally injured is unique, as we do not often think about moral perpetrators meriting our forgiveness or meriting less blame, but hope for repair makes sense “not because the event hoped about is not past, but because one remains uncertain about how it has come out.”¹³²

I pose here that the moral wrongdoer merits a process for repair, because conceptually the process would be legitimate because she holds herself morally accountable while retaining the opportunity to re-characterize the blaming attitudes she and others hold. I imagine that the process of moral repair for the morally injured looks very much like the process of moral repair for victims: it would be a process of creating a new narrative which takes into account moral responsibility for the wrongdoing, and also hopes for the process of repairing the blaming attitudes about the moral wrongdoer being

¹³² Walker, 45.

intransigently *bad*. There would be specific tasks to repair, like addressing wrongs, reinstating moral terms, support for practices that express and reinforce those terms.¹³³ Moral repair, for the morally injured, would be a task of restoring or stabilizing—and in some cases creating—the basic elements of the wrongdoer’s moral goodness, in a way that acknowledges past wrongs and hopes for a renewal of character stability.¹³⁴ I do not address here what moral repair for the morally injured looks like, I only call attention to the possibility for their repair.

My project is not a discussion about therapies or treatments, nor is it a discussion about what moral repair for the morally injured looks like: my project here is to articulate the moral components of the particular moral phenomena of moral injury, the ethical implications to character and wellbeing, which then provides the foundation for a normative discussion about the moral reasons an agent merits moral repair. These subset of moral wrongdoers, the morally injured, are average people as integrated into the moral fabric of their communities as anyone else. Leaving them to the malady of moral injury is perhaps to shirk the duties we have to aid our peers, duties we have even to the serious moral wrongdoers so long as those wrongdoers have satisfied a set of criteria, which I argued earlier they have.

The morally injured, like her victims, are contextualized moral beings located within a fabric of ongoing normative expectations sustained by mutual trust, mutual confidence, and shared moral standards, which can be re-created and renewed. Both

¹³³ Walker, 28.

¹³⁴ Walker, 23.

victim and morally injured perpetrator have the opportunity to be critical of blaming attitudes. Both can legitimately critique attitudes of blameworthiness as a separate and distinct critique from analyzing responsibility or moral accountability.

Chapter Three

Checking the moral intuitions we derive from the Ticking Time-Bomb Hypothetical with resources and constraints in revisionist just war theory. Does the interrogator's susceptibility to moral injury change our moral intuitions?

Preface

In this chapter I analyze the TTB hypothetical and the moral intuitions it elicits through the lens of Jeff McMahan's reexamination of traditional just war theory—*revisionist just war theory* (RJWT). My project here is to check the moral intuitions elicited by the TTB hypothetical against resources within RJWT, in preparation for an upcoming project, where I discuss what I take to be missing within the TTB framework and within our own moral intuitions, which is, adequate and sincere moral concern for the moral costs to interrogators/torturers.¹³⁵ In this chapter I primarily check our moral intuitions; in upcoming work I will form an argument for adequate moral concern for the torturer, after establishing here that our moral intuitions satisfy RJWT constraints.

Examining the *Ticking Time-Bomb* Hypothetical

¹³⁵ Discussions of just wars and just war theory are against the background of Michael Walzer, *Just and Unjust Wars* 4th Edition (New York: Basic Books, 1977).

§3.1.1. The Power of the Ticking Time-Bomb Hypothetical in Intuitive Moral Thinking and in Contemporary Legal Practice

There are certain presuppositions and assumptions implicit and explicit in the Ticking Time-Bomb (TTB) hypothetical: the torturer is pursuing lawful aims within a lawful war, the detainee is an unjust actor implicated in the bomb's imminent detonation who persists in refusing to divulge the bomb's location, the bomb will kill many innocent lives, and the method for extracting the bomb's location would be minimal and interrogative not excessive and punitive.¹³⁶ The scope of the hypothetical is definite and discrete, i.e. no friends or family of the detainee will be tortured. While we may hold that for the most part torture is morally wrong, we also tend to hold that *in this case*, torture is

¹³⁶ *Moderate interrogation torture*, not punitive or sadistic torture, is the only kind of permissible torture. *Psychological* interrogation torture is considered a form of *moderate* torture, and the most widely accepted definition of torture internationally is set out by Article 1 of the United Nations Convention Against Torture and other Cruel, Inhuman or Degrading Treatment or Punishment (UNCAT): "Psychological forms of torture and ill-treatment, which very often have the most long-lasting consequences for victims, commonly include: isolation, threats, humiliation, mock executions, mock amputations, and witnessing the torture of others." from Jose Quiroga and James Jaranson, "The Encyclopedia of Psychological Torture," *IRTC*, Copenhagen, January 2005, <http://www.irct.org/what-is-torture/defining-torture.aspx>.

morally permissible.¹³⁷ As such the TTB hypothetical is a very powerful thought experiment.¹³⁸

Many have challenged the TTB hypothetical's moral and practical legitimacy as wildly unrealistic—that it fails to provide the mechanisms from which we would form an absolute moral prohibition against torture, or that it gives us (the wrong reasons) for bending the torture victim's will against itself, a Kantian criticism.¹³⁹ These criticisms largely conclude that, both in practical reality and in sincere philosophical analysis, the case cannot do any good moral work. I accept that these criticisms have merit: I accept that the TTB hypothetical is a wild fantasy disconnected from real cases, that it fails at eliciting absolutist intuitions and that it does not allow for our sentiments to fully consider the infringement of the detainee's will (although it does give us the opportunity

¹³⁷ Jeremy Waldron's meta-analysis shows that contemporary moral theorists are not arguing for the moral impermissibility of torture, citing reasons of context and consequence. The prevailing characterization by moral philosophers finds that psychological interrogation torture is considered "morally justified," "immoral yet justified," or "morally excusable." Jeremy Waldron, *Torture Terror and Trade-Offs* (Oxford: Oxford University Press, 2012).

¹³⁸ In the US, torture in its various forms had been, in the pre-September 11, 2001 era, deemed illegal, unconstitutional, and against international treaty. The United States had, previous to September 11, condemned the use of all forms of torture by other States, and had ratified anti-torture legislation. The United States Bill of Rights forbids cruel and unusual punishment, which has come to include corporal punishment except execution. In 1994 the US Congress enacted a new federal law to implement the requirements of the Convention against Torture, which characterized certain acts as torture and therefore globally prohibited, regardless of whether their use was for interrogation. The 1994 legislation follows the 1975 UN General Assembly, which passed the Declaration Against Torture, and the 1984 UN General Assembly outlawing specific acts of torture, of which the US is a signatory. In the post-September 11 thinking, America's definition of what constituted torture, and what the stance on torture ought be, was vague, partly due to a vagueness in the UN Convention's language. In the Convention's language there is legal and philosophical room for acts that inflict less than severe pain and suffering; Michael Gross points to the vagueness in the UN language as an indication of how certain acts can border on the UN characterization of torture, and yet not cross a legal line. Those who addressed the vague language began to characterize some acts as causing severe harm and therefore impermissible, and some acts as moderate and therefore permissible. The moderate and severe language demarcates a threshold of legal acts, but it also implies a threshold for moral acts, and it is into this area that I will pursue the moral weight of perpetrating moderate torture. See Michael Gross, *Moral Dilemmas of Modern War: Torture, Assassination, and Blackmail in an Age of Asymmetric Conflict* (Cambridge: Cambridge University Press, 2010).

¹³⁹ David Sussman, "What's Wrong with Torture?" *Philosophy and Public Affairs* 33, (2005): 33.

to weigh the harms done to the tortured against the harms to the many innocents).¹⁴⁰ As such I also accept the criticisms that the TTB hypothetical is a morally pernicious thought experiment which has been radically misused in our contemporary intellectual, legal, and military debate.

There exists, in Darius Rejali's analysis, a phenomenon known as *torture-analysis error*, described as a commitment to the "folklore of torture with social scientific legitimacy," of which the TTB hypothetical plays a pivotal, and privileged, role.¹⁴¹ Rejali cautions legal scholars and academics to be alert to the powerful *folklore* of the TTB hypothetical because actual historical examples of an isolated detainee, who is the only source for information without a web of co-conspirators to investigate, do not statistically pertain. Rejali writes that the TTB hypothetical is false on its face, because the TTB scenario is not how intelligence gathering operates in real life nor is it how impending terrorist attacks are thwarted. Using the TTB hypothetical in rigorous academic discussions, despite its having perhaps never happened in actual life, has been so pervasive and so instrumental to framing and arguing for views on torture that the TTB case acts, in Rejali's analysis, as a "significant and fundamental error" in torture, war, and policy reasoning.

¹⁴⁰ Notable deontological arguments against torture, which can argue that torture is generally wrong or morally absolutely wrong include Michael Davis, "I have now explained what torture is, why it is in general morally wrong, and not only why we have no clear example of morally justified torture but, more importantly, why we are unlikely ever to have one, even in a moral emergency... Absent some unlikely event, torture can in practice never be morally justified." Michael Davis, "Justifying Torture as an Act of War," *War: Essays in Political Philosophy*, ed. Larry May (Cambridge: Cambridge University Press, 2008), 204.

¹⁴¹ Darius Rejali, *Torture and Democracy* (Princeton: Princeton University Press, 2009), 518.

Rejali shows that there is a trend of re-telling and re-framing the TTB hypothetical as a real case in history with real potential to repeat itself, but that in contemporary military engagement, the case is a “fantasy.”¹⁴² Despite its fantastical power, contemporary torture narratives invoke the TTB case and validate it, which provides a neat hypothesis about torture’s efficacy, legality, and morality, but which is, ultimately, a reference to something apocryphal, and thus our contemporary torture analysis has a false starting-point.

David Luban posits that arguments that flow from an appeal to the TTB hypothetical are so far removed from any actual, real-world scenarios that it is “intellectual malfeasance” for us to derive intuitions from about the moral permissibility of real-life interrogation torture, and yet it is central to much thinking on the subject.¹⁴³ In Jessica Wolfendale’s *Torture and the Military Profession*, she shows that psychological interrogation torture is a significant part of military intelligence operations due in large part to military training methods that cultivate the practice based on an unexamined presumption of interrogation torture’s moral permissibility.¹⁴⁴ As such, interrogation torture is perhaps inherent to the machinations of war in an insufficiently reflective way and for under-analyzed reasons. In this way Wolfendale is in agreement with Rejali, who

¹⁴² Rejali, 500.

¹⁴³ David Luban, “Liberalism, Torture and the Ticking Bomb,” *Virginia Law Review* 91 (2005): 1459. Luban urges us to dismiss the hypothetical entirely, but I will take a different approach. I accept that the TTB hypothetical is such a strong thought-experiment, that it is unlikely to be dismissed outright, and instead I will analyze it by appealing to the doctrine and principles which emerge out of revisionist just war theory, such that the hypothetical can be more ethically sound. My conclusion is that my more robust and expanded TTB case will do more work in addressing concerns which are usually outside the case, like appropriate consequentialist and deontological concern for the torturer.

¹⁴⁴ J. Wolfendale *Torture and the Military Profession* (New York: Palgrave Macmillan, 2007).

writes that “[Interrogation torture] and democracy seem to go hand in hand.”¹⁴⁵

Wolfendale and Rejali are both in agreement with Waldon’s meta-analysis on contemporary moral debate about torture, which shows that moral theorists begin from the supposition that interrogation torture is, in certain cases, morally permissible, morally justified, immoral yet justified, or morally excusable. No doubt the TTB hypothetical has played a significant role in informing this moral starting-point.

I agree with Wolfendale, Rejali, Luban, Waldon and others that the TTB hypothetical is so significant that it transforms contemporary, living memory and plays a non-trivial role in legal and moral decision making, including in laws and military direction which were especially prolific during the Bush administration’s “war on terror.” I take seriously the intuitive force of the TTB hypothetical, and I follow McMahan and others who caution that “defenders and opponents of torture alike tend to misinterpret the significance of the ticking bomb case.”¹⁴⁶ Thus we find ourselves in an epistemic dilemma, because the TTB hypothetical is so salient and pervasive that any criticism of the moral prohibition against torture are “implausible,” to Jeff McMahan’s mind, that is

¹⁴⁵ In the criminal trials against US soldiers accused of interrogation torture in Iraq and Afghanistan, descriptions of interrogation torture include threatening a victims family, intimidation using dogs, humiliation with nudity, threats of severe beatings, threats of sodomy with foreign objects, and other psychological intimidations, see Quiroga and Jaranson, “The Encyclopedia of Psychological Torture,” (2005).

¹⁴⁶ When I move to discuss whether the detainee is liable to interrogation torture in the TTB case, I do not make a distinction as to whether the detainee is a terrorist, an unjust combatant, or an enemy combatant. McMahan articulates the differences in status, coercion, conscription, duress, and other factors. It would matter in a particular case whether the detainee is someone with conviction to the cause of terrorism, or a hapless conscripted soldier “on the wrong side.” McMahan notes that it is always morally wrong to kill/threaten to kill innocent bystanders, which is usually what the TTB hypothetical regards, Jeff McMahan *Killing in War* (Oxford: Oxford University Press, 2009), 125-130. In McMahan’s “Torture in Principle and in Practice,” he often refers to the detainee as a terrorist, and in his discussion in *Killing in War; terrorist* implies an *unjust combatant*, an unofficial actor who threatens innocent bystanders.

to say that our moral intuitions about the moral permissibility of torture limit our critical thinking on the ways it ought to be unlawful or impermissible.¹⁴⁷

Checking TTB Intuitions Against Constraints in Revisionist Just War Theory

There are some intuitions about the morality of interrogation torture that we derive from the TTB hypothetical that I will look at more closely: our intuition that the act of interrogation torture is defensive in nature (we accept that torture is the means in which the interrogator is defending innocent people's lives), that the detainee is the legitimate target liable for some form of defensive harm, and that interrogation torture is a proportionate means of defensive harm aimed at extracting the bomb's location. I will examine these intuitions in consultation with McMahan's work in RJWT.¹⁴⁸

The TTB hypothetical stipulates, and I will check our intuitions against RJWT constraints, that the detainee has information about an imminent detonation (a consideration of discrimination) which the detainee refuses to divulge cementing the threat to the innocents (liability), stipulates that many innocent lives hang in the balance (urgency), implies that the interrogator is a lawful combatant with no practical

¹⁴⁷ In "Torture in Principle and in Practice," McMahan says that the case helps most of us see that a morally absolute position against torture is untenable, because in the TTB case we intuit that interrogation torture is morally permissible. Wolfendale and Rejali note that this is our common moral starting point, *the moral permissibility of torture*, from which kinds and degrees are debated. McMahan concludes that our permissive intuitions ought not mean that we should endorse legal cover for torture, akin to what Dershowitz and others have proposed. McMahan offers a set of reasons for this, including reference to the famous Bybee memo and other Bush-era accommodations which show that legal cover enabled a broadening and expanding environment of torture light, pervasive excusing conditions and permissions. McMahan and Rejali conclude that despite our TTB intuitions we ought not, in our contemporary society, provide legal cover and for torture light. But I do not go as far as Luban and others who argue that the TTB hypothetical ought to be abandoned, excised from our collective moral musings: instead I take the intuitive force that the TTB hypothetical offers and check our intuitions against constraints in RJWT.

¹⁴⁸ Augustine, Ambrose, and Aquinas are the founding thought-fathers of historic and contemporary just war theory. According to Robert Emmet Meagher in *Killing from the Inside Out*, (Eugene: Cascade Books, 2014), "Their prescription for just war was a formula for enablement, not deterrence," 9.

alternative, and implies that the option of interrogation torture is of a minimal kind and is a last resort.¹⁴⁹ The TTB hypothetical presupposes that the interrogation torture would likely be effective (justification), and that the interrogator and her state are being wrongly, or unlawfully, threatened by the detainee who is, by virtue of not disclosing the bomb's location, liable to some form of intervention.¹⁵⁰

These implicit and explicit conditions need further analysis. For example, the TTB hypothetical does not offer information as to whether the detainee is a just or unjust combatant, a just or unjust civilian, or an unfortunate actor who is in over his head. Nor do we ask sophisticated questions about the detainee's reasons, cause, or intentions. The TTB hypothetical merely proposes, and intuitively we accept, that there is some causal chain between the detainee and the bomb's pending detonation grounding the detainee's liability to some sort of defensive intervention, and so I will check our assumptions with against resources within RJWT.

§3.2.1. Investigating Principles of Discrimination: Our Intuitions Are That The Interrogator is a Lawful Defensive Agent; Checking Those Intuitions Against RJWT

The particulars of the TTB hypothetical are of course vague and background questions of a just or unjust war are under-determined, but the TTB hypothetical makes

¹⁴⁹ From Michael Walzer, "Political Action: The Problem of Dirty Hands," *Philosophy and Public Affairs* 2, no. 2 (1973), 167. Explicitly stated in what is known as the "minimal harm rule," which refers to minimal infliction of harm, and minimal damage which would not be permanent. Notable deontological arguments against torture, which can argue that torture is generally wrong or morally absolutely wrong include Michael Davis and David Sussman.

¹⁵⁰ Culpable is defined as blameworthy; morally responsible is defined as having been in the causal chain. Lesser evil characterizations are characterizations within the law that have different interpretations, "but for consequentialist thinking it regards the justification of a smaller kind and degree of harm in order to prevent a greater kind and degree of harm," McMahan, "Torture in Principle and in Practice," 105.

certain implications, and from these implications we derive our moral intuitions. The TTB hypothetical is presumed to be contextualized as some sort of military environment where the interrogator and detainee are combatants in war, but the bomb's targets are not.¹⁵¹ The TTB hypothetical implies that the interrogator is some sort of lawful combatant who has a high level of credence in believing that the detainee knows the bomb's location.¹⁵² These presumptions matter, as a lawful agent in the theatre of war has duties and obligations which bear upon actions.¹⁵³ Because the TTB presumes, and we accept, that the interrogator is some sort of *lawful combatant in war*, we can check our

¹⁵¹ Possibly more apt is an analogy to a police officer detaining an informant who is a member of a criminal organization and who have planted a bomb. The TTB hypothetical is also dis-analogous to domestic cases, as the interrogator in the TTB hypothetical cannot call avail of other intelligence and security resources.

¹⁵² From Jeff McMahan, "Can Soldiers be Expected to Know Whether Their War is Just?" in *Routledge Handbook of Ethics and War*, ed. Fritz Allhoff, Nick Evans, and Adam Henschke (New York: Routledge, 2013).

¹⁵³ According to Walzer, as a matter of law, a war exists whenever one state uses armed force against another, regardless of the scale or duration of the conflict. When this happens, something of considerable legal significance does occur: the law of armed conflict begins to govern belligerent relations between the states. But this is a wholly conventional phenomenon and there is no reason to suppose that what is sufficient to activate a certain body of law automatically activates a different set of moral principles as well. The moral world of war is not divided between between morally justified defenders and morally culpable threats. Instead it is divided between those who pose a physical threat (combatants) and those who do not pose a physical threat (non-combatants). From this premise, Walzer argues for the moral equality of combatants (MEC) and noncombatant immunity (NI). MEC claims that combatants are moral equals, regardless of the justness of their cause (*jus ad bellum* considerations), as long as they follow *jus en bello* principles. Notably, Walzer grounds MEC in a sense of shared servitude and shared victimhood among combatants on both the just and unjust sides of a war. NI claims that non-combatants can never be legitimate targets of intentional attack, although they can be harmed in accord with the doctrine of double effect (DDE), Walzer, *Just and Unjust Wars*, 21, 41, 41-45

intuitions about the legitimacy of her actions against a set of moral principles in RJWT which govern actions in war, *jus ad bellum*.¹⁵⁴

Our moral intuitions are, because of what the TTB hypothetical intimates, that the interrogator has legitimate authority to carry out the interrogation; this legitimacy is a necessary requirement for the interrogator's actions to be lawful. Our intuitions are, from what the TTB hypothetical intimates, that the interrogator has a level of professionalization and that the professional norms to which the interrogator subscribes are in accord with the legal mechanisms of fighting a war. This legitimacy is necessary for the interrogator's actions to be circumspect and not wanton. In checking RJWT constraints, we are right to consider the harm that the interrogator will inflict, to use McMahan's term, as *purely defensive torture*. The kind, degree, and scope of the actions, that is, the sort of defensive torture that is morally permitted, needs further examination, as does the liability to defensive harms on the part of the detainee.¹⁵⁵

¹⁵⁴ We can think about wars in two ways: as an analogy to self defense in a domestic case, where states are akin to an individual defending itself. McMahan argues against this *collectivist* approach and for an *individualist* approach, because a domestic analogy is incompatible with concepts we do and ought accept, like how individuals can and should discriminate in defense and in war. Michael Walzer, in *Just and Unjust Wars*, characterizes wars as acts of self defense writ-large. Soldiers are engaged in an *adaptation*, as McMahan refers to it, to the right of self-defense, an adaptation because the moral justifications for war have an extra component of "chivalric morality," McMahan in a reply to Walzer, "Killing in War: A Reply to Walzer," *Philosophia* 34 (2006), 51. Different approaches to the morality of war: definition of war in Steven Pinker and Joshua Goldstein: "a common definition [of war] picks out armed conflicts that cause at least 1,000 battle deaths a year," "War Really Is Going Out of Style," *The New York Times* (New York, NY), *Opinion*, December 17, 2011

¹⁵⁵ It is important to say, again, that just war theory offers a moral lens through which actions are judged, not a legal or legislative one. In McMahan's view, *jus ad bellum* and *jus en bello* are not distinct, they are fundamentally related. The revisionist approach treats war as morally continuous with violence and conflict in ordinary life and so there is no separate and distinct morality which applies in war that does not apply in ordinary life, although moral cases in war can be incredibly rare in ordinary life. I appeal to revisionist just war theory and McMahan's work because it provides for the moral seriousness of doing harm in war just like in doing harm in ordinary life. Killing in war investigates the moral seriousness of harming others in war, and how that seriousness deserves the moral sentiments we bring to doing harm in ordinary life.

§3.2.2. Investigating Principles of Discrimination & Liability: The TTB Hypothetical Stipulates That The Detainee is a Legitimate Target, Who Makes Himself Liable to Defensive Harms

Discrimination constraints in RJWT are satisfied if the detainee has information about a wrongful imminent detonation, and withholding that information can make the detainee liable to defensive harms. Thus, on McMahan's view, discrimination and liability are conceptually linked: defensive harms of some sort (and what kind and degree need analysis) can satisfy the requirement of discrimination only if the detainee is liable to those defensive harms by virtue of their moral responsibility for imminent, significant harms. I'll unpack these conditions and restraints next.

We accept what is stipulated by the TTB case, that the detainee does, either in fact or more likely than not, know the location of the bomb, and so the detainee is an agent who we have discriminately determined to know the bomb's location. Liability, in McMahan's work, is analyzed as sensitive to two considerations: an objectively unjustified threat of harm and a moral responsibility for that harm. This is McMahan's articulation of the *Responsibility Account* for assessing liability to defensive harm.¹⁵⁶ Liability is calculated via culpability, in aggregate, to significant future harms,

The justification is not a necessity or lesser evil justification — that is, it is not that the harm done to the terrorist would be significantly less than the harm that would thereby be prevented. Rather, the justification in this case is liability-based. It is that the terrorist's own moral responsibility for the fact that someone must suffer torture makes it the case that, as a matter of justice, he should be the one to suffer the harm that he has made unavoidable. This is a claim about justice in the *ex ante* distribution of harm, not a claim about greater and lesser harms.¹⁵⁷

¹⁵⁶ Jeff McMahan, "Proportionate Defense," *Journal of Transnational Law and Policy* 23 (2013-2014): 1-36.

¹⁵⁷ Jeff McMahan, "Torture and Method in Moral Philosophy," in *Torture, Law, and War*, ed. Scott Anderson and Martha Nussbaum (Chicago: University of Chicago Press, 2015).

In checking our intuitions against permissibility constraints in RJWT, because we accept the premise that the detainee knows the bomb's location, then as long as the detainee refuses to divulge the information, the detainee is liable to the defensive harms of some sort by a lawful, defensive agent.¹⁵⁸

The primary justification for the permissibility of torture as a mode of defensive harm in the TTB hypothetical is that the detainee has, "through his own culpable action, made himself liable to be tortured in defense of his innocent potential victims."¹⁵⁹ In the TTB case, because we accept that the detainee is morally culpable if the bomb detonates, then that, as long as he refuses to divulge the bomb's location, he makes himself liable to some degree of defensive harm. It is interesting to note that our intuitions about the permissibility of torturing the detainee are not derived through a cost-benefit analysis of what lesser harm can be inflicted on the detainee in comparison to the grave harm of allowing the innocents to die. That is to say that our intuitions about the permissibility of torturing the detainee are not derived through a comparative assessment between harms to the detainee v. harms to the innocent victims. Instead, our intuitions come through the detainee's refusal to divulge the bomb's location, which makes him causally connected to its detonation and thus liable to some sort of defensive harm aimed at discovering the bomb's location.

¹⁵⁸ In JTWT, the soldier is an agent of the state, the state is a collective and has voted/established some procedure in going to war in ways that *de facto* view their war as just. The permissions and constraints on an agent of the state in a just war, according to TJWT, are granted via a) her being an agent of the state which is b) just in its activities.

¹⁵⁹ McMahan, "Torture and Method in Moral Philosophy," 5.

The detainee is not just morally responsible for the bomb, but is also liable for its detonation. For McMahan, the difference between moral responsibility and moral liability is explained in this example: Imagine a conscientious driver who carefully drives her well maintained car carefully to the movie theater, and along the way a freak, unforeseen event causes her to lose control of her car and kill a pedestrian. On McMahan's view, she is morally responsible for the death of the pedestrian because all choices to drive come with uncertain risk, something which all drivers ought to know. The risks, in aggregate, are diffused among all the drivers and all the many times they have driven without killing anyone, and yet each time we drive we take on some small degree of risk.

There is parity of reasoning in this example when we apply the same principles to an innocent bystander having lunch in an outdoor cafe: the bystander ought assume no risk for being hit by a runaway vehicle while she is eating her lunch. More slicing and dicing of the levels of risk and responsibility can be brought to bear upon this example, but McMahan is broadly differentiating *moral responsibility*, which the driver has in abundance, from *moral liability*, which neither driver nor bystander have. McMahan uses *The Responsibility Account* as a resource for calculating *liability* to harm, and I have discussed above how the detainee in the TTB hypothetical moral liability is necessary for the detainee's liability to the defensive harms of torture.

Our intuitions are grounded in our acceptance of what the TTB hypothetical strongly implies if not explicitly states, which is that the detainee has some degree of moral responsibility for the bomb's existence, that the bomb is unjust in that it threatens the lives of many innocents, and that the detainee persists in refusing to divulge the

bomb's location. As long as the detainee refuses to reveal the bomb's location, the detainee is confirming his own moral responsibility, and thus his *liability to defensive harms*. The detainee is, to our minds, *de facto* complicit in any defensive harms the interrogator pursues.

The TTB hypothetical offers us a very circumscribed liability case: the detainee may be a just or an unjust combatant, or the detainee may be minimally responsible for the creation and placement of the bomb. In refusing to divulge the bomb's location we determine that the detainee is liable to some degree, and some kind, of defensive harm,

Liability... arises only when harm is unavoidable and must be distributed. Whether one is liable to harm depends not only on what one has done but also one what others have done. Even if one bears some responsibility for an unavoidable harm that cannot be divided, one may not be liable to suffer the harm if someone else is more responsible.¹⁶⁰

In checking the resources within RJWT, aspects of the detainee's luck, character, or even moral responsibility for creating and placing the bomb *are not factors* for determining liability to defensive harms. Our intuitions pass RJWT liability constraints.

What sort and what kind of defensive harm needs its own analysis, and from here I look at interrogation torture and whether it satisfies conditions of *proportionate defensive harm*. What seems right is that the detainee is increasingly morally culpable for the bomb's detonation as the bomb's timer continues to ticks down.¹⁶¹

¹⁶⁰ Liability is thus contrasted with desert. Liability is something which comes into play when there is unavoidable harm to be distributed—whether one deserves to be harmed is a different consideration from whether one is liable to the harm that is unavoidably at hand, McMahan, *Killing in War*, 157.

¹⁶¹ In JTWT, the soldier is an agent of the state, the state is a collective and has voted/established some procedure in going to war in ways that *de facto* view their war as *just*. The permissions and constrains on an agent of the state in a just war can do are granted via a) her being an agent of the state which is b) *just* in its activities.

It is important to acknowledge that the detainee has multiple opportunities to avoid torture by revealing the location of the bomb. The detainee in the TTB case has at minimum two opportunities to avoid defensive harm: (1) when he chose to plant the bomb or not reveal its location when he found out about it, and (2) when he is threatened with torture if he doesn't reveal it. So long as the detainee has a chance to reveal the location, then at each opportunity and refusal his liability to defensive harm increases, or strengthens.

§3.2.3. Checking our TTB Intuitions With Principles of Necessity & Proportionality

The *necessity constraints* on *permissible* (i.e., not wrongful) *defensive actions* by a *legitimate defensive agent* against a *liable* agent are constraints that take into account the kind and degree of the defensive harms inflicted, which are a *proportionality* considerations.¹⁶² In the TTB case, the *principle of necessity* can be applied to the act of interrogation torture once the other epistemic checks have been satisfied, which I have done here.

It is important to keep *necessity* and *proportionality* distinct. *Proportionality* compares the consequences from doing nothing, compared to the consequences from doing something in the pursuit of a just aim. *Necessity* compares the consequences of one

¹⁶² McMahan's narrow proportionality constraints are sensitive to the intentions of the person who inflicts the harm, and so the distinction between doing and allowing is relevant. Liability to harm considerations are intrinsic, on McMahan's view, to proportionality concerns, as it makes no sense to say that someone is liable to X kind and X degree of harm *simpliciter*. Liability to harm considerations are also instrumental, in that proportionality considerations are an empty analysis without liability. In the taxonomy of proportionality judgements which apply to the detainee in the TTB case, now that I have analyzed that the detainee is in fact liable to defensive harm, then we examine the act of harm which is, in this case, intentional defensive harm to a liable agent.

means of achieving an aim with the consequences of another means of achieving that aim. *Necessity* and *proportionality* are related in complex ways but the fundamental difference is that they depend on different comparisons. The below references to urgency, no practical alternative, and last resort are all components of the *necessity* requirement.

In the TTB hypothetical, our intuitions are that the bomb's detonation is imminent such that *something minimally must be done* (by someone who is in a position to act, urgently, defensively, and *necessarily*) to the detainee who is liable for the bomb's imminent detonation. That which must be done, that which is *necessary*, is that which will cause the least amount of harm to the detainee in the genuine pursuit of saving the innocents.¹⁶³ The *necessity constraint* forbids the use of harms that are unnecessary to the aim sought, i.e. are egregious or gratuitous, and thus punitive or sadistic torture is outside the parameters of the TTB case.¹⁶⁴

Another dimension of necessity constraints regard checking that *defensive torture* causes the least amount of harm to the morally liable detainee: the choice is not just about which method of interrogation is necessary to extract the information, but it is about which method is, among all the different methods and means of defense, morally best and *proportionate* to what the detainee is liable for.¹⁶⁵ So, *threatening* the detainee with harm

¹⁶³ Seth Lazar, "Necessity in Self-Defense and War," *Philosophy & Public Affairs* 3, (2012): 40.

¹⁶⁴ Almost all theorists agree that necessity is a requirement that egregious harms are not morally acceptable. For the interrogator's defensive harms to satisfy necessity constraints, they must not cause egregious foreseeable harms unnecessary to secure the bomb's location, from McMahan, *Killing in War*, 12.

¹⁶⁵ Traditionally, in just war theory, proportionality constraints were almost exclusively focused on collateral damage of civilians and whether those deaths were unintentional, unintentional but foreseeable in the pursuit of a purported good aim. For McMahan, questions of proportional harm inflicted on, for example, civilians are questions about the harms they are liable to, *Killing in War*, 2.

is morally best: it is proportionate to what the detainee is threatening (the pending detonation,) and satisfies conditions of necessity, that some sort of defensive action is warranted.¹⁶⁶ The interrogator ought to, in order to satisfy necessity and proportionality constraints, begin by threatening harm. But, the time-bomb is ticking, and time is of the essence, which increases the necessity conditions for the legitimacy of defensive harm. Our intuitions are not so fine-tuned, what we are intuiting is that there will be some sort of non-trivial harm *actually* inflicted upon the detainee, so we are checking whether interrogation torture satisfies constraints of proportionality.

Proportionality considerations are moral constraints on actions, interpreted differently by different theorists. In Traditional Just War Theory, TJWT, considerations about the moral constraints of proportionality focus on things like collateral damage and whether the loss of innocent lives was unintentional but foreseen, the costs of lost lives is weighed against the intended good aims of the action such that a calculus of proportionality emerges. McMahan revises the TJWT concept of proportionality into *wide* and *narrow* proportionate harm, of which narrow applies in the TTB case.¹⁶⁷

¹⁶⁶ It is an open question to McMahan whether proportionality constraints require the *good* to outweigh the *bad* of the harm, or if commensurate *goods* and *bads* satisfy proportionality constraints. I take McMahan's work to mean that the good or bad harms that are produced are one factor of proportionality concerns.

¹⁶⁷ I am working from Jeff McMahan, "Proportionate Defense." The *narrow* sense of calculations about the permissibility of harm refer to those who are liable, to some degree, for some sort of defensive harm. MacMahan's *narrow proportionality* argues that individuals are only liable to a certain degree of harms, so someone who threatens to punch someone in the arm, for example, is not liable to be killed in self-defense. Narrow proportionality *constraints* are a matter of whether the harm inflicted exceeds that to which the victim is liable.

According to RJWT, the defensive harms cannot be gratuitous, or excessive, to that which is minimally required to achieve the goal at hand.¹⁶⁸ Interrogation torture could satisfy conditions of RJWT proportionality constraints by comparing “what relevant bad effects that the defensive action ... would cause, either directly or indirectly, and the relevant good effects it would cause—in particular the prevention of harms that would otherwise be caused by others.”¹⁶⁹ Specifically, proportionality conditions regard a comparison between the bad effects of harming the detainee are not disproportionate, excessive, to the good aim.¹⁷⁰

In the TTB case, our intuitions are that the permissibility of interrogation torture holds because of the presumptions that many, as Wolfendale, Rejali and others have noted, make: that the defensive harms are not disproportionate. Saving the lives of many is the *good effect* that interrogation torture will purportedly evoke. Interrogation torture, which aims at revealing the bomb's location and defusing it, is a permissible

¹⁶⁸ This interpretation of the necessity constraints are from Jeff McMahan, “Proportionality and Necessity in Jus in Bello” in *Proportionality and Necessity in Jus in Bello*, Ed. Seth Lazar and Helen Frowe (Oxford: Oxford Handbooks Online, 2016).

¹⁶⁹ Discussions of proportionality are discussions about constraints on harmful actions; traditional theories of defensive rights view proportionality in terms of collateral damage, a consideration of double-effect, and intentional v. unintentional killing of innocents. Traditional theories of proportionality also consider urgency, or necessity, of harm. McMahan’s view expands the traditional analysis of proportionality in ways more specific and yet more broad than what we derive when appealing to TJWT. On McMahan’s analysis there are more fine-tuned ways to think about proportionality than traditional accounts offer.

¹⁷⁰ McMahan, “Proportionate Defense.”

proportionate harm so long as it does not cause excessive harms to the detainee, disproportionate to the good aim of saving lives.¹⁷¹

In checking our moral intuitions in consultation with McMahan's proportionality constraints, because we accept what the TTB hypothetical implies, that interrogation torture will more likely than not save lives, and because we accept what the TTB hypothetical implies, which is that the bad effects to the detainee would be non-trivial but also not long-lasting, interrogation torture satisfies considerations of proportionality. The TTB case is perhaps most circumspect, most *fantastical* to use Rejali's characterization, at this level of analysis, because the TTB case stipulates that the interrogation torture will more likely than not be the act that breaks the detainee's resolve. The good effects of a successful interrogation are stipulated and not explicitly stated.

It is important to acknowledge that the form of proportionality that constrains a liability justification means that it could be proportionate to torture the detainee extensively, for a prolonged amount of time, so long as he is not killed. Because the detainee continues to refuse to divulge the bomb's location, and the bomb will kill at least one person, then one could in principle inflict harm upon the detainee that would be a) greater than the sum of the harm to all the victims but b) less than killing the detainee, as

¹⁷¹ McMahan's example of what a pickpocket is liable for is applicable to the TTB case. It may seem that the interrogator could kill the detainee and that death would be morally permissible on proportionality grounds, because the detainee is threatening to kill/allow others to be killed. The only 'goods' that count are that which the interrogator is pursuing, and the interrogator is saving the lives of the many. The TTB case poses that what is needed is information, and that inflicting some non trivial amount of pain will motivate the detainee to reveal it. What the interrogator aims for is information, and the death of the detainee would mean that no more information is forthcoming, McMahan, *Killing in War*, 20-32.

killing would do nothing to reveal the bomb's location. This is a disconcerting, but proportionate, potential eventuality.¹⁷²

External bad effects, like whether torturing detainees influences sympathizers to take up arms, are a nuanced assessment of the permissibility of torture that is outside the scope of the TTB hypothetical. The unintended consequences of emboldening sympathizers to the detainee's cause could count in a real-life analysis of proportionality if torturing the detainee directly caused sympathizers to rise up, but in the fictional thought-experiment of the TTB those sorts of concerns are external to it. That is how these thought experiments work, they offer a small case with clear parameters which invoke our intuitions about the scenario within the parameters of the case. Once we start asking questions about bad effects to agents outside the case at hand we move into the realm of real-life analysis and leave the scope of the thought experiment. Our intuitions are, and the TTB hypothetical intimates that, interrogation torture satisfies constraints of necessity and proportionality.

In consultation with proportionality constraints, it would not be morally permissible to kill the detainee in the TTB hypothetical, because it is only morally

¹⁷² This conception of liability to a proportionate degree of harm is instrumental, that is, "in most cases, for an act that causes harm to be justified, it must be instrumental to the achievement of some valuable goal against which the harm can be weighed and assessed. If the assessment is favorable, the harm is proportionate; if it is unfavorable, the harm is disproportionate," McMahan, *Killing in War*, 19.

permissible to inflict those defensive harms which are proportionate to what the detainee is threatening.¹⁷³

¹⁷³ Killing the detainee would be unnecessary, because any harm that is ineffective is unnecessary. (One might argue that ineffective harm is also disproportionate in that there are no good effects to weigh against the bad.) Liability is calculated via culpability, in aggregate, to significant future harms, and certainly the interrogator seems liable, in the detainee's view, for defensive harm because the interrogator is responsible for (unjustified, to the detainee's mind) threatened harm. To the detainee's mind, the threatened harm of interrogation torture would be unjustified because other options exist: intelligence gathering, capitulation to the detainee's demands, etc. If the detainee could reactively harm the interrogator then it seems that the detainee is justified in defensively harming the interrogator, because the detainee is protecting her own immediate wellbeing which the interrogator threatens.

Conclusion

The Socratic claim that moral wrongdoing is bad for the moral wrongdoer has always been a confounding claim. For the most part it appears to be, on its face, easily dismissible, as we surely all know, or know of, really bad people who have done really bad things, and yet their lives seem to be progressing very well. History is littered with these rich, powerful, successful bad guys. But there is something to this claim that I have not been able to let go of, something about which moral philosophers ought to take note: that ethics and morality are, in a very clear and demonstrable way, an inescapable phenomenon in our own lives as we live as moral beings. Perhaps the corollary, or opposite, also holds: that our our serious moral failings are no small thing to us.

Jonathan Shay brought the concept of moral injury to contemporary psychology, and he has done a tremendous service to medicine as well as to humanity. He called our attention to the particular moral plight of soldiers who return from war injured, in some psychological way that intersects with the moral. The injury, in Shay's characterization, impacts the soldier's view of the moral world and the moral standing of their friends and family (downgraded, less-than, deficient, evil) as well as to their own moral character and moral quality (also downgraded, less-than, deficient, evil).

Moral injury, to Shay, is a broad term which was immeasurably significant to providing a voice, and then a strategy for therapy, to psychologically wounded soldiers for whom post-traumatic stress disorder was an insufficient characterization. I am grateful to Dr. Shay's work and I respectfully narrow and refine his concept to characterize a psychological phenomenon, which I think has tremendous import to moral philosophy: that certain moral wrongdoers are deeply affected by their own moral wrongdoing. Socrates' insight becomes a contemporary, relevant perspective on the ethics of wrongdoing to wrongdoers in ways that are made apparent in broad empirical studies and in personal narratives.

It has been my goal in this investigation to re-visit this ancient, somewhat ridiculed and dismissed concept of character injury or soul injury and examine it through the lens and with the resources of contemporary medical research. If certain moral wrongdoers, for example, soldiers, who committed serious atrocities and moral transgressions in war are describing to us how their beliefs about themselves as moral agents have been drastically altered and diminished, then moral philosophy ought to engage with this phenomenon, and I hope that I have provided a clarity in scope and terminology by offering a typology of moral injury that is useful philosophically and psychologically.

There is still more fertile ground available for philosophical research on this topic. Conspicuously missing from my research on contemporary moral narratives are the voices of moral wrongdoers who were not soldiers but who were another kind of moral

wrongdoer: remorseful murderers, self-critical pedophiles, traumatized child abusers, to name a few. The category of morally sensitive moral wrongdoer is, certainly, far broader than I have offered here, and I hope to apply my work to other areas and other kinds of (im)moral agents. I am indebted to the medical research being done by military psychologists, and thank them for noticing and categorizing the psychological distress of moral wrongdoers v. witnesses to moral trauma, and I hope to find similar research being done in other medical fields.

With the present text as a starting point, I have developed tentative plans for three potential projects, rough ideas for projects that grew out of my research, and which I plan to pursue in the coming years. I would like to explore the dramatic context of Plato's *Gorgias*, which means to bring together the historical information on the characters of the dialogue and pursue their history through the lens of the Socratic claim about moral wrongdoing harming the moral wrongdoer. In my preliminary research I have found that the character given by the interlocutors as an example of a seriously morally bad actor who is doing very well was, in fact, famously ruined by the end of his life. The tragic fall of the hero that the interlocutors offer as an example happened after the events of the dialogue were set, but within Plato's lifetime. The early readers of the dialogue would undoubtedly have had the fall of this hero in mind as they read the interlocutors' presentation of him as a counter-example to Socrates' claim, and I think that a dramatic and historical examination of the dialogue would bring further dimensions to Socrates' moral argument concerning this particular immoral agent.

My second proposed project will look at empirical research into morally sensitive

pedophiles, and what the psychological community is doing to frame the phenomenology of considering oneself to be morally aberrant and how that sort of agent can consider himself in the moral community. I hope to provide the lens and language to which the morally sensitive pedophile, who reviles himself and yet is able to not act on his desires, can be understood to be suffering from the moral injury of a badness of character. I expect to use Margaret Urban Walker's *Moral Repair* as a foundational text and guideline in the pursuit of looking at the reasons certain pedophiles can understand and repair their (im)moral character.

My third project is to argue for international policy that ought to take seriously the moral requirement for the international community to plan for moral repair to postatrocity communities, for example, how a pillar of the UN's Commission on Post-Genocide Human Rights can be that of the moral requirement to re-integrate perpetrators of moral trauma back into the moral community. I have worked on this project with a local chapter of the United Nations General Counsel, and I hope to submit white-papers on the concept of moral injury and perpetrators of genocide, and argue that the international community, as well as local communities, has an interest in the moral repair of perpetrators. I expect to refer heavily to Jeff McMahan's work in *Killing in War* and other research to pursue the concept of moral injury as it can be useful to policy in postatrocity, post-war communities.

Bibliography

- Adkins, Arthur William Hope. 1960. *Merit and Responsibility: A Study in Greek Values*, Clarendon press.
- American Psychiatric Association, and American Psychiatric Association. DSM-5 Task Force. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. Fifth edition / Electronic version. *Psychiatry Online*. Arlington: American Psychiatric Association. <http://dsm.psychiatryonline.org/book.aspx?bookid=556>.
- Andersen, Øivind. 1999. "Review." *Gnomon* 71 (3): 260–62.
- Aristotle. 1984. *The Complete Works of Aristotle: The Revised Oxford Translation*, Vol. 2. Edited by Jonathan Barnes. Princeton University Press.
- Blondell, Ruby. 2013. *Helen of Troy: Beauty, Myth, Devastation*. Oxford: Oxford University Press.
- Bloomfield, Paul. 2008. *Morality and Self-Interest*. Oxford University Press US.
- Boesche, Roger. 1995. *Theories of Tyranny: From Plato to Arendt*. Penn State University Press.
- Bourke, Joanna. 2000. *An Intimate History of Killing: Face-to-Face Killing in Twentieth-Century Warfare*. Basic Books.
- Brentano, Franz, and Tim Crane. 2014. *Psychology from an Empirical Standpoint*. 1 edition. Abingdon, Oxon: Routledge.
- Brownlee, Kimberley, and Adam Cureton. 2009. *Disability and Disadvantage*. Oxford University Press.
- Brunel. 2001. *The Geopolitics of Hunger, 2000-2001*. Lynne Rienner Publishers.
- Burgers, J. Herman, and Hans Danelius. 1988. *The United Nations Convention against Torture: A Handbook on the Convention against Torture and Other Cruel, Inhuman, or Degrading Treatment or Punishment*. Martinus Nijhoff Publishers.
- Burgess-Jackson, Keith, ed. 1999. *A Most Detestable Crime: New Philosophical Essays on Rape*. 1 edition. New York: Oxford University Press.
- Burnet, John. 1916. *The Socratic Doctrine of the Soul*. British Academy.

- Calo, Zachary R. 2008. "Torture, Necessity, and Supreme Emergency: Law and Morality at the End of Law." *Valparaiso University Law Review* 43: 1591.
- Center for Military Health Policy. 2008. *Invisible Wounds of War: Summary and Recommendations for Addressing Psychological and Cognitive Injuries*. RAND, Center for Military Health Policy Research.
- Chambers, Clare. 2007. "Torture as an Evil: Response to Claudia Card, 'Ticking Bombs and Interrogation.'" *Criminal Law and Philosophy* 2 (1): 17–20.
- Clark, Donald Lemen. 1957. *Rhetoric in Greco-Roman Education*. New York: Columbia University Press.
- Claus, David. 1981. *Yale Classical Monographs*. Yale University Press.
- Copp, Terry. 2004. *Fields of Fire: The Canadians in Normandy*. University of Toronto Press.
- Corsetti, Damien. 2016. "Former US Interrogator Recounts Torture Cases in Afghanistan and Iraq." *Informed Comment*. Accessed May 19. <http://www.juancole.com/2007/12/former-us-interrogator-recounts-torture.html>.
- Demosthenes, and Thomas Leland Translator. 1853. *The Orations of Demosthenes: Pronounced to Excite the Athenians Against Philip, King of Macedon, and On Occasions of Public Deliberation*. London: Henry G. Bohn.
- Denis, Lara. 1997. "Kant's Ethics and Duties to Oneself." *Pacific Philosophical Quarterly* 78 (4): 321–48.
- Dershowitz, Alan M. 2003. "Torture Warrant: A Response to Professor Strauss, The." *New York Law School Law Review* 48: 275.
- Diels, Hermann, and Rosamond Kent Sprague. 2001. *The Older Sophists: A Complete Translation by Several Hands of the Fragments in Die Fragmente Der Vorsokratiker*, Edited by Diels-Kranz. With a New Edition of Antiphon and Hackett.
- Dillon, John, ed. 2003. *The Greek Sophists*. Penguin.
- Dover, K. J. 1994. *Greek Popular Morality in the Time of Plato and Aristotle*. Hackett Pub Co Inc.
- Doyle, James. 2010. "Socrates and Gorgias." *Phronesis* 55 (1): 1–25.

- Feinberg, Joel. 1970. *Doing & Deserving; Essays in the Theory of Responsibility*. Princeton, NJ: Princeton University Press.
- . 1984. “Moral and Legal Responsibility of the Bad Samaritan, The.” *Criminal Justice Ethics* 3: 56.
- Figley, Charles. 2006. *Combat Stress Injury*. Taylor & Francis.
- Fischer, John Martin. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge Studies in Philosophy and Law. Cambridge ; New York: Cambridge University Press.
- Foot, Philippa. 1958. “Moral Beliefs.” *Proceedings of the Aristotelian Society*, New Series, 59 (January): 83–104.
- Foucault, Michel, and Jeremy R. Carrette. 1999. *Religion and Culture*. Taylor & Francis.
- Frank, Joseph. 2002. *Dostoevsky*. Princeton University Press.
- Friedman, Marilyn. 2013. “How to Blame People Responsibly.” *The Journal of Value Inquiry* 47 (3): 271–284.
- Furley, D. 1973. *Bulletin of the Institute of Classical Studies of the University of London*. Bulletin of the Institute of Classical Studies.
- Garet, Mary. 1998. “Comparative Rhetoric: An Historical and Crosscultural Introduction by George Kennedy.” *Rhetorica, Rhetorica*, 4 (Autumn): 431–54.
- Gert, Bernard. 1998. *Morality: Its Nature and Justification*. Revised ed. edition. New York: Oxford University Press.
- . 2006. *Bioethics: A Systematic Approach*. Second edition. New York ; Oxford: Oxford University Press.
- Gleason, Maud W. 2008. *Making Men: Sophists and Self-Presentation in Ancient Rome*. Princeton, NJ: Princeton University Press.
- Gorgias. 1982. *Encomium of Helen*. Bristol Gloucestershire: Bristol Classical. <http://hdl.handle.net/2027/uc1.32106008081967>.
- Green, Bonnie L. 1990. “Defining Trauma: Terminology and Generic Stressor Dimensions 1.” *Journal of Applied Social Psychology* 20 (20): 1632–1642.

- Greenburg, Karen J. 2009. "Detention Nation." *The National Interest*, no. 101(January): 49.
- Griswold, Charles L. 1988. *Platonic writings/Platonic Readings*. New York: Routledge.
- Gross, Michael L. 2010. *Moral Dilemmas of Modern War: Torture, Assassination, and Blackmail in an Age of Asymmetric Conflict*. Cambridge ; New York: Cambridge University Press.
- Grote, George. 2010. *A History of Greece from the Earliest Period to the Close of the Generation Contemporary with Alexander the Great*. Cambridge Scholars Publishing.
- Guyer, Paul. 1992. *The Cambridge Companion to Kant*. Cambridge University Press.
- Hare, Richard Mervyn. 1981. *Moral Thinking: Its Levels, Method, and Point*. Oxford University Press.
- Hornblower, Simon, and Antony Spawforth. 2012. *The Oxford Classical Dictionary*. OUP Oxford.
- Huggins, Martha Knisely, Mika Haritos-Fatouros, and Philip G. Zimbardo. 2002. *Violence Workers: Police Torturers and Murderers Reconstruct Brazilian Atrocities*. University of California Press.
- Hunt, Peter. 2010. *War, Peace, and Alliance in Demosthenes' Athens*. 1st ed. Cambridge University Press.
- Hursthouse, Rosalind. 2011. "Virtue Ethics." Accessed April 2. <http://plato.stanford.edu/entries/ethics-virtue/>.
- Information, UN. Department of Public. 1987. *Universal Declaration of Human Rights* (1948). UN.
- Jackson, Frank, and Michael Smith. 2008. *The Oxford Handbook of Contemporary Philosophy*. Oxford University Press.
- Joffe, Hélène. 1999. *Risk and "the Other."* Cambridge, UK ; New York, NY, USA: Cambridge University Press.
- Juratowitch, Ben. 2008. "Torture Is Always Wrong." *Public Affairs Quarterly* 22 (2): 81–90.

- Kahn, Charles H. 1998. *Plato and the Socratic Dialogue: The Philosophical Use of a Literary Form*. Cambridge University Press.
- Kastely, James L. 1997. *Rethinking the Rhetorical Tradition: From Plato to Postmodernism*. Yale University Press.
- Kennedy, George A. 2009. *A New History of Classical Rhetoric*. Princeton University Press.
- Lamb, Sharon. 1996. *The Trouble with Blame: Victims, Perpetrators, and Responsibility*. Cambridge, MA: Harvard University Press.
- Lazar, Seth. 2012. "Necessity in Self-Defense and War." *Philosophy & Public Affairs* 40 (1): 3–44.
- Levinson, Sanford. 2004. *Torture: A Collection*. Oxford University Press US.
- Levitt, Brad. 2005. "Platonic Parody in the Gorgias." *Phoenix* 59 (3/4): 210–27.
- Lewis, Helen B. 1971. *Shame and Guilt in Neurosis*. New York: International Universities Press.
- Litz, Bret et al. 2009. "Moral Injury and Moral Repair in War Veterans: A Preliminary Model and Intervention Strategy. - PubMed - NCBI." *Clinical Psychology Review* 29 (8): 695–706.
- Lorenz, Hendrik. 2006. *The Brute within: Appetitive Desire in Plato and Aristotle*. Oxford Philosophical Monographs. Oxford: Clarendon Press ; Oxford ; New York.
- . 2009. "Ancient Theories of Soul." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2009. <http://plato.stanford.edu/archives/sum2009/entries/ancient-soul/>.
- Luban, David. 2005. "Liberalism, Torture, and the Ticking Bomb," *Georgetown Law Faculty Publications and Other Works*
- . 2007. "The Defense of Torture." *The New York Review of Books*, March 15.
- . 2014. *Torture, Power, and Law*. Cambridge, United Kingdom: Cambridge University Press.

- Macnair, Rachel M. 2002. "Perpetration-Induced Traumatic Stress in Combat Veterans." *Peace and Conflict: Journal of Peace Psychology* 8 (1): 63–72. doi:10.1207/S15327949PAC0801_6.
- Maguen, Shira, Barbara A. Lucenko, Mark A. Reger, Gregory A. Gahm, Brett T. Litz, Karen H. Seal, Sara J. Knight, and Charles R. Marmar. 2010. "The Impact of Reported Direct and Indirect Killing on Mental Health Symptoms in Iraq War Veterans." *Journal of Traumatic Stress*. doi:10.1002/jts.20434.
- May, Larry. 2016. *War: Essays in Political Philosophy*. Cambridge University Press.
- McMahan, Jeff. 2006. "Killing in War: A Reply to Walzer." *Philosophia* 34 (1): 47–51.
- . 2008. "Torture in Principle and in Practice." <http://philpapers.org/rec/MCMTIP>.
- . 2013. "Proportionate defense. (Richard B. Lillich Memorial Lecture)(New Frontiers in the Laws of War)." *Journal of Transnational Law & Policy* 23: 1–36.
- . 2016. "Proportionality and Necessity in Jus in Bello." *The Oxford Handbook of Ethics of War*, Seth Lazar and Helen Frowe. April 2016. Oxford Handbooks Online. <http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199943418.001.0001/oxfordhb-9780199943418-e-24>.
- Meagher, Robert E. 2014. *Killing from the inside out: Moral Injury and Just War*. Eugene, Oregon: Cascade Books.
- Nadeau, Ray. 2015. *Rhetorica Ad Herennium: Commentary and Translation of*, Google Books. University of Michigan Press. Accessed June 22.
- Nails, Debra. 2002. *The People of Plato: A Prosopography of Plato and Other Socratics*. Hackett Pub Co Inc.
- Ober, Josiah. n.d. *Mass and Elite in Democratic Athens: Rhetoric, Ideology, and the Power of Persuasion*. Princeton University Press.
- Pardo, Pablo. 2009. "Torturas: Un Testimonio En Primera Persona | El mundo.es." *El Mundo*. April 24. <http://www.elmundo.es/elmundo/2009/04/24/cronicasdesdeeeuu/1240562878.html>.
- Phillips, Joshua E. S. 2010. *None of Us Were like This Before: American Soldiers and Torture*. Verso.

- . 2011. "Inside the Detainee Abuse Task Force." *The Nation*, May 13. <http://www.thenation.com/article/160560/inside-detainee-abuse-task-force?page=full>.
- Pickard, Hanna. 2011. "Responsibility Without Blame: Empathy and the Effective Treatment of Personality Disorder." *Philosophy, Psychiatry, & Psychology* 18 (3): 209–224.
- . 2013. "Irrational Blame." *Analysis* 73 (4): 613–26. doi:10.1093/analysis/ant075.
- Plato. 1914. *Euthyphro. Apology. Crito. Phaedo. Phaedrus — Plato*. Harold North Fowler Editor. The Loeb Classical Library. No. 183. Harvard U Press.
- . 1980. *Gorgias*. Translated by Terence Irwin. Oxford University Press, USA.
- . 1990. *Gorgias*. Edited by E. R. Dodds. Oxford University Press, USA.
- . 2009. *Gorgias, Menexenus, Protagoras*. Schofield, Malcolm. Cambridge University Press.
- . 1924. *Laches Protagoras Meno Euthydemus*. James Loeb, and W.R.M. Lamb Editor. The Loeb Classical Library. No.165,; Harvard University Press.
- . 1967. *Plato in Twelve Volumes*, Vol. 3 trans W.R.M. Lamb. Cambridge, MA, Harvard University Press; London, William Heinemann Ltd. Perseus Digital Library. <http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.01.0178%3Atext%3DGorg>.
- . 2013. *The Republic*. Edited by C. J. Emlyn-Jones and William Preddy. Loeb Classical Library ; 237, 276. Cambridge, MA: Harvard University Press.
- . 1935. *The Republic*. Loeb Classical Library. Cambridge, Mass: Harvard University Press ; London.
- Plochmann, George Kimball, and Franklin Edward Robinson. 1988. *A Friendly Companion to Plato's Gorgias*. SIU Press.
- Quiroga, Jose. 2008. "Torture." *The Encyclopedia of Psychological Torture*. John Wiley & Son.
- Randall, Glenn, and Ellen L. Lutz. 1991. *Serving Survivors of Torture: A Practical Manual for Health Professionals and Other Service Providers*. American Association for the Advancement of Science.

- Raz, Joseph. 1999. *Engaging Reason: On the Theory of Value and Action*. Oxford University Press US.
- . 2007. “Reasons: Explanatory and Normative.”
- Rejali, Darius. 2009. *Torture and Democracy*. Princeton University Press.
- Roisman, Joseph, and Ian Worthington. 2011. *A Companion to Ancient Macedonia*. John Wiley & Sons.
- Roochnik, David. 1996. *Of Art and Wisdom: Plato’s Understanding of Techne*. University Park, PA: Pennsylvania State University Press.
- Rowe, Christopher. n.d. “A Problem in the Gorgias: How Is Punishment Supposed to Help with Intellectual Error?”
- Ryvicker, Miriam. 2009. “Preservation of Self in the Nursing Home: Contradictory Practices within Two Models of Care.” *Journal of Aging Studies* 1 (23): 12–23. doi:10.1016/j.jaging.2007.09.004.
- Sabini, John, and Maury Silver. 1982. *Moralities of Everyday Life*. Oxford University Press.
- Sandler, T., and Keith Hartley, eds. 1995. *Handbook of Defense Economics*. 1 edition. Amsterdam ; New York: North Holland.
- Scanlon, Thomas. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass: Belknap Press of Harvard University Press.
- Schacter, Daniel L. 2000. *Memory, Brain, and Belief*. Cambridge, Mass: Harvard University Press. <http://hdl.handle.net/2027/mdp.39015042476955>.
- Segal, Charles P. 1962. “Gorgias and the Psychology of the Logos.” *Harvard Studies in Classical Philology* 66: 99–155.
- Shay, Jonathan. 1994. *Achilles in Vietnam: Combat Trauma and the Undoing of Character*. First edition. New York : Toronto : New York: Atheneum.
- . 2002. *Odysseus in America: Combat Trauma and the Trials of Homecoming*. New York: Scribner.
- . 2014. “Moral Injury.” *Psychoanalytic Psychology* 31 (2): 182–191.

doi:10.1037/a0036090.

- Sherman, Nancy. 2010. *The Untold War: Inside the Hearts, Minds, and Souls of Our Soldiers*. Reprint edition. W. W. Norton & Company.
- Strawson, P. F. 1974. *Freedom and Resentment, and Other Essays*. London: Methuen distributed in the USA by Harper & Row, Barnes & Noble Import Division.
- Strayer, R. 1975. "Vietnam Veterans: A Study Exploring Adjustment Patterns and Attitudes," 31, 31 (4): 81–93.
- Sussman, David. n.d. "What's Wrong With Torture?" *Philosophy & Public Affairs* 33 (1): 1–33.
- Tännsjö, Torbjörn. 2009. *From Reasons to Norms: On the Basic Question in Ethics*. Springer.
- Tarnopolsky, Christina H. 2010. *Prudes, Perverts, and Tyrants: Plato's Gorgias and the Politics of Shame*. Princeton University Press.
- The Nation*. 2011. "Part II: The Soldier Who Chose Suicide After She Refused to Go Along With Torture." Accessed December 30. <http://www.thenation.com/blog/154648/part-ii-soldier-who-chose-suicide-after-she-refused-go-along-torture>.
- Thomson, Judith Jarvis, Philip Fisher, Martha C. Nussbaum, and Barbara Herrnstein Smith. 2003. *Goodness and Advice*. Princeton University Press.
- Vargas, Alison Flipse, Thomas Hanson, Douglas Kraus, Kent Drescher, and David Foy. 2013. "Moral Injury Themes in Combat Veterans' Narrative Responses From the National Vietnam Veterans' Readjustment Study." *Traumatology* 19 (3): 243–50. doi:10.1177/1534765613476099.
- Waldron, Jeremy. 2010. *Torture, Terror, and Trade-Offs: Philosophy for the White House*. Oxford ; New York: Oxford University Press.
- Walker, Margaret Urban. 2003. *Moral Contexts. Feminist Constructions*. Lanham, Md: Rowman & Littlefield Publishers.
- . 2006. *Moral Repair: Reconstructing Moral Relations after Wrongdoing*. 1 edition. Cambridge University Press.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, Mass: Harvard University Press.

Walzer, Michael. 1973. "Political Action: The Problem of Dirty Hands on JSTOR." *Philosophy & Public Affairs* 2 (2): 160–80.

———. 1977. *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. New York: Basic Books.

Weiner, Ross L. 2008. "Office of Legal Counsel and Torture: The Law as Both a Sword and Shield, The." *George Washington Law Review* 77: 524.

Williams, Bernard. 1982. *Moral Luck*. Cambridge Cambridgeshire; New York: Cambridge University Press.

Williams, Kayla, and Michael E. Staub. 2006a. *Love My Rifle More Than You: Young and Female in the U.S. Army*. W. W. Norton & Company.

Wolfendale, Jessica. 2007a. *Torture and the Military Profession*. Palgrave Macmillan.

Wooten, Cecil W., editor. 2001. *The Orator in Action and Theory in Greece and Rome*. Essays in Honor of George A. Kennedy. Leiden: Brill.

Yoo, John. 2006. *War by Other Means: An Insider's Account of the War on Terror*. Atlantic Monthly Press.