Moral Disillusion

Shattering Moral Illusions for the Sake of Taking Responsibility

by

Zachary J. Goldberg

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

ARIZONA STATE UNIVERSITY

May 2012

ABSTRACT

I present in this dissertation a theory of moral disillusion. In chapter 1 I explain moral innocence and its loss. I show that becoming morally responsible requires shattering the illusion that one is not an appropriate candidate for the reactive attitudes. The morally responsible individual must understand that she can be an agent of wrongdoing. In chapter 2 I explore the nature of the understanding that accompanies the different phases of disillusion. I show that moral disillusion is an ability, not to follow moral principles, but to question them. In chapter 3 I argue that another phase of disillusion involves an acquaintance with evil. One shatters the illusion that only malicious individuals can be evildoers. Morally good people can also bring about evil. I conclude that evil is the exploitation of the extremely vulnerable. In chapters 4 and 5, I analyze more complex phases of moral disillusion. These stages are characterized by an understanding that one can be an agent of unchosen evil, that one might bring about evil even when pursuing the morally best course of action, and that one can be morally responsible for doing so. In order to understand unchosen evil and the tragedy of inescapable moral wrongdoing, the individual sees that moral responsibility ought to track what we care about, rather than what we believe. In chapter 6 I show that Kierkegaard's conception of the self is a philosophy of moral disillusion. I argue that his prescription that we shatter moral illusions is congruent with Harry Frankfurt's prescription that we ought to care about some things and not others. From this discussion emerges the explicit distinction between moral disillusion and moral goodness. Moreover, I conclude that the

i

morally disillusioned are morally accountable for more than those still harboring

moral illusions. Although moral disillusion does not entail becoming morally

good, by acquiring the ability to raise questions about moral principles and to

affect the content of one's cares, one acquires the ability to take responsibility for,

and potentially minimize, evil. To have and understand these abilities, but not to

care about them, increases one's moral accountability.

DEDICATION

For Alena and Ilja

ACKNOWLEDGMENTS

TABLE OF CONTENTS

v

CHAPTER 1

INTRODUCTION

It is an interesting historical fact that some of the greatest evils have been committed in the name of the most ambitious and inflexible ethical norms. The history of Christianity, Communism, and Nazism supply sufficient evidence of this fact. I take it that many of the followers of these doctrines who committed horrible atrocities truly believed that they were doing the morally right thing. But clearly, they acted under an illusion. Thus, it is evident that holding fast to moral illusions can have dreadful consequences. However, it is not only morally corrupt ideologies, like Nazism or Communism, that generate illusions. Legitimate moral theories can also engender illusions. Holding moral illusions keeps one from taking responsibility for the evil one brings about.

Because perpetrators of evil have often been motivated by an illusory morality rather than no morality at all, it is important to provide a theory of moral disillusion, which I do in the following chapters. By "disillusion" I mean the shattering of moral illusions, not becoming morally good. I do not provide here a theory of moral goodness. It would be simplistic to claim that moral disillusion entails moral goodness. The negative act of shattering illusions does not necessarily include the positive act of forming and maintaining a positive and non-illusory ethics. One might replace one set of moral illusions for another. But moral disillusion allows at least for the possibility of taking responsibility for one's actions. Part of moral disillusion is recognizing the inadequacy of many moral principles and moral directives to account for the evil we bring about. One

sees this inadequacy by examining the locus of our responsibility to others, how our lives are vulnerable to evil that we cause, and how wrongdoing can be unavoidable. It is impossible to address the evil we cause if we remain ignorant of how it comes about. But once made aware of evil's nature and genesis, we can take responsibility for its occurrence.

Taking responsibility is important because it allows one to respond appropriately to the harm one has, either directly or indirectly, caused. Sometimes the appropriate response to a harm is to ensure that one does not repeat it in the future. For example, when an individual recognizes that her lies hurt others, she may resolve not to lie in the future. But sometimes the harm we do to others is unchosen and inescapable. Due to such features of the harm, the individual cannot ensure that she will not bring it about in the future. In these cases, the appropriate response is regret, remorse, confession, contrition, apology, restitution, or reparation. Whichever of these attitudes fits the specific circumstances, the individual takes responsibility for her actions. In order to take responsibility, one must shatter illusions of moral purity.

Moral disillusion is a process, not a status. It involves various phases of coming to situate wrongdoing and evil in terms of one's own agency. To do so one must shatter the illusions promulgated by many ethical theories that it is possible to cultivate for oneself a state of moral purity. Such theories maintain that it is possible, although admittedly unlikely, to avoid all wrongdoing as long as one follows certain universal moral principles. Judeo-Christian morality, Kantian ethics, and utilitarianism clearly fall into this category. It is an essential

element of these theories that wrongdoing is avoidable. But to claim that moral

purity is a possibility is to ignore certain facts about the moral landscape. It is to

ignore the complexity of moral life; it is to harbor a moral illusion. One can

shatter such illusions by raising questions about moral principles that purport to

establish moral purity.

The disillusions I consider most important and will analyze in the present

work include understanding the following: one's capacity for wrongdoing, the

nature of evil, unchosen evil, inescapable wrongdoing, and the importance of self-

reflection and what we care about.

I begin my analysis with the disillusion that characterizes becoming a

morally responsible agent. At this phase we shatter the illusion that our actions do

not matter, and we realize that we are appropriate candidates of the reactive

attitudes. This understanding allows us entry into the moral community. But

becoming morally responsible is only one phase in a larger process of disillusion.

We may, and probably do, still harbor other moral illusions. We might believe

that we need only follow God's laws, or the categorical imperative, or the greatest

happiness principle. But rule-following proves to be inadequate for a world rife

with evil.

Evil can be brought about unchosenly. Moral disillusion involves realizing

that one can be held morally responsible for bringing about unchosen evil.

Furthermore, one might be faced with a situation in which wrongdoing and

evildoing are inescapable. Even by choosing the morally best course of action, the

agent may transgress a moral value. Thus, it is not the case that following the ten

commandments, or bringing about an optimific state of affairs, or respecting another's autonomy, sufficiently describe moral obligation in every set of circumstances. Our characters and relationships are complex enough to give rise to moral obligations that preclude the possibility of moral purity.

Understanding our characters and relationships involves understanding what we care about. Cares help structure the will and as such are powerful motivators of action. They are powerful enough that if what we care about conflicts with what we believe, then they sometimes motivate us to act despite our beliefs. Because much of our acts have moral import, and our acts are often motivated by what we care about, then it follows that what we care about has moral import as well. Although the content of our cares is not entirely within our control, the understanding that one acquires through the various stages of moral disillusion can affect what one cares about. Thus, the morally disillusioned understanding is a kind of ability, or abilities. It is an ability to participate more fully in the moral community through an awareness of the nuances of one's moral interaction with others, an ability to question the adequacy of moral principles and directives, and an ability to care about the evil one brings about. This kind of ability allows one to take responsibility for, and even potentially minimize, the evil one causes.

But understanding the evil one brings about does not make a person morally good. Moral disillusion does not entail moral goodness. As one acquires a more developed understanding of moral life, one might care about bringing about more evil, or one might understand, but not care at all about the evil one

4

causes. Therefore, I do not claim that shattering moral illusions entails becoming morally good. At each phase of disillusion the agent becomes, not (necessarily) morally good, but critical. Since people can do evil by following moral directives, it is often more important to raise questions about moral behavior than to follow moral directives.

Although moral disillusion does not entail moral goodness, I do think that it brings with it an enhanced moral responsibility. The person who understands how unchosen evil occurs, that situations of inescapable wrongdoing arise, that moral responsibility should track our cares rather than our beliefs, and yet does not care about taking responsibility for, and potentially minimizing, evil is more morally responsible for the evil she causes than a person who lacks an understanding of these features of moral life. Those devoid of moral illusions are to be held morally accountable for much more than those who retain them. The reason for this is not because they ought to know better, but because they do know better, and do not care.

I will argue for these conclusions by presenting a theory of moral disillusion. In chapter 1 I explain moral innocence and its loss. This is the first stage of the process of moral disillusion. I show that becoming morally responsible requires shattering the illusion that one is not an appropriate candidate for the reactive attitudes. The morally responsible individual must understand that she can be an agent of wrongdoing. In chapter 2 I explore the nature of the understanding that accompanies the different phases of disillusion. Following Wittgenstein's analysis of understanding as an ability to go on, I show that moral

disillusion is an ability, not to follow moral principles, but to question them. In chapter 3 I argue that another phase of disillusion involves an acquaintance with evil. One shatters the illusion that only malicious individuals can be evildoers. Morally good people can also bring about evil. By examining particular examples of evil as well as theories of evil, I conclude that evil is the exploitation of the extremely vulnerable that results in irreparable life-wrecking harm. In chapters 4 and 5, I analyze more complex phases of moral disillusion. These stages are characterized by an understanding that one can be an agent of unchosen evil, that one might bring about evil even when pursuing the morally best course of action, and that one can be morally responsible for doing so. In order to understand unchosen evil and the tragedy of inescapable moral wrongdoing, the individual sees that moral responsibility ought to track what we care about, rather than what we believe. In chapter 6 I show that Kierkegaard's conception of the self is a philosophy of moral disillusion. His philosophy shows that the ability to question moral principles is essential to taking responsibility for what one does. Further, I argue that his prescription that we shatter moral illusions is congruent with Harry Frankfurt's prescription that we ought to care about some things and not others. From this discussion emerges the explicit distinction between moral disillusion and moral goodness. Moreover, I conclude that the morally disillusioned are morally accountable for more than those still harboring moral illusions. Although shattering moral illusions does not entail becoming morally good, by giving one the ability to raise questions about moral principles and to affect the content of one's cares, one acquires the ability to take responsibility for, and potentially

minimize, the evil one causes. To have and understand these abilities, but not to

care about them, makes one more morally accountable.

CHAPTER 2

SHATTERING THE ILLUSIONS OF MORAL INNOCENCE

1. Introduction

In this chapter I describe the shattering of moral illusions that comprise moral innocence. Moral innocence characterizes those who cannot be held morally responsible. In *Responsibility and Control*, Fischer and Ravizza explain that in order to enter the moral community, the morally innocent must recognize that their actions have an upshot in the world and that they are appropriate candidates for the reactive attitudes. This recognition requires shattering the innocent notion that one is immune from blame. In his essay "Losing Innocence for the Sake of Responsibility," Peter French argues that the morally innocent must be acquainted with evil in order to shatter the illusions that comprise moral innocence and enter the moral community. Like Fischer and Ravizza, I think that an acquaintance with something milder than evil is required for entry into the moral community. The morally innocent must understand wrongdoing in terms of their own agency and develop the reactive attitudes; they do not necessarily need to become acquainted with evil. But French's argument is significant for two reasons. First, he provides a phenomenological description of the shattering of innocent moral illusions necessary for moral responsibility. Second, although I disagree with his definition of evil, he reminds us that moral life is far richer than barely being the appropriate target of the reactive attitudes. Harm to others cannot be completely understood by examining wrongdoing. There is also evil in the world. Moreover, this evil is sometimes brought about by our actions. Hence, the

8

illusions that must be shattered in order to provide entry into the moral

community are not the only moral illusions we harbor. By juxtaposing Fischer

and Ravizza's argument for the necessary conditions of moral responsibility with

French's account of the loss of innocence via the acquaintance with evil, I begin

to show that fully understanding moral life requires different shatterings of

various moral illusions. In this chapter I focus on one phase of moral disillusion:

recognizing the ability to do wrong. I contrast moral innocence with simple

ignorance of moral concepts to show that moral innocence is an inability: an

inability to participate in the moral community. Finally, I argue throughout the

chapter that the illusions that comprise moral innocence are not the only moral

illusions we may have. I begin to show in this chapter, and give a full description

in chapter 3, that an acquaintance with our ability to do or be complicit in the

occurrence of evil is also a significant phase of moral disillusion.

2. Moderate Reasons-Responsiveness

In *Responsibility and Control* Fischer and Ravizza provide the structure of

the functional properties of moral responsibility.[1] Structurally, moral

responsibility has two necessary elements. An agent is morally responsible for an

action when the action issues from the agent's own moderate reasons-responsive

mechanism; that is, it is necessary that the mechanism is the agent's own and that

it is moderate reasons-responsive. What it means for an agent's mechanism to be

moderate reasons-responsive is that the mechanism is reasons-receptive in a

strong way and at least reasons-reactive in a weak way. The mechanism is

---

[1] John Martin Fischer and Mark Ravizza, *Responsibility and Control* (New York: Cambridge University Press, 1998).

receptive in a strong way when it is regularly receptive to reasons. The agent

must be able to respond in a regular fashion to reasons so that her responses and

actions create a pattern or seem logical. She need not always act on these reasons

but it must be possible that she do so. The mechanism is reactive in weak way

when it can react to a sufficient reason to do otherwise. It need not actually react

to a sufficient reason, but if there is a possible world where it does, then it is weak

reasons-reactive. For example, it is conceivable that a car thief reacts to a moral

reason to which she is receptive not to steal the car, and yet she steals the car.

Then she is morally responsible for the car theft even if it is only a remotely

conceivable possibility that she would react appropriately to the moral reason she

recognizes, or is receptive to, as sufficient for not stealing the car by not stealing

it. Combining strong reasons-receptivity with this weak reasons-reactivity forms

moderate reasons-responsiveness.

However, moderate reasons-responsiveness is not alone sufficient for

moral responsibility. The agent's mechanism could be subject to responsibility-

undermining conditions. Therefore, Fischer and Ravizza introduce the second

necessary element of moral responsibility: the action-issuing mechanism must be

the agent's own. "Taking responsibility" is a necessary condition of the action-

issuing mechanism becoming one's own. Taking responsibility has three

necessary elements: the agent recognizes that her actions have some "upshot" in

the world; the agent recognizes that she is an appropriate candidate for the

reactive attitudes; and these two beliefs must be based on appropriate evidence.

This last element serves to rule out freedom-undermining conditions such as

brainwashing, behavioral engineering and the like. Since this aspect of their argument, viz. the metaphysics of free will, is not my present concern, I will concentrate on the first two components of taking responsibility.

Fischer and Ravizza explain that a child comes to realize that her actions have some upshot in the world through her moral education. When a young boy tears open a present belonging to the birthday girl, his parents might send him to his room.[2] This punishment serves two purposes. First, it helps teach the boy that the present did not open on its own. Rather, it opened due to an exercise of his agency. He starts to learn that his agency has certain upshots in the world; that he is the source of certain events and that these events are not caused by accidents or other agents. Second, the punishment shows the boy that others feel indignation toward him in these circumstances. He starts to learn that he can be fairly praised or blamed for his behavior. He sees that he is an appropriate candidate for the reactive attitudes.

These two elements of moral responsibility, moderate reasons-responsiveness and taking responsibility for the agent's own action-issuing mechanism, give us the structure of what is required for an agent to be held morally responsible for her actions. As we can see with the example of the boy with the presents, in order for a person to acquire the capacities necessary for moral responsibility, certain illusions must be shattered. One shatters the illusions that one's actions have no upshot in the world and that one cannot be blamed for one's actions.

---

[2] Fischer and Ravizza, 208.

The way in which we come to shatter innocent illusions is described by Peter French in his essay, "Losing Innocence for the Sake of Responsibility." His account provides us with a more developed understanding of the phenomenological content of "taking responsibility" of which Fischer and Ravizza supply merely the structure. French argues that one must become acquainted with evil in order to enter the moral community. He defines evil as undeserved harm. Understood in this way, the acquaintance with evil is indeed necessary to enter the moral community. But understood in this way, evil is not distinct from wrongdoing. The boy who opens the birthday girl's presents causes her undeserved harm. Part of realizing that one's actions have upshot in the world, and that one is an appropriate candidate for the reactive attitudes is to recognize that one can do wrong. But recognizing one's ability to do wrong does not sufficiently account for the all of the degrees of badness one can bring about. In chapter 3, I argue that evil is distinct from wrongdoing and that recognizing one's capacity to do or be complicit in the occurrence of evil, is a significant moral disillusion. Nevertheless, French's essay has a twofold significance for my discussion. First, he provides us with a phenomenological account of the shattering of the illusions necessary for entry into the moral community. Second, he reminds us that evil, not just wrongdoing, exists in the world.

3.   Innocence and its Loss

Before giving my account of moral innocence it would be helpful for me to clarify what I do *not* mean by the term "innocence". Certain authors equate moral innocence either with the status of moral purity or with not being morally

blameworthy for a particular act in specific circumstances. For example, Peter

Johnson claims that "innocence as moral purity implies an inability to inflict

harm."[3] The innocent are those who do not have the ability to do wrong. Elizabeth

Wolgast argues against those who claim that innocence as moral purity is

something to be valued.[4] Nevertheless, she still defines innocence as a state of

moral purity. In legal scenarios "innocent" means not guilty for a particular act.

To be found innocent of a crime does not mean that one is morally pure, but

simply that one did not commit, or is not culpable for committing, a particular act.

But both of these approaches are misleading when it comes to

understanding moral innocence. The legal definition of innocence is appropriate

for legal cases but not for understanding moral innocence. By moral innocence we

mean a certain status, rather than being not guilty for committing a particular act.

Thinking of moral innocence as moral purity also misses the mark. It does not

capture what we mean when we think of moral innocence. French notes that

"[m]oral purity, if there is such a thing, is determined by evaluating someone

from the perspective of or against the standards set by moral rules and

principles….Moral virginity is the condition or state of not being a proper subject

of those standards."[5] He defines moral innocence, not as moral purity, but as

moral virginity.[6] It is not that the innocent are morally pure, but morally virginal.

---

[3] Peter Johnson, *Politics, Innocence and the Limits of Goodness* (New York & London: Routledge, 1988).

[4] Elizabeth Wolgast, "Innocence." *Philosophy* 68, no. 265 (Jul., 1993): 297-307

[5] French *Responsibility Matters*, 30.

[6] We might also construe "moral purity" as moral sainthood. See J.O. Urmson "Saints and Heroes" in *Essays in Moral Philosophy*, ed. A.I. Melden (Seattle: University of Washington Press, 1958) and Susan Wolf "Moral Saints" in *Virtue Ethics*, ed. Roger Crisp and Michael Slote (New York: Oxford University Press, 1997).

An innocent has not yet acquired the cognitive and volitional structures necessary to be morally responsible. I will not argue here for or against the value of moral virginity; rather, I will describe it and explain why understanding it is helpful to understanding moral disillusion.

According to French, being innocent is characterized by lacking a descriptive framework rich with the vocabulary of responsibility. He rejects the view that Adam and Eve lost their innocence by being made aware of their own freedom through an act of disobedience. French rightly points out that children can be disobedient and still maintain their innocence. Therefore, disobedience is not necessary to a loss of innocence. What must occur when innocence is lost is a conceptual shift that promotes describing familiar things in a new way, a way laden with the language of responsibility. The illusions of innocent description, description of the moral world that does not cohere to reality, must be shattered. This conceptual shift occurs after confronting evil and the recognizing that one can both do and be done evil.[7] I argue in section 5 of this chapter, and in chapter 2, that this conceptual shift is not merely cognitive. It is volitional as well, and it also gives the agent a certain ability to participate in the moral community.

This acquisition of the requisite descriptive framework can be seen by examining the difference between seeing and seeing that.[8] An innocent sees something without always seeing that. For example, one can see a child being led into a van without seeing that the child is being kidnapped. Seeing that involves a

---

[7] Throughout my presentation of French's argument, I follow his use of the word "evil". However, I think that an acquaintance with wrongdoing is sufficient for moral responsibility. Nevertheless, the acquaintance with evil is an important moral disillusion that I will discuss more in chapter 3.

[8] I examine this distinction in detail in chapter 2.

conceptual understanding or insight into the significance of an event or state of affairs. Innocence means not having the training to use moral conceptual descriptions with insight. Loss of innocence means acquiring this insight and the ability to describe moral and immoral persons and actions.

Not only must the loss of innocence have a certain content, but it also must occur in a certain way. French relies on Russell's distinction between knowledge by description and knowledge by acquaintance to argue that descriptions of morality will not always be sufficient to promote entry into the moral community. What is required is an acquaintance, specifically an acquaintance with evil.

> Acquiring knowledge of good and evil…relevant to the loss of innocence must be different from the acquisition of what we call purely objective knowledge, like knowledge of geometry, which is merely a matter of learning the rules and the angles. Suppose innocents were taught Kant's categorical imperative, all three formulations, and then Bentham's utility calculus. Would that transform them into mature members of the moral community? It might bore them to tears, but they would not be crying for the loss of innocence….The learning has to be about oneself and it has to be, in some sense, active. The illusions that are shattered must be your own…[9]

We can clearly imagine how ineffectual it would be to teach moral responsibility simply by describing various ethical theories to someone unfamiliar with the importance of moral action. In fact, one reason someone may take interest in studying an ethical theory in the first place is that the person already has some

---

[9] Ibid., 39-40.

acquaintance with evil and thereby acknowledges the importance of ethical

behavior both on her own part as well as others. Before this acquaintance takes

place, it is not just that attention to the intricacies of Kantian ethics or

Utilitarianism would fall on deaf ears, rather the student would not grasp the

significance of moral vocabulary at all. The student cannot *go on* in the

conversation. The innocent must recognize through acquaintance with evil that

acts often have moral significance. More importantly she must realize that *her*

acts often have moral significance. There is no specific standard of evil or

threshold of acquaintance that can be set in advance for the innocent to transcend.

Rather innocence is a scalar notion, and one person's push into the moral

community might not work for another. What is necessary though is that the

innocent's illusions concerning the moral order are shattered.

Moral disillusion is a process. It is a scalar process that has many phases.

The phase of disillusion necessary for moral responsibility is the recognition that

one's actions have upshot in the world and that one is an appropriate candidate for

the reactive attitudes. But one can shatter the illusions of moral innocence and still

retain other moral illusions. It follows then that moral responsibility is not

coextensive with complete moral disillusion. I agree with Fischer and Ravizza

that the process of becoming a morally responsible agent is scalar, and I will show

in the following chapters that there are significant phases of disillusion beyond

that required for moral responsibility.

To see that becoming a morally responsible agent is only one phase on a

scale of moral disillusion, let's define the pertinent terms. In most discussions of

16

moral responsibility, moral agency is defined in terms of the capacity to be held morally responsible for one's actions. An agent is considered a moral agent insofar as she may be held responsible for what her agency brings about. A moral agent is "one who qualifies generally as an agent open to responsibility ascriptions."[10] Based on this view, a moral agent simply is an agent who can be held morally responsible for her actions. It is, however, possible to distinguish moral agency from moral responsibility if we distinguish between attributability and accountability.[11] We can attribute an action to an agent even if we do not hold the agent accountable for it. If the action has a moral quality or moral upshot, then we can attribute a moral action to the agent. Attributing a moral action to an agent does not entail finding the agent accountable for the action. Children and the mentally handicapped are capable of performing actions with moral qualities or moral upshot. Therefore, we can attribute moral actions to them as agents. However, if a certain lack of guidance control or lack of understanding accompanies the actions of children or the mentally handicapped, then we usually do not find them morally accountable. In this sense, someone can be a moral agent, but not be an appropriate candidate for ascriptions of moral responsibility. Hence, moral agency can be distinguished from moral responsibility.

One might reply that this kind of agency is not moral agency as we ordinarily understand it. When we say that someone is a moral agent we mean that she is an agent whose actions are open to responsibility ascriptions. Either

---

[10] Andrew Eshleman, "Moral Responsibility", *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, ed. Edward N. Zalta, URL = <http://plato.stanford.edu/archives/win2009/entries/moral-responsibility/>.
[11] See Gary Watson, "Two Faces of Responsibility," *Philosophical Topics* 24 (1996): 227-248.

approach is consistent with my present account. Whichever way we define the connection between moral agency and moral responsibility, they do not comprise full moral disillusion. Even after becoming a morally responsible agent, one might still hold fast to certain moral illusions.

Whether moral agency is defined in terms of moral responsibility or considered distinct, the necessary and sufficient conditions of moral agency and of moral responsibility (whether the same or not) constitute only one phase of moral disillusion. Moral agency is defined either as capable of bringing about actions that have a moral quality, or one who qualifies generally as an agent open to responsibility ascriptions. Moral responsibility is defined as being an appropriate candidate for the reactive attitudes and being moderate-reasons responsive. I will argue in chapter 3 that a significant moral disillusion is understanding that one can do and be done evil. But the acquaintance with evil is not necessary for becoming a moral agent or morally responsible; it is necessary for losing a certain moral illusion. Hence, one can be a moral agent and morally responsible and still retain some moral illusions.

In the example of the boy opening the birthday girl's presents, Fischer and Ravizza do not treat the boy as a responsible moral agent, but as one who has an opportunity to begin the process of becoming a responsible moral agent by being the target of very mild reactive attitudes. He gets acquainted with the notions of right and wrong and the fact that behaving wrongly raises the reactive attitudes in others. The boy probably does not apply any reactive attitudes to himself. He may not really feel guilty or even grasp what feeling guilty involves. His primary

concern is most likely not to raise such reactions in his parents again. He may also gain a sense of ownership of his actions when they are appraised by others. The reactions from the parents serve heuristically to teach the boy about moral agency. This example indicates that the process of becoming morally responsible is precisely that: a process. Similarly, the process of moral disillusion is also scalar, one of the phases of which is becoming morally responsible. The boy need not be acquainted with evil to be held morally responsible. Moreover, despite being held morally responsible he may still retain some moral illusions.

The boy with the presents might at some stage, however, understand that not only is it wrong for him to do what he did because it spoils another child's birthday party, but that he has the capacity or ability to do such things and that they could be done to him. Further, he might eventually express indignation at those who act towards others as he did to the birthday girl. As he matures he may grasp that he can do more than open another's presents to destroy the happiness of another. In other words, he may understand his broader capacities to do wrong, and eventually, evil. This scalar process is one comprised of the shattering of different illusions. These shatterings will occur at different times for different people and it is best left to child psychologists to inform us when they usually occur. What I present here is an analysis of the kind of moral understanding that accompanies the various shatterings of moral illusions. The loss of moral innocence is a significant, but only one of several, disillusions. One needs to recognize one's capacity for wrongdoing to be held morally responsible. This recognition is one phase of moral disillusion. But I think that moral life is richer

19

than merely being an appropriate candidate for the reactive attitudes. One shatters a different moral illusion by recognizing one's capacity to bring about evil. Another phase of moral disillusion is the recognition that one can bring about or be complicit in the occurrence of evil. I shall return to the notion of evil in chapter 3.

The shattering of moral illusions need not be sudden or dramatic (although it makes for good literature when it is); it need only motivate the innocent's grasp of her own ability to do something wrong. French points out that oftentimes the innocent individual needs a push from adult members of the moral community who have already lost their innocence. Thus we might say that what is required is a teleological suspension of innocence in order to enter into the moral community.[12]

What is important to note at this juncture is the significance of the notion of moral illusions. As we have seen, French shows why the illusions of the innocent must be shattered to enter the moral community. I have already mentioned and will later show that the stable sense of morality that replaces the moral illusions of the innocent is also often riddled with illusions.[13] After realizing that one can be an agent of wrongdoing, one might think that following the categorical imperative or the principle of utility will allow one to avoid doing wrong, and that one can live a life of moral purity. These theories suppose that becoming a moral saint is possible as long as one follows the correct moral

---

[12] In chapter 6 I show how Kierkegaard's philosophy can be interpreted as a philosophy of shattering moral illusions.
[13] See chapters 3-5.

principle. For example, Kant argues that human nature is fundamentally good and one need only make choices under the maxim of the categorical imperative to lead a morally pure life. Evil is the corruption of our primarily good nature and thus can be avoided.[14]

> The human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it). The law rather imposes itself on him irresistibly, because of his moral predisposition; and if no other incentive were at work against it…he would be morally good.[15]

This idea that an agent can be perfectly good, or already is so by virtue of human nature, is an example of the kind of moral principle that replaces the moral illusions of the innocent. By endorsing these sorts of moral principles, one already understands that evil occurs and even that it can occur because of one's own actions. However, one still retains the idea that evil can be avoided, human nature is fundamentally good, and that evil is a corruption of that which is good. But I will later show that this idea is also a moral illusion.[16] Through the acquaintance with moral tragedy and unchosen evil one learns that the natural order is indifferent to moral categories, there is no fundamentally good (or evil) human nature, and evil cannot always be avoided. Recognizing these facts of moral life constitute shatterings of other moral illusions. But more of this later on. First, we must be clear on what we mean by "innocence" and its loss.

---

[14] Kant admits that moral purity is unlikely, however, it is still possible. His view might be seen as a kind of pessimistic optimism.

[15] Kant, *Religion within the Boundaries of Mere Reason and Other Writings*, trans. and ed. Allen Wood and George Di Giovanni (New York: Cambridge University Press, 1998),58.

[16] See chapters 4 and 5.

In his essay "Loss of Innocence and the Things that Remain"[17] Michael McKenna raises questions regarding French's theory of the loss of innocence. First, McKenna points out that the loss of innocence does not guarantee the acquisition of the knowledge of one's capacity for wrongdoing that is necessary for entry into the moral community. One can confront wrongdoing without subsequently gaining the knowledge that one is capable of wrongdoing. In such a case, the individual loses her illusory understanding of the moral universe but does not replace it with a more realistic one.

> The loss of innocence is not, at least not clearly, a guarantor of entry into the class of morally responsible agents. This is because what is lost need not be accompanied by the conceptual gains that French finds necessary for becoming a moral agent. A young person might lose the option of seriously using the illusion of innocent description, and yet not be able to supplant it with a stable understanding of good and evil, of morality.[18]

McKenna is correct that one's understanding can be shattered without a new one taking its place. We can imagine the case of a molested child who, by living through the trauma of the abuse, realizes that her previous conception of morality was naïve. Nevertheless she does not take the further step of establishing a more nuanced understanding of moral action, nor internalizes the wrong act in a non-traumatic manner such that she sees herself as a potential agent of wrongdoing.

McKenna's point is well-taken and raises the question of what it means for someone not to enter the moral community. As we have described it thus far,

---

[17] Michael McKenna, "The Loss of Innocence and the Things that Remain" (*APA Newsletter* 7, no 2 Spring 2008), 5-9
[18] Ibid., 7.

losing moral innocence is a necessary condition for entry into the moral community. What we mean when we say that an agent is a member of the moral community is that she can be held morally responsible for what she does and that she is an appropriate candidate for the reactive attitudes. By looking at Strawson's description of the moral community and why certain people are excluded from membership, we will gain insight into the different kinds of moral illusions that agents have. We will then also see why some agents might have their moral illusions shattered, but still not become members of the moral community.

Strawson notes the great importance we place on the attitudes and intentions of others toward us. Our personal feelings and reactions depend on our beliefs about these feelings and intentions of others. It matters a great deal to us whether another person's actions reflect attitudes toward us of goodwill, affection, or esteem on the one hand or contempt, indifference, or malevolence on the other. As he puts it:

> If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with a malevolent wish to injure me. But I shall generally feel in the second case a kind and degree of resentment that I shall not feel in the first. If someone's actions help me to some benefit I desire, then I am benefited in any case; but if he intended them so to benefit me because of his general goodwill towards me, I shall reasonably feel a gratitude which I should not feel at all if the benefit was an

23

incidental consequence, unintended or even regretted by him, of some plan of action with a different aim.[19]

The manner in which I react to another's actions, or the feeling I have toward them, depends on the attitude I attribute to the other agent. The range and intensity of our reactive attitudes is quite wide, from resentment to gratitude and in between, and varies due to how we understand the attitudes and intentions of those with whom we interact. To experience these attitudes is both natural and essential to human interaction.

Strawson points out that there might be occasions on which we suspend a reactive attitude. There are situations in which I feel resentment toward another. Then there are mitigating circumstances which might motivate suspending my resentment. Taking a clue from J.L. Austin's *A Plea for Excuses*, Strawson notes that when an agent proffers excuses such as "I didn't mean to," "There was no other way," "I was pushed," etc., the feeling of resentment is mollified or removed altogether. But even when an excuse mollifies resentment after an injury, it does not indicate that reactive attitudes are in any way inappropriate when applied to the agent generally. The agent is still a fully morally responsible agent. She just happens not to be responsible for this particular injury at this time. Therefore, I forswear my feeling of resentment toward her for this particular act, but not the application of the reactive attitudes more generally. This agent is still an appropriate object of demands of goodwill and the injury she caused is not inconsistent with this more general demand. He adds,

---

[19] P.F. Strawson, *Freedom and Resentment and Other Essays* (London: Methuen and Co., 1974), 3.

[t]he offering of such pleas by the agent and their acceptance by the sufferer is something in no way opposed to, or outside the context of, ordinary inter personal relationships and the manifestation of ordinary reactive attitudes. Since things go wrong and situations are complicated, it is an essential and integral element in the transactions which are the life of these relationships.[20]

Then there are those suspensions of the reactive attitudes that are due to the agent's general incapacity to live up to the demand of goodwill. Strawson includes children, schizophrenics, and individuals acting under hypnosis in this category. In these cases we suspend our reactive attitudes towards the agent at all times. We should not feel resentment towards these agents when they cause injury. We adopt an objective attitude toward such individuals treating them differently than others from whom we expect just actions. In other words, we do not hold them morally responsible for their actions. Therefore, being an appropriate target for the reactive attitudes indicates that the agent is a member of the moral community; that is, she can be held morally responsible for her actions.

Strawson distinguishes three kinds of reactive attitudes all of which are necessary for moral agency and moral responsibility. These attitudes are the ones that after losing moral innocence an agent adopts towards others, adopts towards oneself, and that others adopt towards the agent. First, we hold reactive attitudes towards others' intentions towards us. We expect others to have goodwill, or at least not active ill will, towards us. These are personal reactive attitudes. We also have expectations about others' wills towards others. When one person treats

---

[20] Ibid., 4.

another poorly without justification, then we feel not resentment, but indignation or moral disapproval. These are vicarious reactive attitudes. Just as there are personal reactive attitudes associated with the demands on others for oneself and vicarious reactive attitudes associated with the demands on others for others, so too are there self-reactive attitudes associated with demands on oneself for others. "And here we have to mention such phenomena as feeling bound or obliged (the 'sense of obligation'); feeling compunction; feeling guilty or remorseful or at least responsible; and the more complicated phenomenon of shame."[21] If being held morally responsible requires understanding oneself as capable of wrongdoing, and someone becomes acquainted with wrongdoing without gaining this understanding, then she cannot be held morally responsible for her actions. In this case, this individual would lack the participant reactive attitudes, especially the self-reactive attitudes. If she does not see that she can do wrong in the world, then she would not understand why others resent her for some of her actions. Thus this individual would not feel obliged to act appropriately toward others, and would not feel guilt and shame when she failed to do so. Or, if an individual experienced the personal and self-reflective, but not the vicarious reactive attitudes, she would be, as Strawson points out, a moral solipsist.

> One who manifested the personal reactive attitudes in a high
> degree but showed no inclination at all to their vicarious analogues
> would appear as an abnormal case of moral egocentricity, as a kind
> of moral solipsist. Let him be supposed fully to acknowledge the
> claims to regard that others had on him, to be susceptible of the

---

[21] Ibid., 9.

whole range of self-reactive attitudes. He would then see himself as unique both as one (the one) who had a general claim on human regard and as one (the one) on whom human beings in general had such a claim. This would be a kind of moral solipsism.[22]

This individual has not lost (for reasons biological, psychological, or otherwise) the illusion that what happens to others has nothing morally to do with her. Part of moral disillusion is to see not only that I am morally obligated to others and they to me, but that our moral responsibility extends beyond responsibility for our own actions to responsibility for seeing that the appropriate moral resentment of others is seconded by our own indignation. I shatter the illusion that the only actions of moral import are those I do or those done to me. But without this recognition, the individual retains illusions about moral behavior that exclude her entry into the moral community.

By distinguishing three different kinds of reactive attitudes, Strawson has helped delineate three different ways in which an individual might be excluded from the moral community. That is, failure to be an appropriate candidate for any of the three categories of reactive attitudes may be sufficient not to hold a person morally responsible for what they do. These three ways of failing could be due to a shattering of moral illusions without the illusions being replaced by a stable moral structure, or due to a failure to shatter moral illusions altogether. The individual might grasp that others have moral obligations to her, but not that she has any to others. Or, she might understand that other people make moral demands on others, but not on her. Or, she might think that the only morally

---

[22] Ibid., 9.

important actions are those that involve her. Whatever the deficiency in moral understanding, this schema also shows us that it is not always true that those individuals excluded from the moral community, and specifically of interest here the innocent, lack all moral understanding whatsoever. Rather, they lack *sufficient* moral understanding. Moreover, we can see that acquiring the reactive attitudes is not an all or nothing affair. An individual might slowly come to realize her obligation to others, their obligation to her, and the obligation of others to others. Perhaps, the agent first feels resentment that another has shown her ill will, then feels shame when she shows another ill will, and eventually feels indignation when another shows another ill will. The loss of innocence is a scalar process. Hence, there can be an incipient morality for the morally innocent.

McKenna explores the kind of morality that functions for innocents. One might incorrectly suppose that a moral innocent has no understanding of morality at all. Their innocence is simply a lack of awareness of moral action and appraisability, and they cannot parse actions or people into good or bad. But this is simply not true. Innocents are very aware of a moral order, however theirs is a simplistic one where "the superhero is always pure, strong, physically superior, handsome or beautiful, and the villain always tainted, physically enfeebled, ugly, and so on."[23] The moral order in which innocents operate is one where the categories of good and evil are clear to discern. More disillusioned members of the moral community have grasped that actions are not always so morally black and white. In fact, one might come to the conclusion that there are no good and

---

[23] McKenna, "Loss of Innocence and the Things that Remain", 7.

evil people as such, but only good and evil actions. This realization could be based on the observation that even good people can do evil things. But this latter notion is lacking in the moral understanding of the innocent. It follows then that one way to describe innocence is that it is a certain kind of ignorance. But ignorance is not equivalent to innocence.

4. Ignorance

When we think of who in the actual world is morally innocent we tend to think of children. Children are of course morally innocent. However, being morally innocent does not belong exclusively to children. Adults could also be morally innocent. Although it is possible for adults to be morally innocent, it is more likely that an adult is ignorant about the moral nature of a particular kind of act. Thus, we need to be careful not to conflate innocence with ignorance. Recall that moral innocence is a status. Moral innocence is a *general* status that an agent has, whereas ignorance can be both general *and* particular.

Norvin Richards[24] distinguishes innocence into two kinds: first, an agent could do something wrong but not know the act fits under a category we ordinarily consider wrong, and second, one could be incapable of conceiving of a certain kind of (wrong) act as possible. Richards thinks that both kinds of innocence appear in children and adults. I agree that it is possible for both kinds of innocence to appear in children and adults, but I think that when it appears in adults it is usually ignorance, not innocence. For example, Richards asks us to imagine a little girl who likes telling stories when she is asked questions. All she

---

[24] Norvin Richards, "Innocence," *American Philosophical Quarterly* 31, no. 2 (Apr., 1994): 157-167.

knows is that it is fun to imagine events and tell others that they actually took place. She does not know that she is lying; that is, even if she knows what lying is, she does not realize that she is lying by telling her stories. She does not realize that her act fits into a category of acts one ordinarily considers wrong. In this case she is doing something wrong, but does not recognize it as wrong.[25] She does not place her action under the category of lying. To illustrate the second kind of innocence, Richards asks us to imagine a boy who enjoys looking at pictures of other children. He enjoys it, not for any sexual reason, but because he finds it fascinating that other children can appear in so many different ways. Although we know that some people do look at pictures of children as sexual objects, the boy does not. He does not know that such an act is even possible. He completely lacks the conception of this kind of act. Here we have two examples of children who are generally innocent, but who manifest their innocence in different ways. The girl shows her innocence in this case by doing something wrong, but not knowing that it fits a certain kind of wrongdoing, and the boy shows his by being incapable of even conceiving of a certain kind of wrongdoing. Both kinds of innocence are due to their being generally innocent.

It is possible, though unlikely, that these different kinds of innocence, viz. not knowing that what ones does is wrong and not conceiving of a kind of act as possible, can also belong to adults. However in the case of adults, I believe that usually one is not innocent, but ignorant. Richards argues that adults in the following examples are innocent because they do not comprehend that a particular

---

[25] Given her innocence, she is not morally blameworthy for the wrong she does.

action is morally wrong. However, if moral innocence is a general status, one cannot be morally innocent of a particular act; rather one can be (culpably or inculpably) ignorant that a particular act is wrong or has a moral component. For example, a rapacious inside trader sees nothing wrong with what he considers brilliant steps in a competitive game. He is doing something wrong, but he does not, for whatever reason, recognize it as wrong. Perhaps he has rationalized the act as simply the only possible way to win in the game of capitalism, or perhaps he thinks that if it is a "victimless crime," then an act cannot be classified as wrongdoing. Whatever the reason, he does not understand his wrongdoing as wrong. In this way he is similar to the girl who tells stories. But he differs from the girl in that he might be familiar with the idea that what he is doing could be wrong, but he rejects it. His familiarity with moral notions, and his ability to understand that he is an appropriate candidate for the reactive attitudes, entail that he does not have the general status of "morally innocent." For the girl the idea that telling stories is wrong is something entirely new because she is morally innocent. Although Richards calls the inside trader innocent, he is really ignorant.[26]

Consider next a woman working as bookkeeper in a company. She is new to the company and notices a discrepancy while pouring over the company's records. She recalculates the numbers multiple times, but cannot get them to come out right. She brings her report to the boss who is an alumnus from the college she attended and a pillar in the local Lutheran church. He says he will take care of it.

---

[26] Below, I consider culpable ignorance in more detail.

Later talking to her friends on the phone she is astonished that they think the boss is embezzling money. She cannot understand how that would be possible. He went to the same college, he is an upstanding member of the local Lutheran church, and such people do not do such things. She cannot even conceive of the possibility, just as the little boy could not conceive of the possibility of looking at children as sexual objects. She differs from the boy in that the boy lacks understanding of the whole dimension of sexuality (or sexual perversion) because he is morally innocent; she lacks more specific understanding that "embezzler" can belong to the same list as "pillar of the local Lutheran church," not because she is innocent, but because she is ignorant. Because she is a member of the moral community she has already lost her moral innocence, however she retains certain moral illusions. These cases of the two adults illustrate that losing moral innocence does not entail that the agent cannot be ignorant of moral truths. In fact, it is precisely because members of the moral community retain some illusions that there are different phases of moral disillusion. Each advanced stage of moral disillusion involves a shattering of ignorant illusions.

Although the adults in the above examples are ignorant without being innocent, the children's innocence is characterized by a specific kind of ignorance. As we have discussed so far, what the innocent lack is knowledge by acquaintance with wrongdoing; that is, they are ignorant that their actions have upshot in the world and that they are appropriate candidates of the reactive attitudes. We have also discussed that this lack of acquaintance explains an absent cognitive ability either to conceive of their act as wrong or even to conceive of the

possibility of the act at all. This ignorance is experiential as well as cognitive. The innocent have not yet had the requisite experiences nor adopted the requisite concepts or descriptive vocabulary to comprehend the nature of wrongdoing or of evil. More importantly, an innocent has not yet comprehended that she is capable of wrongdoing or evil, and therefore lacks the ability to participate in the moral community. Her innocence is intimately tied up with her ignorance. But the ignorance of the innocent is not just ignorance of the concepts of good and bad, as Aristotle argues. Rather, it is ignorance that the agent herself is capable of wrongdoing or evil. It is an *inability* to *see that* one's acts have moral significance. It follows that the loss of moral innocence is not just a gain in conceptual knowledge, but the acquisition of an ability—an ability to participate in the moral community. I present the structure of this ability in the next two sections and then examine it in detail in the next chapter. First, I contrast the ignorance of the morally innocent with Aristotle's definition of ignorance. This contrast will help motivate the interpretation of moral understanding as an ability.

In Book III of *Nicomachean Ethics* Aristotle distinguishes between two kinds of ignorance. It is important that Aristotle define ignorance and distinguish it into its different kinds because "every wicked man is ignorant of what he ought to do and what he ought to abstain from."[27] Therefore, understanding the nature of virtue and becoming virtuous includes an understanding of ignorance. According

---

[27] Aristotle, "Nicomachean Ethics" in *The Complete Works of Aristotle Volume II* ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), 1110b28-29.

to Aristotle an action can be done *in* ignorance or *caused by* ignorance.[28] An agent who does an action *in* ignorance does it without knowledge of the universal. One does not have general knowledge of good and bad, and therefore of what one ought to do. One experiences no pain or regret due to one's ignorance of this general knowledge. This kind of ignorance is inexcusable because one should have this knowledge of the universal. In Strawson's terminology, we would not suspend the reactive attitudes towards this individual.[29]

One could also be ignorant of particular circumstances. The agent knows what is good and bad, but is kept ignorant of the badness of her action until after its occurrence. Once the individual realizes what she has done, she experiences pain and regret. This kind of ignorance mitigates responsibility as long as the agent was not responsible for her state of ignorance by a prior act. Unless someone's ignorance about what she should do is itself culpable, then she is not morally blameworthy for her act. Given this kind of ignorance, we suspend the reactive attitudes toward the agent for this particular wrong act.

This distinction of kinds of ignorance shows that it is morally permissible to be ignorant of facts, but not about moral principles. Aristotle believes that the first kind of ignorance is sometimes exculpatory and the latter kind never exculpatory because actions based on the first kind of ignorance are involuntary and actions based on the second kind are not involuntary.

---

[28] I use these terms because they better correspond to our ordinary English usage of terms. But there are various translations of the Greek in different versions of the text.
[29] The children in the example above are of course not morally responsible for their ignorance. This kind of culpable ignorance applies only to members of the moral community.

Everything that is done by reason of ignorance is *non*-voluntary; it is only what produces pain and regret that is *in*voluntary. For the man who has done something owing to ignorance, and feels not the least vexation at his action, has not acted voluntarily, since he did not know what he was doing, or yet involuntarily, since he is not pained.[30]

According to Aristotle the two kinds of ignorance line up with involuntary or non-voluntary action. Although involuntary action is sometimes not morally blameworthy, non-voluntary action is blameworthy. Although the agent did not know what she was doing, she should have known and she is further morally blameworthy for not knowing. Given Aristotle's schema, the innocent person would be the one who is inculpably ignorant of the universal. However, I think that defining moral innocence simply as ignorance of the universal is incomplete. Exposing what is missing in Aristotle will help to illuminate what is significant about the present account.

It is not entirely correct that the innocent are ignorant of the universal. As I discussed above, the innocent often recognize a naïve moral system where good and bad are clearly defined. But even if Aristotle were correct that the innocent person is ignorant of general knowledge of good and bad, she is not culpable for her ignorance. Ignorance of the universal is blameworthy for those capable of grasping it, but the ignorance that characterizes the innocent is not. The innocent are not the kind of individuals who should have knowledge of the universal. Even so, losing the kind of ignorance that Aristotle thinks constitutes innocence is not sufficient for the loss of moral innocence.

---

[30] Aristotle, *Nicomachean Ethics*, 1110b17-22.

For the loss of moral innocence, it is not sufficient merely to acquire general knowledge of good and bad. As French notes:

> Knowledge of good or what one morally ought to do, however, is probably not as important as knowledge of evil or rather, of one's capacity to do and be done evil…Experiencing evil in the loss of innocence is grasping for the first time the possibility that things might have gone differently, and so seeing what would have been good in the situation, and so seeing oneself as capable of evil.[31]

One must not only acquire general knowledge of good and bad, one must also see oneself as the potential source of wrongdoing or evil. The individual must make the connection between the occurrence of evil, and her ability to produce it. In other words, one must locate evil in terms of one's own agency. As Fischer and Ravizza put it, the agent must be aware that her actions have some upshot in the world. I added to their account by saying that she must also be aware that her actions could have a moral upshot in the world. What we will continue to see at each stage of moral disillusion is the significance of comprehending, not just what is good and bad or right and wrong in the form of moral principles, but that the good and bad or right and wrong action comes from oneself.[32] In the next chapter I examine this kind of understanding in more detail and argue that it is not merely conceptual because it is a kind of ability—an ability to participate in the moral community and to exercise the moral imagination. In the following section, I

---

[31] French, *Responsibility Matters*, 39.

[32] Thomas Nagel makes a similar point when he argues that the problem with theories like utilitarianism is that they ask the moral agent to reflect only upon good or bad states of affairs. They miss a crucial feature of any act, viz. that you are the one performing the act. One ought to consider not only which states of affairs may obtain, but what one must do such that specific states of affairs obtain. Nagel argues for this reason that there are agent-relative deontological constraints on one's actions. See Nagel, *The View From Nowhere* (New York: Oxford University Press, 1986), 175-180.

begin to lay the foundation for the argument in chapter 2 by arguing that the condition necessary for being held morally responsible, namely guidance control, is an ability.

5. The Ability to Guide

In section 2 of this chapter, I explained that to be held morally responsible for one's actions the agent must be moderate reasons-responsive. Fischer and Ravizza argue persuasively that the agent's mechanism must be strong reasons-receptive and at least weak reasons-reactive. What it means to be moderate reasons-responsive is to have "guidance control." Inspired by Frankfurt-type cases, Fischer and Ravizza distinguish two different kinds of control, only one of which is necessary for moral responsibility. Regulative control is the kind of control where the agent has alternative possibilities. In such situations, the agent can do one action or another. Frankfurt shows that there may be situations where an agent does not have alternative possibilities open to her and yet is still held morally responsible for what she does. For example, Jones shoots Smith while Black stands waiting to interfere and ensure that Jones decides to shoot Smith should Jones decide not to shoot Smith. Because Black is a counterfactual intervener, Jones lacks an alternate possibility to deciding to shoot Smith. Yet, he shoots Smith without any coercion or intervention by Black, and we hold him morally responsible for it. The example illustrates that an agent may be held morally responsible even if she lacks any alternative possibilities. Because Jones goes through with the action on his own and without Black actually intervening, he is morally responsible for what he does. The intervention

occurs in the alternate scenario, not the actual one. Therefore, having open alternative possibilities, or regulative control, is not necessary for ascriptions of moral responsibility.

However, it does not follow from this discovery that no control at all is required for moral responsibility. Fischer and Ravizza explain that although the morally responsible agent in Frankfurt-style cases does not have regulative control, that is, does not have open alternative possibilities, she does have guidance control. "Guidance control of an action involves an agent's freely performing that action. Regulative control involves a *dual* power: for example, the power freely to do some act *A*, and the power freely to do something else instead."[33] In a situation where an agent has guidance control, she freely does whatever she does although she could not have done otherwise. It does not matter whether she could have freely done something in an alternative sequence.[34] Fischer and Ravizza ask us to imagine a case where the driver of a car, Sally, turns the car to the right. However, this car is a driver instruction automobile with dual controls for the driver instructor. If Sally had not turned or guided the car to the right, the instructor was ready to engage his controls to ensure the car turned to the right. Sally does not have regulative control because she could not have done otherwise than turn the car to the right. However, she does exercise guidance

[33] John Martin Fischer and Mark Ravizza, *Responsibility and Control*, 31.
[34] There has been extensive discussion regarding the kind of freedom in Frankfurt-style cases and whether these "flickers of freedom" are robust enough to ground our ascriptions of moral responsibility. See John Martin Fischer, *The Metaphysics of Free Will: An Essay on Control*, Aristotelian Society Monograph Series, vol. 14 (Cambridge Mass.: Blackwell Publishers, 1994), 131-159, and also David Widerker and Michael McKenna, ed. *Moral Responsibility Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, (Burlington, VT: Ashgate, 2003).

control. She freely turned the car to the right. This unusual case points out

something significant regarding action and control. One can have a certain sort of

control without having the sort of control that involves alternative possibilities.

What both the Frankfurt-type cases and Fischer and Ravizza point out is

that the traditional association of moral responsibility with control is correct. But

the relevant sort of control does not involve alternative possibilities. The sort of

control associated with moral responsibility is guidance control. But what does

guidance control mean specifically? In what way does Sally freely guide the car

such that she can be considered morally responsible for turning the car to the

right? Fischer and Ravizza have already shown that she need not have any

alternative possibilities open to her. Rather, what is required is that her actual

action-issuing mechanism be reasons-responsive. Although the *agent* could not

have done otherwise, her actual action-issuing *mechanism* was reasons-

responsive. She could receive sufficient reasons to do otherwise, and there is

some possible world where she acts on one of those reasons. Furthermore, the

mechanism that operates in Sally's brain is her own. Thus, Fischer and Ravizza

define guidance control as action performed by an agent's reasons-responsive

mechanism that is also her own. The agent is able to receive and possibly react to

sufficient reasons to do otherwise.

But guidance control need not be entirely cognitive. The agent's ability to

receive and possibly react to reasons and then to guide one's actions in a certain

direction can have a volitional component as well. Reasons can move an

individual to act given what the agent is willing to bring herself to do. That is, one

39

might receive a reason based on purely cognitive grounds, but she may only react to those reasons, or react to those reasons in a possible world, volitionally. One can decide based on reasons that one course of action is best. But one might still be incapable of bringing oneself to act on those reasons.[35] In Harry Frankfurt's words, "one runs up against the limits of his will."[36]

For Frankfurt, an agent's will is the desire by which she is motivated in some action or the desire by which she will be motivated when or if she acts. "An agent's will, then, is identical with one or more of his first-order desires."[37] Thus, the notion of will as Frankfurt uses it is the notion of an effective desire. Not only does *A* have certain desires, but she also desires to desire in a certain way and wants those desires to move her effectively to act. In other words, *A* wants her will to be a certain way. When *A* wants a certain desire to be her will, then she has a second-order volition. Second-order volitions are often motivated by what one cares about. Those desires that a person wants to have and wants to be effective are those which comprise her will and her identity, and which move her act. We may not always choose which desires move us to act, but we reflect upon the desires we have and have a second-order volition that they be effective in our action. Therefore, guiding ourselves is not simply a cognitive endeavor. It is also a volitional ability.

---

[35] In chapter 4 I discuss in more detail how an agent is guided by what he cares about. It is sufficient at this point of the argument to show that guidance control may have a non-cognitive component.

[36] Frankfurt, 182.

[37] Ibid., 14.

Both Frankfurt and Fischer and Ravizza allow us to see that being a morally responsible agent demands the ability to guide one's actions. Being moderate reasons-responsive and having guidance control require a certain kind of understanding. In order to be receptive to the kinds of reasons that are germane to moral acts, one must *see that* certain actions have moral import. Moreover, in order to react to a sufficient reason to act one must be able to guide one's actions in a particular direction. Thus, the loss of moral innocence is not merely an acquisition of moral concepts; it is the acquisition of an ability to *see that* one's actions have moral significance and to be able to guide one's actions accordingly. In this section, I have only begun to indicate that the moral understanding that accompanies the loss of moral innocence is an ability. In the next chapter I examine this ability in more detail.

6. Conclusion

In this chapter, I analyzed the shattering of moral illusions that characterizes the loss of moral innocence. In order to be held morally responsible one must recognize one's capacity to do wrong and be an appropriate candidate of the reactive attitudes. One loses moral innocence, but gains the ability to participate in the moral community. However, the moral disillusion that characterizes entry into the moral community does not rid us of all of our moral illusions. I will continue to examine the more complex phases of moral disillusion in the subsequent chapters. I distinguished innocence from ignorance by showing that although innocence includes ignorance of moral concepts, ignorance does not necessarily include innocence. That is, one can be culpably ignorant. Finally, I

41

have begun to show that what one gains from losing innocence is an understanding seen as an ability to participate in the moral conversation and the moral community. If being a member of the moral community requires that one has guidance control, then not being eligible for the moral community means that one lacks this ability. Therefore, moral innocence is characterized not only by a lack of moral concepts, but an inability to practice guidance control. One cannot practice guidance control if one is not moderate reasons-responsive. Becoming moderate reasons-responsive, i.e. being able to receive and possibly react to sufficient moral reasons to do otherwise, requires an ability to discern sufficient moral reasons. This discerning ability is the ability to understand that one's action has moral significance. Thus, the understanding that accompanies the loss of moral innocence is not merely a conceptual comprehension; rather, it is understanding in the sense of an ability to participate in the moral conversation. In chapter 2 I examine this ability in more detail.

CHAPTER 3

UNDERSTANDING, SEEING THAT, AND THE MORAL IMAGINATION

1. Introduction

In the previous chapter I analyzed the loss of moral innocence. I explained how the loss entails a gain; one gains enough moral understanding to be held morally responsible for one's actions. At the end of the chapter I explained that impersonal conceptual knowledge of good and evil is not sufficient for the loss of innocence. One must also see oneself as a potential source of wrongdoing. That is, the individual must understand wrongdoing in terms of her own agency. In this chapter I shall say more about this sort of understanding that accompanies this phase of moral disillusion. Based on Wittgenstein's distinction between knowing and understanding, I argue that knowing good and evil does not entail understanding good and evil. However, both knowing and understanding involve a transition from "seeing" to "seeing that." I argue that we can "see that" actions have moral significance in different ways. One of these ways involves imagining different ethical consequences, or what another person might do in my situation, or what another way of life is like. The moral imagination allows the agent both to explore future possibilities as well as to play out what one should have done in the past. It is more than a recognition that one's actions have moral upshot in the world and that one is an appropriate candidate for the reactive attitudes. By imaging hypothetical ethical scenarios, one has the ability to raise questions about moral principles.

Before I continue, I need to clarify a possible confusion; one might conflate moral disillusion with moral goodness, but it is important to distinguish losing moral illusions from being morally good. Losing moral illusions does not entail becoming morally good, however being morally good requires the loss of certain moral illusions. We can easily see this by noticing that one could use their knowledge of good and evil to do evil deeds, but one cannot do good deeds until one recognizes her capacity to do good and evil.

Moral disillusions are necessary but not sufficient for becoming a morally good person. Losing moral innocence is required for participation in the moral community and being held morally responsible for what one does. But clearly, one need not be morally good to be a member of the moral community. One might think that after losing moral innocence one replaces the shattered illusions with a stable moral system like Aristotle's universals, or the categorical imperative, or the greatest happiness principle, or any other moral directive. This may indeed occur. However, such knowledge characterizes (perhaps) the morally *good* person. What is distinctive about losing moral illusions, although it also allows for the possibility of grasping moral directives, is the ability to raise ethical questions, not (always) to answer them. The morally good person can recite and follow moral directives. The person who has shattered moral illusions can raise questions as to the inadequacy of those moral directives. As we examine the different phases of moral disillusion we will see that this ability develops such that the agent can raise ever more difficult questions, especially those that cannot

44

be easily answered.[38] Thus, the inability to answer difficult ethical questions may not be an indication of an inability or a lack of understanding (although it could be), but rather a distinguishing feature of an understanding that grasps the complexity of moral life.[39] In what follows, I unpack the kind of understanding that accompanies the various phases of moral disillusion and show that raising ethical questions is one of its defining characteristics.

2. Adam and Eve

There is of course a famous myth in our culture concerning the loss of moral innocence. It is the story of Adam and Eve. God commands Adam and Eve not to eat from (or even touch!) the tree of the knowledge of good and evil. But God does not merely tell the two not to eat from the tree. He includes a threat for noncompliance as well. He tells them that they will die if they eat from the tree. "And the Lord God commanded the man, 'You may freely eat of every tree of the garden; but of the tree of the knowledge of good and evil you shall not eat, for in the day that you eat of it you shall die.'"[40] Now immediately the reader might be confused. Knowing the end of the story one might ask why God would tell them they will die, when He knows that they will not die. Also, why did God place this tree in the middle of the Garden of Eden (along with the tree of life) if he did not want them to eat from the tree? Perhaps it is because God did not want humans to have knowledge of evil. But if God did not want them to eat from the tree so that they would never know evil, did he also not want them to know good? After all it

---

[38] Socrates is the paradigmatic gadfly challenging naïve ethical notions.
[39] The specific complexities I will later address are moral tragedy, inescapable moral wrongdoing, and unchosen evil.
[40] Gen. 2:16-17 (New Revised Standard Version)

is the tree of the knowledge of good and evil, not just evil. Quite a lot of ink has been spilled on biblical exegesis to interpret this story as a myth that explains free will, the origin of sin, etc., but outside of a religious context God's actions seem quite peculiar. In any event, in the story the serpent tells Eve that she and Adam will not die, and God only forbade eating from the tree so that humans would not become like him. Later, God admits as much.

> Then the Lord God said, 'See, the man has become like one of us, knowing good and evil; and now, he might reach out his hand and take also from the tree of life, and eat, and live forever'—therefore the Lord God sent him forth from the garden of Eden…[41]

After God discovers that they ate the fruit, He casts Adam and Eve out of the Garden of Eden. He does so not only to punish them for disobedience, but also because they have already gained knowledge of good and evil, and if they added immortality to the mix by eating from the tree of life, they would be gods like him.[42]

The narrator of the story tells us that after eating the fruit "the eyes of both were opened."[43] Where they had previously *seen* each other naked, they now *saw that* they were naked. In other words, they acquired a certain kind of knowledge that allowed them to see the significance of a certain state of affairs. Losing moral innocence can be described as an acquisition of knowledge such that one sees *that* where before one could only see. Adam and Eve now have knowledge of good and evil. In other words they know the difference between good and evil.

---

[41] Gen. 3:22-23.
[42] I am not sure why this would have been such a bad state of affairs.
[43] Gen. 3:7.

And yet, it seems a stretch to claim that they *understand* the difference between good and evil. Their "knowledge" in the story seems to indicate "awareness of" or "becoming acquainted with." But awareness or becoming acquainted are not understanding. I am aware, i.e. I know, that when I spill coffee on my desk, the liquid pools on top of the hard surface and does not drip through it. But unless I have a conception of the kinetic energy of molecules and how this energy differs between solids and liquids, I do not understand this state of affairs. Similarly, Adam and Eve know good and evil, but they do not yet understand good and evil. Nevertheless, it would not be infelicitous to utter "I see that…" in contexts of understanding as well. This indicates that we can utter "I see that…" in contexts of both knowing and understanding, even though there is also sometimes a distinction between knowing and understanding. What is the difference between knowing and understanding good and evil? The difference, as we will come to see in this chapter, is one of moral disillusion.

3. Wittgenstein on Knowing and Understanding

In *Philosophical Investigations*[44] (hereafter *PI* [45]). Wittgenstein puzzles over this difference between knowing and understanding. There are some passages in *PI*, where it seems as though Wittgenstein distinguishes between knowing and understanding, and there are other passages where it seems he uses them interchangeably. Similarly, in ordinary language we sometimes distinguish between them and we sometimes use them interchangeably. But where does the

---

[44] Ludwig Wittgenstein, *Philosophical Investigations*, trans. G.E.M. Anscombe (Upper Saddle River, NJ: Prentice Hall, 1958.).
[45] I refer always to the passage number, not the page number.

distinction lie and what does the distinction tell us about moral disillusion? I will present Wittgenstein's description of "to understand" and "to know" first to establish that they are both abilities rather than mental states. Understanding them both as abilities is crucial to arriving at what I mean by moral disillusion. Then I ask whether he distinguishes between the two abilities. I answer this question in the affirmative and then show how we can get clear on the distinction by looking at an analogous distinction between reading and being literate. What we learn from Wittgenstein is that although there is indeed a distinction between knowing and understanding, neither one is a mental state. Rather, they are each an ability. In a moral context, they are each an ability to interact, or "go on," in the moral conversation. I will show that although each is an ability to go on, there is a difference in what we can go on to do. Knowledge of good and evil gives us the ability to experience, and be an appropriate candidate for, the reactive attitudes. Understanding good and evil gives the ability to raise nuanced ethical questions.

 First, it is important to be clear about what Wittgenstein says concerning "understanding" and "knowing." In presenting his general approach to understanding, I will initially gloss over differences between "understands" and "knows" in order to establish that they are abilities rather than mental states. After we are generally clear concerning what he wants to say about "understanding" and "knowing" as abilities, I will return to passages where he indicates how they might be distinct.

 Wittgenstein discusses with an interlocutor what it means to know or to understand. The discussion is couched in a discussion about applying a rule. The

interlocutor argues that one's ability to go on is the application of the understanding but not the understanding itself. By analyzing how one applies a rule and solves a problem Wittgenstein and the interlocutor approach the question of how one understands. The interlocutor represents the approach standard in the history of philosophy. He wants to define the essence of the thing. Wittgenstein wants to move away from looking for essences to looking at how words are used.

> When philosophers use a word—"knowledge", "being", "object", "I", "proposition", "name"—and try to grasp the *essence* of the thing, one must ask oneself: is the word ever actually used in this way in the language-game which is its original home?—What we do is bring words back from their metaphysical to their everyday use.[46]

The interlocutor, as the representative of the view coming out of the history of philosophy, is that the understanding is a mental state that is the source of one's ability to go on and use rules and language correctly. That is its essence. Under this theory, when a person hears and understands a word a certain thought appears in the mind. This thought then logically compels a certain application of a rule, which in turn allows the person to go on. Hence, the interlocutor thinks, there is a strong connection between what occurs in the mind when someone understands a word and how that person makes use of this understanding. The discussion moves to an example of a teacher teaching a student to complete a numerical series. The question is how the teacher knows that the student has *understood* the series. Again, the interlocutor wants to defend the position of traditional philosophy, which argues that the essence of the understanding is that it is a mental state.

---

[46] Ibid., 116.

Perhaps you will say here: to have got the system (or, again, to understand it) can't consist in continuing the series up to *this* or *that* number: *that* is only applying one's understanding. The understanding itself is a state which is the source of the correct use.[47]

The interlocutor concludes that the understanding is a mental state lying behind the correct use of a rule. How the student correctly and repeatedly completes the numerical series is by applying to the series the rule generated or grasped by the understanding. Wittgenstein resists this conclusion. He points out that we do not use the word "understands" to refer to the process of a mental state applying a rule. Rather, we say the student understands when he completes the series; that is, when he is able to go on. Understanding simply is the ability to go on.

Consistent with the main goal of *PI*, Wittgenstein tells us that we need to look to the grammar of "understand" in order to know what we mean when we use the term.

We must do away with all *explanation*, and description alone must take its place. And this description gets its light, that is to say its purpose, from the philosophical problems. These are, of course, not empirical problems; they are solved, rather, by looking into the workings of our language, and that in such a way as to make us recognize those workings: *in despite of* an urge to misunderstand them. The problems are solved, not by giving new information, but by arranging what we have always known. Philosophy is a battle against the bewitchment of our intelligence by means of language.[48]

---

[47] Ibid., 146.
[48] Ibid., 109.

The grammatical investigation will show that our concept of understanding functions very differently from the philosophical picture painted by the history of philosophy.[49] Wittgenstein's analysis will then shed light upon what we mean by an agent acquiring moral understanding and how this understanding differs from moral knowledge.

Wittgenstein objects to the idea that something in the mind compels a person to use a word in a particular way. He does this by asking us to imagine a cube. It is perfectly possible that on being told to imagine a cube, one person pictures a cube and another person a triangle. There is nothing in the drawing of a cube that forces a particular use on us. In terms of the example with the student and the numerical series, Wittgenstein freely admits that it may be a criterion of the student understanding the series that the formula comes to the student's mind. But if this is so, then it is only because the formula is used in a certain way and not because it is the source of the understanding. As with the example with the cube, there are circumstances where the formula comes to the student's mind but we would not say that she understands the series. Thus, Wittgenstein quickly rejects the notion that there is some relation of logical compulsion between the understanding and the use of a word or application of a rule.

It is at this point in his discussion with the interlocutor, that Wittgenstein turns to look more closely at the grammar of understanding. Again, his main goal is to resist the notion that the understanding is a mental state. It is only by failing

---

[49] I am not using "concept" with any specific technical or philosophical usage in mind. Quite the opposite. I am using the term in an ordinary way, and therefore am not entering into the discussion of Wittgenstein's position on concepts understood in the philosophical sense.

to attend to grammatical distinctions among our use of words that we end up

positing misleading philosophical theories which obfuscate rather than elucidate.

This grammatical investigation will also allow us to understand "moral

understanding" as an ability to go on, rather than as a mental state that applies

moral principles to specific situations. In terms of the loss of innocence, the

ability to go on is the ability to participate as a moral agent in the moral

community. We can understand this participation as an ability to go on in the

moral conversation.

   One grammatical distinction between mental states and understanding is

exposed by noticing that when we speak of mental states temporal concepts like

duration and continuity make sense. However, such concepts do not make sense

when we try to apply them to understanding.

> "Understanding a word": a state. But a *mental* state?—Depression,
> excitement, pain, are all called mental states. Carry out a
> grammatical investigation as follows: we say
> "He was depressed the whole day".
> "He was in great excitement the whole day".
> "He has been in continuous pain since yesterday".—
> We also say "Since yesterday I have understood this word":
> "Continuously", though?—To be sure, one can speak of an
> interruption in understanding. But in what cases? Compare:
> "When did your pains get less?" and "When did you stop
> understanding that word?"[50]

When we try to apply the same notions of temporality and continuance to

understanding, we end up with strange questions like at the end of the passage.

[50] Ibid., 151 footnote (*a*)

When a theory results in nonsense, then the theory itself might be nonsense! The above passage shows that even if understanding is some kind of a state, it is grammatically distinct from mental states. Wittgenstein wants to show that the concept of understanding is linked not to mental states, but instead to ability. The example of the student and the numerical series helps show that when we use "understand," we do not include any notion of some internal mechanism distinct from what the understanding does. What it does is allow the person to go on. But what does it mean to go on?

Going on is a practice. Practices are steeped in a person's (persons') form of life. The context within which "understanding" gets its sense (i.e. its use) is the form of life revealed in the way the people in question speak and act. This form of life includes their past history as well as their future ways of acting. Thus, to understand "understanding" we have to look at practices and performances of understanding.

> If there has to be anything "behind the utterance of the formula" it is *particular circumstances*, which justify me in saying I can go on – when the formula occurs to me.[51]

This passage indicates that there are various contexts in which we would say that someone understands. There is no single definition of "understanding" abstracted from the particular contexts and uses of "I understand" or "she understands." The example of the student with the numerical series shows us that there are a number of situations where the student might have the formula, but we would still not conclude that she understands. Wittgenstein acknowledges that the meaning of

---

[51] Ibid., 154.

"understand" is quite complicated. However, he also shows that the standard philosophical picture of the understanding further complicates matters. The grammatical investigation of "understanding" helps to bring its meaning to light even if it remains somewhat in the dark. What we can know is that the understanding is not a mental state. We have to look to the particular circumstances in which we use the words "Now I understand" or "Now I can go on" to elucidate what we mean by "understand." Thus, my use of these words connects with a form of life in which I have been acculturated and that reveals itself through past performances. Words have their significance only in this form of life. Consequently, "understanding" is not something that can be understood by looking at an isolated event. Understanding requires a context (i.e. form of life) as a backdrop to have any significance and communicative force. A form of life is a necessary condition for successful understanding. When we consider the moral community, or even better the moral conversation, as a form of life, we can think of "knowing" and "understanding" as each an ability; the former to experience the reactive attitudes, and the latter the ability to raise ethical questions about one's own and others' actions. These are two different ways that one can go on. As we will see, the first relates to knowing and the second to understanding.

Now that we have established that both knowing and understanding are abilities, I turn to see how they differ. This difference will explain different phases of moral disillusion. In the following passage we are told more explicitly that understanding is the ability to go on.

54

> The grammar of the word "knows" is evidently closely related to that of "can", "is able to". But also closely related to that of "understands". ('Mastery' of a technique,)[52]

There are two important parts of this quotation to notice. The first is that he is relating "to know" and "to understand" to "can" or "is able to."[53] This relation shows how knowing/understanding is an ability: the ability to go on. The second important part of this quotation is that Wittgenstein here explicitly distinguishes between knowing and understanding. He states that "knows" is closely related to "can" and "is able to" *and is also* closely related to "understands." Thus as far as this passage indicates, he does not consider "knows" and "understands" to be indistinguishable. As we will see the distinction between knowing and understanding in a moral sense is a distinction between two different kinds of ability and two kinds of moral disillusion.

Before I go on to show that Wittgenstein sometimes distinguishes between "know" and "understand," I want to look at how we use the these two terms in ordinary language. Because Wittgenstein is an ordinary language philosopher it is important that his explanation of these two words matches our use of them in ordinary language.

Clearly there are some important distinctions between "to know" and "to understand" in ordinary language. Above we said that although Adam and Eve know good and evil they do not understand good and evil. Similarly, I can say "I

---

[52] Ibid., 150.
[53] The original German is "wissen", "können" or "imstande sein", and "verstehen". These words are translated as "know", "can", and "understand" respectively. There is no significant change in meaning that results from the translation. Therefore, although we are dealing with the English terms in this paper, we can be confident that we are being true to Wittgenstein's message in these passages.

know you" in the sense of being familiar with who you are, but nevertheless say "I don't understand you" in the sense of not comprehending your motivations or a statement you have uttered. In both cases understanding is more robust than knowing.

For a more detailed example let's take the Czech sentence "Moje kočka je černá."[54] Now even if I do not speak Czech I can still state that I know what this sentence says. It says "Moje kočka je černá." However, if I do not speak Czech, then I certainly cannot truthfully assert that I understand what this sentence says. This distinction seems to be correct on the surface. However, Wittgenstein might say that even if I do not speak Czech I might understand the sentence depending on the context. Imagine the following situation: I am in an apartment in Prague and a black cat walks into the room. A Czech speaker says the above sentence to me. It is reasonable to suppose that given the right conditions, I might understand what she is trying to communicate. But as Wittgenstein also points out in other places in *PI*, there is no necessary connection between the words and the object. The circumstances of the situation would have to be such that I understand how the sentence is being used to communicate that someone's cat is black.

This example is helpful in showing us that Wittgenstein does not want to say that the meanings of "understands" or "knows" are always clear and easily definable. In fact, that they are not is what is wrong with the traditional philosophical approach. The philosophical approach tries to determine the essences of both "understanding" and "knowing." That is, it seeks clear and

---

[54] The English translation is "My cat is black".

distinct definitions. However, meaning is use and use depends on context and a form of life. Therefore, we can conclude that the meanings of "understands" and "knows" vary depending on the relevant context and specific language game. This variance means that there may be contexts where "I know" and "I understand" can be used interchangeably. So, we have to look to the specific context.

In the context of the moral conversation, we use "knows" as an ability to experience the reactive attitudes and thus to morally interact with others. "She knows the right thing to do" can only be uttered in reference to an individual who has knowledge of good and evil (or right and wrong). "She doesn't know any better" is uttered in reference to a moral innocent. For more morally disillusioned agents, we might use "understands" to indicate a deeper comprehension of good and evil. I contend that this comprehension is the ability to raise ethical questions. "She understands evil" can only be uttered in reference to an individual who has the ability to raise questions concerning evil. That is, part of her understanding is recognizing the complexities of a moral term like "evil." Because it is complex, she raises questions about it. What is the nature of evil? How might it differ from wrongdoing? How can it be that good people do dreadful things? Posing such questions requires understanding. I show later that this understanding involves the use of the moral imagination. I now turn to the passages in *PI* where Wittgenstein is also pondering the connection between "understands" and "knows" to further elucidate this distinction.

We already saw that in passage 150 it seems as though Wittgenstein distinguishes between "to know" and "to understand." He seems to claim that

they are related, but not identical. However, in the following passage he seems to contradict himself. He states,

> But there is also *this* use of the word "to know": we say "Now I know!"—and similarly "Now I can do it!" and "Now I understand!"[55]

At first glance it may seem as though Wittgenstein is contradicting what he wrote in 150. However, such a contradiction may not be accidental on Wittgenstein's part. By emphasizing the word "this" he is (possibly) explicitly acknowledging that there is more than one meaning (i.e. use) of "knows" or "understands." After he states in 150 that there is a relation between "knows" and "understands" implying that they are distinct, he states in 151 there is also a usage whereby we equate the two words. There are two important lessons to gather from this passage. The first is that we sometimes mean the same thing with "knows" and "understands". The second is that what we mean is "Now I can do it" or "Now I have the ability to go on."

In the remainder of 151 he describes the student with the numerical series who is able to complete the series and then yells "Now I can go on!" and "Now I know the series!" Perhaps Wittgenstein has tested and discarded the idea in 150 that "knows" and "understands" are distinct and we can conclude that they are used in the same manner. However, in the next passage (152), Wittgenstein asks the question, "But are the processes I have described here *understanding*?" He immediately questions the conclusions of the previous passage. The motivation for this questioning is that there are situations we can imagine where someone

---

[55] Ibid., 151

completes a numerical series, feels that they can go on, but yet we would not say that they understand. At this juncture in the discussion, he has both given us examples when we use "knows" and "understands" interchangeably, and also examples where we distinguish between the two. I think the reader is meant to feel confused at this point. As he states in 153, "I am in a muddle."

This feeling of being in a muddle might only be a result of wanting to define "knows" and "understands" clearly and distinctly, that is, philosophically rather than grammatically. What Wittgenstein does in these passages is show that it is simply empirically true that we sometimes use "knows" and "understands" interchangeably and we also sometimes distinguish between them. Yet, he still claims that both mean the ability to go on. The question now arises how it is that they can be used differently and yet both be used to communicate "Now I can go on." At 156, Wittgenstein writes that "This will become clearer if we interpolate the consideration of another word, namely 'reading.'"[56] This consideration will draw an analogy between reading and being literate on the one hand and knowing and understanding on the other. In passage 156, Wittgenstein remarks that when he considers "reading" he is not considering the understanding of what is read as part of "reading." By "reading" he simply means the rendering out loud of what is written or printed. This remark is key to distinguishing between reading and being literate and will shed some light on the distinction between "to know" and "to understand."

---

[56] Ibid., 156.

As Wittgenstein points out we can distinguish between reading and reading with understanding. (For the purposes of our present discussion, I am going to call reading with understanding "being literate.") The first question to ask is why this discussion of reading shows up in the section of *PI* on knowing and understanding. It does not, however, merely show up. Wittgenstein explicitly uses the discussion of reading to help clarify the meaning of "knowing" and "understanding." He does so because reading is clear example of being able to go on. When one reads, one literally goes on to the next word or sentence in the series. Thus, Wittgenstein can help clarify what he means by "being able to go on" by analyzing reading. In the context or language game of reading, one cannot go on to the next word or even pronounce a single isolated word without having mastered the skill of reading, i.e. rendering out loud what is written or printed. As we saw in passage 150 understanding, knowing, and being able to go on are all connected to mastery. For example, mastering chess entails having mastered the rules of the game. Once you have mastered the rules of the game you no longer need to consult the rules to be able to go on.[57] Being able to read entails having mastered the rules of the game of reading. One no longer needs to consult the rules of pronunciation (i.e. sounding out) to render the printed word aloud. Now we can better see the connection between reading and understanding, knowing, and being able to go on. Yet, a question remains: why does Wittgenstein make sure to distinguish between reading and reading with understanding (i.e. being

---

[57] But mastering the rules does not entail mastering chess. Mastering the rules of chess does not make one a master of chess. It simply means that one need not consult the rules to make a play. Conversely, mastering chess does entail having mastered the rules.

literate)? This is curious given that he states that reading is being able to go on, understanding is also being able to go on, but that reading does not entail understanding what is read.

Let's return to my previous example of the Czech sentence. In this example, we said that I know what "Moje kočka je černá" says. It says "Moje kočka je černá." However, I do not understand what it says (assuming I am not in a context that makes it clear to me how it is being used). If we take "understanding" and "knowing" to be related to each other in this way, then there is an analogy with reading and being literate. I know what the sentence says simply by identifying the words present in the sentence. I read the sentence aloud simply by pronouncing the sounds. In neither case would we say that I understand the sentence. However, when I understand the Czech sentence I understand what the speaker is communicating and how she is using the sentence. I can then reply to her or raise a question about what she just said. Similarly, when I am literate I read the sentence and also understand what is being communicated to me as the reader. Thus, it seems we have an analogy between knowing/understanding and reading/being literate. Perhaps we are out of the muddle!

Now we are in a position to see how knowing and understanding are both an ability to go on, but in different ways. Reading and being literate both involve mastery and the ability to go on. However, they do so in different ways. These different ways correspond to different phases of moral disillusion and is the key to the distinction (we sometimes make) between "knowing" and "understanding." On the one hand, it is correct to say that both knowing and understanding can be

used to communicate the ability to go on. On the other hand, this ability is not always the same. The circumstances of our use of these terms expose differences in their meaning(s). Sometimes certain circumstances dictate that we use "to know" and "to understand" differently from one another as we saw in some of our examples above. Other circumstances might show their usage to be similar or even the same. For example, we can imagine a situation where these two sentences are equivalent: "I know what you mean" and "I understand what you mean." This variance is what Wittgenstein tries to illustrate through his own confusion and discussion with the interlocutor about "to know" and "to understand." Various circumstances result in various uses and, therefore, meaning of terms. He states,

> Thus what I wanted to say was: when he suddenly knew how to go on, when he understood the principle, then possibly he had a special experience—and if he asked: "What was it? What took place when you finally grasped the principle?" perhaps he will describe it as much as we described it above—but for us it is *the circumstances* under which he had such an experience that justify him in saying in such a case that he understands, that he knows how to go on.[58]

Wittgenstein seems very much aware that the meaning of "knows" or "understands" is confusing. Much of this confusion arises because we want to define them as mental states or clearly and distinctly. If, however, we recognize that differing circumstances result in different ways in which one can go on, then we have understood "knows" and "understands." They are both the ability to go

---

[58] Ibid., 155.

on, but how we use the words depends on specific circumstances in a form of life.[59] They are both the ability to go on but sometimes in different ways.

One of the different ways shows up in the context of the moral conversation. Adam and Eve know good and evil, but they do not understand good and evil. They can go on in one way but not another. They can go on as appropriate candidates of, and individuals who experience, the reactive attitudes, or as agents who are receptive and reactive to reasons. When God punishes them they have the ability to feel regret, attempt to excuse their action (Adam blames Eve and Eve blames the serpent), and feel that God's wrath is either warranted or not. However, they cannot go on in the conversation beyond stating what they have been told is good or evil. They cannot raise questions about what God has said is good or evil. They cannot ask why God planted the tree in the garden if he did not want them to eat from it, or how knowledge of good and evil makes one god-like, or even how they could know that an action could be wrong if they did not already have knowledge of good and evil and thus wouldn't it be fair to excuse them based on their lack of knowledge. They cannot raise questions because they have not understood. Understanding good and evil is different from knowing good and evil and includes an ability to question moral principles because one understands the complexities of moral life. I contend that to

---

[59] See, Margaret Urban Walker, *Moral Understandings* (New York: Oxford University Press, 2007). Walker's notion of narratives of relationships is one way to interpret the ability to go on. A narrative of relationship is a story of the relationship's acquired content and developed expectation and its possibilities for continuation. Telling the narratives of relationships is a needed component of moral understandings. Similar to my view, Walker argues that there are various moral understandings rather than one privileged universal moral position from which to pass judgment. I elaborate on the moral need to understand our relationships in Chapter 5.

understand the complexities of moral life, one must shatter moral illusions. In the subsequent chapters I will explain these phases of disillusion.

Wittgenstein helps explain that "to know" and "to understand" are not always identical. In the myth of Adam and Eve we are told that when the two acquire knowledge of good and evil "their eyes were opened" and they see that they are naked. In the story "knowing" is tied to "seeing that." If seeing that is knowing, and understanding is distinct from knowing, then there is more to moral understanding than seeing that. However, we can also use "I see that…" in contexts of understanding as well. This dual use of "I see that" indicates that the context will dictate whether the utterance refers to knowledge or understanding. In the context of the moral conversation there are two levels of seeing that: one is knowledge and the other is understanding. I next explain in what way "seeing that" is knowing. After, I explain what kind of "seeing that" gives us understanding.

4. Seeing That

The Adam and Eve myth describes the loss of innocence as a transition from "seeing" to "seeing that." "Seeing that" is laden with knowledge or understanding of the significance of the event or object seen. One person might see a child being ushered into a van without seeing that the child is being kidnapped. "Seeing that" involves a "knowing that" or an "understanding that" as well. The innocent certainly see what they do. When they lose their innocence they *see that* what they do often has a moral content as well. They come to know that their action sometimes has moral significance. But they may not yet

64

understand this moral significance. This is why "seeing that" can correspond either to knowing or to understanding. It is crucial to my account to examine more closely the kind of knowledge that is inherent to "seeing that." To do so, I turn now to N.R. Hanson's essay "Observation."[60]

Hanson begins his essay by asking whether Kepler, who regarded the sun as fixed, and Tycho Brahe, who thought the earth was fixed and all the celestial bodies orbit around it, see the same thing in the east at dawn. Is there a sense in which the two individuals do not see the same thing although they are visually aware of the same object? Hanson answers this question in the affirmative by equating "seeing" with "seeing that." Because the two individuals differ in their knowledge of what they see, they see different things. Although I believe his answer to this question is incorrect, his analysis of "seeing that" is quite helpful to our understanding the kind of knowledge acquired through the loss of innocence.

Being visually aware of the same object is not equivalent to seeing the same thing. This is because "[s]eeing is an experience."[61] An image reflected onto one's retina is not sufficient for seeing. If it were, then it would be appropriate to claim that a camera sees. "People, not their eyes, see."[62] In some sense Tycho and Kepler see the same thing – namely a bright yellow luminescent disc in the sky. But in another sense they do not. Tycho sees one of earth's satellites and Kepler sees the center of the galaxy. Hence, something is very different about their visual experiences. A visual experience is not merely the physical state of light

---

[60] Norwood Russell Hanson, *Patterns of Discovery* (New York: Cambridge University Press, 1965).
[61] Ibid., 6.
[62] Ibid., 6.

reflecting upon the eye. Rather, it involves interpretation of what is seen.

However, this is not an ex post facto interpretation. It is not a mental state

applying a principle to an experience just had. He illustrates this point by

examining various optical illusions.

Hanson presents the reader with a number of images that portray different

objects when seen from different mental perspectives. One might see an old hag

or a young girl, a rabbit or a duck.[63] These images serve to illustrate that "one

does not first soak up an optical pattern and then clamp an interpretation on it."[64]

Rather, "for you and me to have a different interpretation…just is for us to see

something different. This does not mean we see the same things and then interpret

it differently." [65] The individual who sees the rabbit simply sees differently from

the one who sees the duck. The different perspectives on a single image are

examples of different things being seen without any interpretation being

superimposed on the sensation. "Seeing is not only the having of a visual

experience; it is also the way in which the visual experience is had."[66] The

experience is had in a particular way due to the knowledge that one has of the

object seen. Therefore, seeing differently rests on knowing differently.

Hanson defines seeing this way: "seeing an object x is to see that it may

behave in the ways we know x's do behave."[67] The seeing of x is thus shaped by

prior knowledge of x. Since Tycho and Kepler have different prior knowledge of

---

[63] Wittgenstein plays with these images as well.
[64] Ibid., 9.
[65] Ibid., 9.
[66] Ibid., 15.
[67] Ibid., 22.

66

the sun, they see differently when they see the sun. Thus, for Hanson seeing simply is seeing that. That is, seeing always already includes knowledge.

Hanson's description of "seeing that" as laden with knowledge is correct, but his equating "seeing that" with "seeing" is incorrect.[68] If it were true that any act of seeing requires prior knowledge of the object seen, then ignorant people would be blind. In his essay "'Seeing' and 'Seeing That', 'Observing' and 'Observing That,'"[69] Peter French asks us to consider an example of a child and an astronomer alternately looking through a telescope pointed in a particular direction. The child reports seeing only a white blotch in a black background. The astronomer tells the child that it is a nebula. Under Hanson's explanation, the child cannot be said to have seen a nebula. But rather than accepting such a counterintuitive claim we need only recognize that the child did indeed see a nebula, but he did not see *that* it was a nebula. The knowledge that Hanson argues is laden in every act of seeing is actually laden in every act of seeing that. In order for the child to see that it is a nebula he needs to possess some knowledge (but not necessarily any understanding) of what nebulas are. The presence of knowledge, rather than being a defining characteristic of seeing generally, is what distinguishes "seeing" from "seeing that."

Now if in every "I see that" an "I know that" can be unpacked, it does not follow that an "I know that" can also be unpacked from all of the cognates of "I see that." Hanson does not distinguish "seeing that" from its cognates, but French

---

[68] Hanson often describes seeing that as "theory-laden" which strikes me as too strong. One can have knowledge of an object without grasping any theory that can be applied to it.
[69] Peter A. French, "'Seeing' and 'Seeing That', 'Observing' and 'Observing That'," *American Philosophical Quarterly* Monograph Series: Studies in Epistemology no. 9 (1975): 89-97.

is careful to do so. One distinction that is germane to the present discussion is that between "seeing that" and "observing that."

Hanson assumes that "seeing that" and "observing that" are more or less equivalent terms. French notes, however, that "observing *x*" is something different from both "seeing that *x*" and "seeing *x*." Someone who observes as a profession, say a marine biologist who observes whales, would utter an infelicity were she to say "I see whales" rather than "I observe whales" when reporting what she does for a living. The observing entails a looking in order to gather information. "'To observe' is to look with a purpose."[70] One can be better or worse than others at observing. One can be a careful observer or a careless observer. As French puts it: "To be a careful observer is to pay heed to the object of observation so that one could answer a great number of potential questions about it. The careless observer finds himself in the position of having to answer many questions with 'I do not know' or 'I did not notice,' etc."[71] Described in this way, we can see that observing is a skill; a skill one must develop. Seeing, however, is not a developed skill, but rather a capacity.

There is another important aspect to the distinction between "seeing that" and "observing that" that French mentions only in passing, but which I think is quite significant to understanding the kind of knowledge acquired by the newly non-innocent and how it is distinct from a more developed moral understanding. He notes that one can utter "I observe that" only when the proposition following "that" is of a non-hypothetical sort.

---

[70] Ibid., 95.
[71] Ibid., 95-96.

One cannot, for example, say "I observe that if I throw a rock at the window it will shatter." One can, of course, see that a rock thrown at the window will shatter it. One can observe that the window is glass, and when the rock is thrown one can observe that the window did break. What one cannot observe is that something *will* happen. No one can observe that it will rain this afternoon, but many of us can see that it will.[72]

This is an important point that French raises. Utterances with hypothetical propositional content can follow "I see that," but not "I observe that." Here we begin to see where the "seeing that" of understanding departs from the "seeing that" of knowing. It is hypothetical content that partially makes up the *understanding* of the non-innocent and gives them the ability to raise ethical questions. There is a more developed "seeing that" beyond that which describes Adam and Eve or the newly non-innocent. The ability to consider hypothetical situations distinguishes knowing as seeing that from a more developed moral understanding as seeing that. The consideration of hypothetical situations relies on the imagination.

5. The Moral Imagination

In *Meaning and the Moral Sciences*[73] Putnam argues that imagination is a necessary element of practical knowledge. In attempting to answer the question of how to live (which he considers to be more fundamental to ethics than deontic questions), he says that an agent must be able to consider various hypotheses. The imagination provides the agent with the capacity for hypothetical thinking.

---

[72] Ibid., 97.

[73] Hilary Putnam, *Meaning and the Moral Sciences* (Boston: Routledge and Kegal Paul, 1978).

Although these hypotheses may not be verifiable in the way that scientific

hypotheses are, they are not thereby stripped of their value for practical or moral

knowledge.[74] Moral knowledge (what I call moral understanding) requires the

capacity to consider circumstances that could possibly obtain or the capacity to

entertain different ways of life.

For example, Putnam asks us to consider a mountain climber who figures

out how to climb a mountain by climbing it in his head.

> A man is climbing a mountain. Halfway up he stops, because he is
> unsure how to go on. He imagines himself continuing via one
> route. In his imagination, he proceeds on up to a certain point, and
> then he gets into a difficulty which he cannot, in his imagination,
> see how to get out of. He then imagines going up by a different
> route. This time he is able to imagine himself getting all the way to
> the top without difficulty. So he takes the second route.[75]

This example shows that using the imagination to entertain various future

possibilities is a perfectly reasonable way to solve a problem. Therefore, moral

reasoning, or the attempt to solve a moral problem, may require not just logical

faculties in the narrow sense, but also "our full capacity to imagine and feel."[76]

(Significantly, but perhaps unaware of its significance, Putnam says that the

climber's uncertainty keeps him from being able "to go on.")

This use of the imagination leads Putnam to see how useful literature is to

our moral sensibility. Literature can confront us with various hypotheses

concerning how to act or how to live. These hypotheses may not be empirically

---

[74] Putnam does not distinguish between knowledge and understanding as I do. Therefore, his use of "knowledge" is often what I mean by "understanding".

[75] Ibid., 85-86.

[76] Ibid., 86.

verifiable, but they give us conceptual knowledge. "Thinking of a hypothesis that one had not considered before is *conceptual* discovery; it is not empirical discovery, although it may result in empirical knowledge if the hypothesis turns out to be correct."[77] Literature gives us this conceptual knowledge by allowing the reader to see the world as it looks to someone (a literary character) who holds certain beliefs about the world or about how to live. One might disagree with a novel's hypothesis, but one is now aware of a possibility of which one was previously ignorant. I think that Putnam's account is quite illuminating, but I will add that the imaginative consideration of hypothetical situations gives us not only conceptual knowledge, but understanding exercised as an ability to participate in the moral conversation and an ability to raise ethical questions. Therefore, it is more appropriate to say that after reading literature I have gained understanding rather than knowledge.

Putnam's example of the mountain climber illustrates that one function of the moral imagination is to reveal the various consequences of different courses of action. Just as the climber imaginatively runs through different paths he can take to ascend the mountain, the moral agent can imagine the way in which different courses of action will have different ethical outcomes. For example, on Saturday I discover that my bike has been stolen from my back porch. On Sunday, someone shows up at my house with my bike. He tells me that he stole it the previous day, but then regretted his actions and decided to return it. I imagine how things will

---

[77] Ibid., 90.

play out both for me and for him if I simply accept his apology or if I call the police and report his crime.

Imagining consequences is only one use of the imagination. Other uses of the imagination include imagining similar situations and asking myself what I would think or do in that situation. Or, I can imagine what it is like to be someone else. I can imagine what it is like to be more virtuous than I in fact am, or conversely what it is like to be a scoundrel. For example, as I am walking through downtown Tempe I pass by a homeless person. He asks for a handout. Although my first inclination is to keep on walking, I can imagine what a more compassionate person might do in this situation. Or, as I write a check to donate money to the Arizona Animal Welfare League I can imagine what it would be like not to care about the suffering of others. Life would be more carefree. Furthermore, I might imagine that I am a participant in a way of life that is not my own. For example, when I visit my in-laws in the Czech Republic I can imagine what they might do in a certain situation for the purpose of figuring out what would be the kind or respectful thing to do. Perhaps I should kiss them when I greet them. In all of these situations I imagine hypothetical situations to figure out what to do, or how to go on.

But I might also imagine how to go on to bring about evil. I can imagine what the worst consequences for someone else would be, how a vicious person might act, or imagine a different way of life in order to act cruelly or disrespectfully. In fact, it seems that most evil people throughout history were quite imaginative. How else could the Nazis have come up with such innovative

72

ways to humiliate and torture their victims? Thus, the moral imagination need not lead one to be morally good. It can easily lead to wickedness. It does, however, give one the ability to go beyond simple knowledge of good and bad.

There are numerous forms a hypothetical, or conditional, proposition could take on. I will focus on the three main forms that grammarians label conditional 1, 2, and 3. "Conditional 1" is used to make a claim about the future the speaker thinks is likely to occur. The examples French uses in the above passage are of this form. For example, "If I throw this rock at the window, it will shatter." The speaker uses the simple present tense in the antecedent and the verb "will" in the consequent to indicate her belief that is likely that the window will in fact shatter should the first condition obtain. "Conditional 2" is used to make a claim about an event that the speaker believes is not likely, but not impossible, to occur. For example, I might utter, "If I won the lottery, I would buy a mountain lodge in the Czech Republic." Here I use the simple past in the antecedent and the verb "would" in the consequent to indicate that I do not think it very likely that I will win the lottery. This distinction is best seen by examining the following pair of conditionals: "If I become President, I will lower taxes" and "If I became President, I would lower taxes." The first can be uttered by a candidate who believes it is likely (or at least wants to communicate to voters that she thinks it is likely) that she will become President. The second of the pair is uttered by an ordinary person (any non-candidate), perhaps sitting in bar discussing politics with her friends. It is unlikely that this person will become President in the near future. In both situations the hypothetical could follow "I see that," but in neither

73

case "I observe that." For example, one could discover upon reflection that she would lower taxes if she ever became President and then utter, "I see that if I became President, I would lower taxes." But one cannot say, "I observe that if I became President I would lower taxes." Hypothetical utterances, and hence hypothetical knowledge, accompany the sort of understanding belonging to "seeing that."

There is another conditional, the one grammarians call "Conditional 3," that is important, not for exploring future possibilities, but for imagining how the past might have been different and thereby possibly influencing my future decisions. This is the conditional that refers to the unreal past. That is, it describes what would have happened, had some condition obtained. For example, after getting back a test he failed Johnny utters, "If I had studied harder, I would have passed." He uses the past perfect in the antecedent to indicate that he did not in fact study very much and "would have" in the consequent to indicate how things might have gone had the condition in the antecedent actually taken place. This kind of hypothetical thinking is an integral element of the more developed stage of "seeing that." It is to understand and acknowledge what occurred and how things might have been different. This conditional shows us that it is not only future-oriented hypothetical propositions that help comprise the sort of understanding accompanying "seeing that," but past-oriented ones as well.

When one contemplates what one will, would, could, would have, might have, could have, or should have done given certain conditions, one has the capacity to experience some of the reactive attitudes. For example, I cannot regret

74

a past action if I cannot imagine how I could have and should have acted differently. Now we can see that in addition to "seeing that" as a kind of knowing acquired by losing innocence, moral understanding relies upon the moral imagination to entertain hypothetical (or possible world) situations. That is, the shattering of innocent moral illusions includes the imaginative ability to grasp both what one might or should do, as well as what one might have or should have done. One *sees that* certain actions are morally significant and one understands how things could have been or could be otherwise.

These various uses of the imagination rely on shattering the illusion that morality is "all about me." To imagine what it is like to be someone else, or to participate in a way of life that is not my own, or how different consequences might affect others, depends on recognizing that my experiences are not the only ones of significance. To imagine hypothetical situations is to entertain different possibilities. The imaginative (re)creation of possible scenarios allows one to gain insight into the significance of different ways of life and courses of action. Not only can I imagine various situations, I can also morally evaluate the relevant possibilities in terms of good and evil. The moral imagination explores future possibilities and can help us avoid repeating past mistakes. These exploratory and corrective functions may help me see that the moral principles I had adopted are not adequate for every situation I can imagine. I gain the ability to question the moral principles I had previously accepted as adequate.

But "seeing that" is not merely a Gestalt shift as Hanson wants to describe it. French notes this as well:

> The metaphor of the eyes being opened suggests that when people lose innocence they undergo a Gestalt shift, seeing the old hag where earlier they had seen the young woman, or the rabbit where before the duck was evident….I would rather suggest that those who lose innocence learn in a very personal way how to redescribe their situations, their experiences, and their actions. In effect, they learn first hand or in the first person how to appropriately use the language of responsibility with respect to themselves.[78]

I already discussed in the first chapter how those who lose their moral innocence acquire this moral vocabulary and learn how to redescribe the occurrence of wrongdoing or evil with respect to their own agency. I have explained in this chapter that one can both know that one has the potential for wrongdoing as well as understand that one has the potential for wrongdoing. Wittgenstein helps us see that both knowing and understanding are abilities. The ability that one acquires is minimally the ability to have the reactive attitudes and to have guidance control. But upon shattering other moral illusions one acquires the ability to raise ethical questions. We can come to question whether the moral directives we have adopted are in fact sufficient to avoid wrongdoing.

At the minimal level one has the ability to experience the reactive attitudes and to have guidance control. Strawson and Fischer and Ravizza explain this minimal condition of moral responsibility. This is the level of being and knowing that one is morally responsible. But this ability is not the only moral ability we have. As we discussed above, the ability that accompanies moral disillusion also involves employing the moral imagination to contemplate hypothetical situations.

---

[78] French, "Losing Innocence" 35.

This ability is not necessarily the kind of ability exhibited by the student who completes the formula; it does not entail answering a question. That is, the ability does not necessarily give us answers. Rather, I contend that the ability to go on in a moral sense, involves *raising* ethical questions. The non-innocent can ask: What ought I do? Am I blameworthy? What should I have done? What would I do in such and such a situation? The ability to raise these ethical questions is the ability to engage in ethical contemplation or ethical discourse. Now the individual can go on and actively participate in the moral conversation. In subsequent chapters I will show that this ability to raise ethical questions is further developed by the confrontation with, and contemplation of, evil, unchosen evil, and inescapable moral wrongdoing. The ability to raise ethical questions is a defining mark of moral disillusion.[79]

We see this ability first-hand when we teach ethics courses. The ethics courses I took as a student and now teach as an instructor do not teach students how to be morally good (that might truly be a case of the blind leading the blind!), rather we teach them to shatter naïve illusions about moral life. In my courses I present arguments that support various ethical theories, and I always include a discussion of relevant objections to whatever theory we are discussing. I think most philosophy instructors do something similar. I used to worry that my ethics courses, and those taught similarly, were churning out moral skeptics or moral nihilists semester after semester. On more than one occasion a student has said to

---

[79] Kierkegaard's philosophy of moral development, which I will discuss in chapter 6, will further illustrate how the raising of ethical questions is required for a more individual and personal sense of moral responsibility.

me, "I came into this course one hundred percent certain of my ethical convictions, but now I am not so sure!". Have I taught these students to believe that there is no point in trying to act ethically if no single ethical theory stands without objection? Perhaps. Although I hope that my students learn something about being a good person, I do not consider it my job to teach them to be morally good. It is my job to help them contemplate ethical principles and problems. When they lose their naïve understanding of good and evil, they gain an ability to raise ethical questions. This ability is a mark of their disillusion.

6.   Conclusion

In this chapter I have distinguished two levels of moral disillusion: knowledge and understanding. The former includes knowing the difference between good and evil and knowing that one is an appropriate candidate for the reactive attitudes. The latter includes the ability to imagine different future courses of action as well as how the past could have or should have been different. In both knowing and understanding, one can *see that* one's action has moral significance, but in different ways. These abilities then allow one to enter and participate in the moral conversation. One is an appropriate candidate for the reactive attitudes, one can experience the reactive attitudes, and one has the ability to raise ethical questions. Raising, rather than answering, ethical questions, is the hallmark of moral disillusion. We will continue to see that this ability provides the means for taking responsibility for one's actions. In chapter 1 I stated that evil is distinct from wrongdoing. In the next chapter I argue for this assertion and explain the kind of disillusion involved with recognizing evil.

CHAPTER 4

EVIL

1. Introduction

In chapter 1, I disagreed with French that an acquaintance with evil is necessary for entry into the moral community. One can enter the moral community through an acquaintance with wrongdoing. However, becoming acquainted with evil does characterize another significant moral disillusion. But what is evil? In this chapter I unpack the concept of evil by examining its distinct qualities and explain what the acquaintance with evil involves. Although I argue that defining evil too stringently may lead to excluding some acts that deserve to be called evil, we can still discern certain features that are always present in an evil act. One of these features is that the victim is always in a situation of extreme vulnerability. Understanding the state of extreme vulnerability comprises a moral disillusion. For, it is by *seeing that* another is extremely vulnerable, and that I am capable of either directly or indirectly[80] exploiting this extreme vulnerability, or being causally complicit in its occurrence, that I become personally acquainted with evil. In this chapter I first explain why evil is a necessary part of moral discourse and how we ought to define it. Then, I examine some well-known theories of evil to formulate some general characteristics of evil acts. Next, I use Harry Frankfurt's conception of a person to describe a particular kind of evil act. Finally, I will be in the position to explain that extreme vulnerability is an essential feature of evil acts. Understanding this fact coupled with the realization

---

[80] In chapter 4 I examine how we can indirectly, complicitly, or unintentionally bring about evil.

79

that one has the ability, either directly or complicity, to abuse another's extreme vulnerability comprises the acquaintance with evil necessary for this phase of moral disillusion.

2. Why "Evil"?

Outside of a religious context, there has been little philosophical discussion of a concept of evil until recently. Religious thinkers have long had to grapple with the problem of evil. Philosophers, however, have not spent much time discussing evil as a moral category. The few philosophers who do discuss evil usually do so by equating it with wrongdoing. In *De Malo* (*On Evil)* Aquinas treats the subject of evil, but defines it as sin by which he means wrongdoing. There is no indication that he conceives of evil as distinct in any way from wrongdoing. Such discussions of evil have been the most common in the history of philosophy. "Evil" is considered simply to be the opposite of good, or that which is most undesirable, or that which is wrong. Of course one cannot think of "evil" and the history of philosophy without thinking of Kant. He famously discusses radical evil in *Religion within the Bounds of Mere Reason*. I will say more about Kant's account below, but his conception of evil does not give us real insight into how evil differs from wrongdoing. That is, he argues that evil is the propensity to act on maxims contrary to the moral law. But even if all maxims contrary to the moral law are wrong, it does not follow that they are evil. Furthermore, his definition of evil cannot be accepted without also accepting his unrealistic metaphysical dualism.

One might think that the reason philosophers have not spent much time analyzing evil is because it is not, in fact, conceptually distinct from wrongdoing. The term "evil" is therefore used emotively, dramatically, poetically, and even inappropriately to refer simply to wrongs. For example, in his account of the loss of innocence French calls undeserved harm "evil." However, undeserved harm can refer simply to a wrong. In *Mortal Questions* Nagel asks about the evil of death. "I want to ask whether death is in itself an evil; and how great an evil, and of what kind it might be."[81] Throughout the rest of his chapter it becomes clear that he means to ask what is bad about death or why death is not good. He uses "evil" as a device to attract the reader and to motivate interest in his discussion; he does not conceptually distinguish evil from bad. This usage of "evil" is not uncommon in philosophy. Given this backdrop, it might be tempting to conclude that "evil" does not refer to a moral category distinct from that already circumscribed by the term "wrong." Hence, an analysis of evil does not yield any valuable information about the moral landscape that cannot already be discerned via an analysis of wrongdoing. Furthermore, some might think that "evil" is used mainly to demonize the wrongdoer, to cast aspersions on her character, to imply that she is irredeemable or cannot, and should not, be forgiven. In these cases, it is not simply that "evil" is superfluous, but actually dangerous. The use of the term would skew our understanding of the wrong act and the wrongdoer thereby confusing our moral judgment. It would lead to inappropriate emotional categorizations of individuals or their acts as evil. Because the term carries more

---

[81] Thomas Nagel, *Mortal Questions* (New York: Cambridge University Press, 1979), 1.

reproach and contempt than "wrong" or "bad", use of the term would condemn the wrongdoer unfairly.

However, even if it is true that the term "evil" is sometimes used emotively, dramatically, poetically, and occasionally inappropriately to refer to acts which should be labeled "wrong," it does not follow that there are no acts that are, in fact, "evil." The fact that "evil" is used to dramatize a wrongdoing does not entail there are no evil acts that are distinct from wrongdoings. It means simply that we have to be careful when appealing to ordinary language when defining "evil." When we examine immoral acts, it is not difficult to see that acts come in all degrees of moral badness. Clearly, there is a moral difference between robbing a bank and genocide. Lumping both acts into the category of "wrongdoing" does not adequately capture the difference between the two. The difference is, however, marked by the concept "evil." In order for our moral language, censure, and opprobrium to be accurate we need to distinguish evildoing from wrongdoing.

Here I present what I believe are a few uncontentious examples of evil acts. I include them not to titillate, but to provide real evidence of the need to think of evil as distinct from wrongdoing.

> In October 1978 a man named Lawrence Singleton offered to take a 15-year-old girl, Mary Vincent, from Berkeley to Los Angeles. On the way she fell asleep, and after she fell asleep he took her to a canyon in Nevada, where he beat her, threw her into the back of his pickup truck, ripped off her clothing, tied her hands, raped her several times, later dragged her from the truck, held her hands down and chopped them off with his hatchet. "He chopped it three times. The blood was spurting all over". She was then tossed over a guard rail, stuffed into a culvert beneath a road, and left for dead.

But somehow she didn't die, somehow she survived, and she was later found, dazed and bleeding, naked with both arms chopped off below the elbows, blood streaming from them.[82]

In 2009 in Canada an Afghani man murdered his three daughters for dishonoring his family name. Because the daughters wore Western "revealing" clothing, and had boyfriends, the father likened them to prostitutes. The girls were found dead in car that had been pushed into a canal. They were 19, 17, and 13 years old. The father was quoted saying, "I would do it again 100 times".[83]

In February 2012, a Nepalese woman, Dhegani Mahato, was accused of being a witch and burned alive. She was a mother of two. She was attacked and set on fire by members of her family and others after a shaman accused her of casting a spell to make one of her relatives sick. She was beaten with rocks and sticks before being doused with kerosene and set on fire. The attack was witnessed by her 9 year-old daughter.[84]

In the late 1990's and early 2000's, the purchase of some diamonds helped fund devastating wars in Africa. Profits from the trade of conflict diamonds were used to fund armed conflict costing the lives of 3.7 million people. In Sierra Leone the profits supported the guerrilla army, the Revolutionary United Front, which cut off the hands, feet, lips, ears, and noses of civilians to keep them from

---

[82] Associated Press dispatch of March 1979, quoted in Marcus G. Singer, "The Concept of Evil." *Philosophy* 79, no. 308 (2004): 193.

[83] Ian Austen, "Afghan Family, Led By Father Who Called Girls a Disgrace, is Guilty of Murder," *New York Times*, January 19, 2012, http://www.nytimes.com/2012/01/30/world/americas/afghan-family-members-convicted-in-honor-killings.html (accessed February 27, 2012).

[84] Manesh Shrestha, "Nepalese Woman Accused of Witchcraft and Burned Alive," *CNN*, February 18, 2012, http://www.cnn.com/2012/02/18/world/asia/nepal-witchcraft-burning/index.html?hpt=hp_t3 (accessed February 27, 2012).

harvesting crops to feed the national army. To inspire fear and maintain power over villagers, children were kidnapped from villages, injected with cocaine, given a gun and told to kill their parents.[85]

The first act is an example of malicious intentional (perhaps psychopathic) evildoing. The second act is motivated by honor, and the third by religious belief. The fourth example shows that one might be causally complicit in bringing about evil, without doing so directly. We see from these examples, and I will argue below, that evil must not be sadistic or diabolical; there are many roots of evil. In addition to the above acts, we can add genocide, war rape, torture, and most cases of abuse of children. I do not claim this list to be exhaustive. In fact, in order to correctly describe an act as evil, we may need to consider the specific agents and the particular circumstances involved, rather than kinds of acts. This sort of fine-grained approach has the disadvantage of lacking a clear and concise definition of evil, but it has the advantage of accurately including all of those acts that deserve to be labeled evil.

3. How to Define Evil

The above examples show that there are convincing reasons for distinguishing evil from wrongdoing. How might we conceive of the distinction? Much of the philosophical literature on the subject over the past decade has focused on whether evil is quantitatively worse than wrongdoing, i.e. one murder

---

[85] Dick Durham, "Diamond Trade Fuels Bloody Wars," CNN, January 18, 2001, http://www.amnestyusa.org/our-work/issues/business-and-human-rights/oil-gas-and-mining-industries/conflict-diamonds (accessed March 13, 2012).

is wrong, but one thousand murders is evil, or whether evil acts have a unique quality or qualities which distinguish them from wrong ones.

But definitions of moral concepts, such as the good, are notoriously inadequate. Philosophers have built careers on pointing out the inadequacies of prior moral theories and definitions of moral terms. When a moral concept is strictly defined, a counterexample, improper exclusion, or contradiction is often found. One of the reasons this occurs is because moral life is quite rich, and strict definitions cannot usually account for all of its nuances. This reason also explains why abstract command theories of morality, like Kantian deontology or utilitarianism, run up against so many objections. Human interaction is often more complicated than a universal command theory can allow for.[86] I think we need to be careful not to be too hasty in defining evil too stringently. Definitions are static; they draw rigid lines around the concept defined. But evil appears in many forms. Therefore, we need to be careful that any definition of evil is not too restricting, thereby excluding acts that should be deemed evil. Although it might be practically onerous, we may need to examine particular acts as they occur to determine whether they are evil.

In his essay "Drawing Lines" James Rachels argues that we must determine our moral responsibilities to others in such a fine-grained fashion. Only by looking at the particular characteristics of another being (human or animal), can we know what sort of acts may harm it. Once we know what sort of acts may harm another, we can know what moral responsibilities we have toward it. The

---

[86] I shall address this issue in my chapter on Moral Tragedy.

particular ways in which a being is vulnerable to harm may be determined either by the characteristics of the species to which the individual being belongs, or may be given by the individual characteristics of the being. For example, if someone were to attack John, a classical pianist, and cut off his thumbs, he is harmed in (at least) two ways. First as a sentient being, he has been forced to undergo immense physical pain. This harm is determined by the fact that he belongs to a species that has the capacity to experience pain. It is harm qua human, not qua individual. Second, he is harmed because his ability to pursue his passion (playing the piano) has been destroyed. This harm is determined by the fact that he has a particular interest, namely playing the piano. It is harm qua individual, not qua human. Similarly, evil may be done to someone (human or animal) due to its species-determined characteristics or its individual characteristics. This realization is part of the acquaintance with evil necessary for this phase of moral disillusion.

Rachels points out that there are different approaches to defining moral standing each of which attempts to clarify to whom we owe a direct moral duty. At first philosophers thought that simply being human confers moral standing. This approach had the advantage of being nondiscriminatory, but has the disadvantage of being too vague. What is it about being human that gives us this status? Philosophers have often connected specific human characteristics or capacities to moral standing. For Aristotle, the human capacity for rationality gives humans importance. For Kant, human self-consciousness and the capacity to exercise autonomy obligate moral consideration. Utilitarians defend the theory that to have moral standing it is necessary only that one feel pain. But why must

86

we choose which of these characteristics gives moral standing? Each of these characteristics can convey moral standing on a being, and therefore any abuse or harm to any one of these characteristics could constitute a wrong.

If a being is autonomous, then harming its ability to exercise its autonomy is wrong. We usually do not consider children to be autonomous and we treat them differently because of it. Treating a rational and autonomous adult as a child would ignore or directly harm her ability to exercise her autonomy. The harm could come in the guise of physical restraint or harm to the capacity itself. Similarly, some beings are self-conscious and thus can reflect upon their own character and conduct and conceive of themselves as extended through time. As Rachels notes,

> [t]here are…a number of goods that self-consciousness makes possible: self-confidence, hope for the future, satisfaction with one's life, the belief that you are someone of value, and the knowledge that you are loved and appreciated by other people. Without self-consciousness, there could be no sense of pride or self-worth. [87]

Just like being autonomous, so too does being self-conscious make one vulnerable to a host of harms. "[Y]ou may feel embarrassed, humiliated, guilty and worthless. Because you can think about your own future, you may despair and lose hope."[88] Therefore, there are ways of treating beings that are objectionable based upon their capacity for self-consciousness. Of course, the ability to feel pain

---

[87] James Rachels, "Drawing Lines" in *Animal Rights*, ed. Cass R. Sunstein and Martha C. Nussbaum (New York: Oxford University Press, 2004), 168.
[88] Ibid., 168.

is a capacity that obviously makes one vulnerable to various harms. The fact that someone feels pain supplies others with a reason not to cause her pain.

These capacities, and perhaps others, give us grounds for not treating others in certain ways. It may be that a creature's autonomy obligates a certain kind of treatment, or it might be a creature's ability to feel pain that obligates a certain kind of treatment. For example, I should not walk up and hit you with a stick, not because you are autonomous, but because you feel pain. I should not treat you paternalistically, not because you feel pain, but because you are autonomous. Rachels adds the example that one should not tell her husband's friends that he is impotent, not because he is autonomous, but because he is self-conscious and would be humiliated.

As noted above, individual characteristics play a role in moral consideration as well. For example, if a student has a dream to become a philosophy professor, and I steal and destroy her philosophy books, I have harmed her in a way that is different than if I were to steal the philosophy books from a student who just finished her philosophy course and has no interest in philosophy at all. What makes certain kinds of treatment morally objectionable depends on the individual and the specific circumstances.

> Facts about people often figure into the reasons why they may or may not be treated in this or that way. Adam may be ejected from the choir because he can't sing. Betty may be given Prozac because she is depressed. Charles may be congratulated because he has just gotten engaged. Doris may be promoted because she is a hard worker. Notice, however, that a fact that justifies one type of

treatment may not justify a different type of treatment: Unless
something unusual is going on, we could not justify giving Betty
Prozac on the grounds she can't sing or throwing Adam out of the
choir because he has become engaged.[89]

Therefore, moral standing is always with respect to some particular mode of treatment. There is quite a long list of characteristics, belonging both to one's species and to one's individual interests that constitute morally good reasons why someone should or should not be treated in various ways. A sentient being should not be forced to undergo unnecessary pain, an autonomous being should not be unfairly coerced, a self-conscious being should not be humiliated, and those goods someone deems necessary for her pursuit of a good life should not be destroyed unnecessarily.

Rachels's account of moral standing helps us understand that there are different kinds of evil acts as well. Given the characteristics of the victim, an act can be evil for a variety of reasons. If any of the harms I have described above are excessive, or especially undeserved, or maliciously caused, then they may be evil. They may be evil for other reasons as well, and I take this up in the next section. Rachels admits that his approach makes it practically quite difficult, if not impossible, to enumerate all of the different ways every being ought not be treated. Similarly, it may be difficult to enumerate all evil acts. That may be the price of understanding the many forms an evil act can take on. It is a feature of moral disillusion to accept that moral life may not fit neatly into one definition or theory. Furthermore, as we have seen, a given act may be evil if it involves one

---

[89] Ibid., 167.

particular agent but not another. An act may be evil if it destroys one's ability to exercise one's autonomy, or it may be evil if it destroys the means which a particular individual requires to pursue her life plan and conception of the good. Hence, we may need to examine the particular features of acts to be able to correctly describe them as evil.

Despite this indication that evil acts may appear under many guises, there are shared characteristics. When we think about the above examples of evil one common denominator in evil acts is that they all require a certain kind of vulnerability on the part of the victim. I argue that acknowledging that others are sometimes extremely vulnerable and that one has the ability to exploit this vulnerability, either directly or indirectly, is the personal acquaintance with evil that shatters certain moral illusions. Before I turn to examine this vulnerability more directly, I look at prior theories of evil to provide some other general characteristics of an evil act.

4. What is Evil?

First, I look at why Kant's theory of radical evil is an insufficient account. His description is not only inadequate because one would have to accept his metaphysical dualism; there is also a glaring omission in Kant's theory. He focuses solely on the harm done to the moral character of the perpetrator, or potential perpetrator, and ignores the harm done to the victim. Such a perpetrator-centered approach leaves out one of the most salient features of an evil act: the significant or excessive harm done to the victim. It may be true that an evil act requires that the perpetrator of the act has certain intentions whether they are

motivated by malice, thoughtlessness, ambition or otherwise. But it cannot be that an evil act does not also require a certain level of harm[90] experienced by the victim. Kant does not adequately define evil because he ignores the harm done to the victim. As we will see in this section, this harm is necessary for an act to be an evil act.

Kant departs from theological tradition by rejecting the notion that sin is inherited from Adam and Eve. Because his notion of moral responsibility depends upon the free exercise of the will, he has no use for hereditary sin or guilt. For one to be morally responsible for her wrongdoing or evildoing, she must freely choose the maxim that results in a wrong or evil deed. Allen Wood and George di Giovanni point this out in their translation of the *Religion* by noting that Kant does not use the German word "Erbsünde," which means literally "hereditary sin," but opts for the Latin *peccatum originarium*, which does not imply heredity.[91] Kant does think that doing evil is innate to human nature, but only if we understand "innate to human nature" as referring to the ability or in his words "propensity" to do evil.

Kant says that "[w]e call a human being evil, however, not because he performs actions that are evil (contrary to law), but because these are so constituted that they allow the inference of evil maxims in him."[92] Immediately we see that the definition of evil depends not on empirical or external deeds, but on the intentions, or constitution, of the perpetrator. Kant believes that empirical

---

[90] I later define the relevant harm as irreparable life-wrecking harm.
[91] Immanuel Kant, *Religion within the Boundaries of Mere Reason*, ed. and trans. Allen Wood and George Di Giovanni (Cambridge: Cambridge University Press, 1998), xii.
[92] Ibid., 46.

acts are not sufficient to allow us to conclude that an individual is evil. We can

only infer from a number of consciously evil actions, an underlying evil maxim.

Although one is evil according to her constitution, one is not evil without

choosing to be so. The ground of evil cannot lie in natural impulses, for then one

did not choose freely and is not morally responsible, but only in the rule on which

one chooses to act.

> Whenever we therefore say, "The human being is by nature good,"
> or "he is by nature evil," this only means that he holds within
> himself a first ground (to us inscrutable) for the adoption of good
> or evil (unlawful) maxims, and that he holds this ground qua
> human…[93]

Because the disposition (i.e. the first subjective ground of the adoption of

maxims) is adopted through the power of choice, it is noumenon. In other words,

choosing one's maxims is an activity of pure reason, which is noumenal.

Therefore the evilness of one's deeds lies not in the suffering experienced by the

victim, which is empirical, external, and phenomenal, but by the kind of maxim

adopted by the perpetrator.

Kant's framework provides three ways in which one can do wrong. One of

these ways of doing wrong is what he labels "radical evil." To understand the

propensity for wrongdoing and evildoing, we have to understand the

predispositions that are directed towards the good, but can be corrupted into doing

wrong or evil.

---

[93] Ibid., 47.

Kant says we have three predispositions to good and three corresponding propensities for wrongdoing and evildoing. The first predisposition is to satisfy physical and psychological needs dictated by the preservation of the individual and species. This is a necessary predisposition for humans *qua* human, but various vices can be grafted on to it. Kant includes gluttony, lust and "wild lawlessness." The wrong associated with the corruption of this predisposition is *weakness*. In cases of weakness one has adopted the categorical imperative, but is sometimes too weak to follow it. Here one does wrong empirically.

The second predisposition is to rationally evaluate the satisfaction of basic and culturally conditioned needs in the light of our conception of happiness. One compares one's own status with that of others in the community and desires equal worth in order to be happy. Because of a "constant anxiety that others might be striving for ascendency" comes gradually "an unjust desire to acquire superiority for oneself over others."[94] The propensity for evil in these cases is *impurity*. One follows the categorical imperative, but for the wrong reasons. One does not do it out of duty, but for fear of a bad reputation, or out of inclinations. This is worse than weakness for Kant even though the outward empirical actions may be in accordance with the moral law.

The third predisposition is to submit our will to the command of the moral law, and thus to universalize the principles upon which we act. In the noumenal realm one adopts a supreme principle for oneself, such as the categorical imperative. At the phenomenal or empirical level one performs in accordance

---

[94] Ibid., 51.

with the supreme principle one has adopted. Here enters depravity or radical evil. The propensity for evil corresponding to this predisposition is the subordination of the moral law to the maxim of self-love. Radical evil consists in making self-interest one's supreme practical principle, and subordinating the moral law to it. In Kant's words, "the statement, 'The human being is evil,' cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it."[95] This is an intelligible (or noumenal) action. Therefore, evil does not consist fundamentally in temporal (empirical) violations of the moral law. We can be radically evil even if our temporal actions happen not to violate the moral law. This evil is radical because it comes about through one's own choice to subvert the moral law. The moral purity of the agent has been corrupted.

The difference between a propensity and predisposition is that the former is something we bring upon ourselves whereas the latter is natural and original. We bring this propensity upon ourselves by being too weak to follow the moral law, by following the moral law for the wrong reasons, or by choosing a maxim to act upon which deviates from the moral law. The propensity to evil is in human nature but not in the same way as the predisposition to good. Our predispositions, that is, our natural pursuit of happiness and satisfaction of our needs are good in themselves and should be guided by the moral law. Evil occurs when we are not so guided and choose self-love to guide our actions. Not following the moral law is due to frailty, impurity, or corruption. Therefore, humans are evil in that they

---

[95] Ibid., 55.

94

subordinate the moral law to some other incentive to act, namely self-love. Consequently, the sole difference between good and evil persons is which maxim guides them.

Clearly, Kant's main concern is the moral purity or character of the perpetrator or potential perpetrator of evil. This account leaves out that feature of evil acts which is integral to the act being evil, namely significant harm to the victim. Furthermore, Kant's three propensities to evil cannot explain the difference between a wrong act such as stealing and an evil act such as genocide. Even if it is true that the perpetrator was either too weak to follow the moral law, followed the moral law for the wrong reasons, or subverted the moral law to self-love, these three possibilities do not sufficiently capture what we mean when we refer to an act as evil, and not merely as wrong. Kant's account might be useful in thinking about the moral psychology of why some people act immorally, but it does not fully explain the degree of immoral action that the category of evil helps demarcate. I now turn to some more contemporary theories[96] of evil that do so. These theories will provide us with some general characteristics of evil acts and lay the ground for seeing that victims of evil are in situations of extreme vulnerability.

Claudia Card provides us with a coherent and fairly correct account of evil. She defines evil as "foreseeable intolerable harm produced by culpable

---

[96] The following theories of evil are in the tradition of analytic philosophy. For an engaging discussion of radical evil situated at the intersection of critical theory and philosophy of religion, see Martin Beck Matuštík, *Radical Evil and the Scarcity of Hope: Postsecular Meditations* (Bloomington, IN: Indiana University Press, 2008). Matuštík meditates on the loss and recovery of intransitive hope in the face of excessive and irreparable violence.

wrongdoing."[97] She notes that the intolerable harm experienced by the victim is the most salient aspect of an evil act. The harms, not the perpetrators' psychological states, distinguish evil from wrongdoing. But neither intolerable harms nor culpable wrongdoing alone is sufficient for evil. They must both be present for an act to be evil. Card's approach to evil is attractive because it focuses on that aspect of an evil act that strikes us as necessary for an act to be evil; namely, immense suffering or life-wrecking harm experienced by the victim (caused by human agency not natural causes). Although she emphasizes the intolerable harm condition of her definition of evil, she is careful also to include a "perpetrator component." This is the "foreseeable" element. An evil act is not just one in which intolerable harm obtains. The intolerable harm must have been foreseeable (although not necessarily foreseen) by the agent who caused the ham. Thus, "foreseeable" entails that there was some agent who was able to foresee the intolerable harm she caused. But the perpetrator need not have maliciously foreseen the intolerable harm she caused. "Foreseeable" does not mean maliciously intended. One need not intend to cause intolerable harm for the act to be evil. Finally, we are responsible for acts that have such consequences even without intending them. "For we can be responsible for causing what is reasonably foreseeable, even if it is not what we aimed for."[98] We can cause evil without directly intending to, and we can be morally responsible for doing so.

---

[97] Claudia Card, *The Atrocity Paradigm: A Theory of Evil* (Oxford: Oxford University Press, 2002), 16.
[98] Card, "Torture in Ordinary Circumstances," in *Moral Psychology: Feminist Ethics and Social Theory*, ed. Peggy DesAutels and Margaret Urban Walker (New York: Rowman and Littlefield, 2004), 151.

If Card's account of evil is correct, it is important to ask why it is correct. Answering this question will tell us why the elements of her definition are necessary for any definition of evil. I turn now to Paul Formosa's argument for why a "combination approach" to the concept of evil is appropriate.

Formosa points out that there are four different approaches to a theory of evil. On a victim approach, it is something about the kind and degree of harm inflicted upon the victim that constitutes an evil act. On a perpetrator approach, like Kant's, it is something about the perpetrator, such as an intention or motive, that constitutes an evil act. On a bystander approach it is something about the bystanders or evaluators of an act, such as the horror it inspires or its incomprehensibility[99] that constitutes an evil act. On a combination approach, which Formosa endorses, it is something about the combination of these factors that constitutes an evil act.

Formosa believes, as does John Kekes,[100] that there can be many roots of evil. It is not simply that one kind of motive or intention is a necessary and sufficient condition for evil. "Motives such as envy, malice, greed, hatred, boredom, honor, pride, revenge, ambition, thoughtlessness, a lack of self-esteem, ideology, and faith can all, at times, be roots of evil."[101] It may be empirically true that some of these factors result in the occurrence of evil more often than others, but any one of them may lead to evil.

---

[99] See Joel Feinberg, *Problems at the Roots of Law: Essays in Legal and Political Theory* (New York: Oxford University Press, 2003), p.183-192. Feinberg argues that evil is harm causing and wrongdoing that is incomprehensible. What distinguishes evil from other bad things is its capacity to generate puzzlement. The actor's motives and intentions are unintelligible to others.
[100] I examine Kekes's conception of evil in the following chapter.
[101] Formosa, 220.

Victim approaches are attractive because they can account for the many roots of evil. If the "evil-making" component of an act lies in the harm done to the victim, then an evil act need not be caused by specific motives or intentions. We need not try to learn about the psychology of the perpetrator, which even Kant admits we can only make assumptions about based on empirical acts, and we can focus on the palpable harm done to the victim. This approach has the advantage of excluding acts that, although maliciously motivated, result in only minor harm.

However, this strength of victim approaches also reveals a weakness. "The general problem with victim approaches is that they must require that any culpably wrongful act that inflicts much harm is necessarily evil, because no other factors, besides the amount of harm, are at all relevant to a judgment of evil."[102] Therefore, while focusing on the culpable harm done to victims is necessary for knowing whether an act is evil, it is not the only relevant factor. The motives and intentions of the perpetrator are also relevant to judging an act to be evil. Hence, victim approaches to evil are inadequate.

Perpetrator approaches have the advantage of picking out the maliciousness or depravity of the evildoer. For example, Mary Midgley argues that evildoers lack motives that ordinarily stop people from committing evil acts.[103] As we saw above, Kant believes that radical evil consists in subverting the moral law to the maxim of self-love. But neither of these accounts can explain the many roots of evil. Midgley's approach, and those like hers, cannot explain evil

---

[102] Ibid., 221.

[103] Mary Midgley, *Wickedness: A Philosophical Essay* (London: Routledge & Kegan Paul, 1984), 22-24.

done from relatively ordinary motives. Evil acts are not always motivated by malicious or sadistic motives. Greed, thoughtlessness, ambition, and honor can all motivate evil acts. Therefore, the perpetrator of evil need not have some abnormal psychology. On the other hand, Kant's account, and those like his, cannot explain evil done from especially sadistic motives. Kant explicitly argues that one does not do evil for evil's sake. Given the previous explanation of how the various propensities to evil corrupt the good predispositions, Kant concludes,

> The depravity of human nature is therefore not to be named *malice*, if we take this word in the strict sense, namely a disposition (a subjective *principle* of maxims) to incorporate evil *qua evil* for incentive into one's maxim (since this *diabolical*), but should rather be named *perversity* of the heart…. An evil heart can coexist with a will which in the abstract is good. [104]

Because of the naturally good predispositions, a human agent cannot will to do evil for evil's sake. Such an agent would be what he calls diabolical, and for Kant, humans may be perverse or corrupt, but never diabolical.[105] Yet, we are all familiar with stories of people who committed evil acts for the sake of evil. Therefore, perpetrator approaches are inadequate because they cannot include the many roots of evil.

Furthermore, whether the emphasis is on acts done out of ordinary or extraordinary motives, perpetrator approaches cannot explain that we ordinarily do not consider an act to be evil unless a certain kind or degree of harm is also

---

[104] Kant, *Religion*…60.
[105] Although Kant does not think that it is possible for humans to be diabolical, he does think that agents can be. Satan is a diabolical agent. In this way he differs from the Greeks, who thought it was impossible for any agent to do evil for its own sake.

present. One can shoplift out of malice, but the deed is still not evil. "All perpetrator approaches must require that any act perpetrated in a particular fashion, no matter how small and trivial the harm inflicted, must be evil."[106] Hence, perpetrator approaches are inadequate because they ignore the significance of the harm done to the victim.

Formosa's analysis shows us that both victim and perpetrator approaches are inadequate on their own, and yet each presents relevant features of an evil act. We are still left with the bystander approach that says an act is evil given a certain kind of negative response by a hypothetical bystander. This approach claims that an act is evil if incomprehensible or horrific to bystanders; it completely excludes both the victim and the perpetrator. By excluding these essential components, it cannot adequately explain evil. If something about the victim and something about the perpetrator are both relevant to a theory of evil, then the bystander approach must be inadequate. Now we are in the position to formulate some general guidelines about an evil act.

Formosa's analysis provides us with three necessary features any theory of evil must have:

> First there must be a perpetrator component, which identifies what it is about the way evil is perpetrated that makes them deserving of our very strongest moral condemnations. Second, there must be an unjustifiability component, which identifies what it is about evil acts that make them morally unjustifiable. Third, there must be a

---

[106] Formosa., 223.

victim component, which identifies what it is about the amount of harm that evil acts inflict that makes them so morally abhorrent.[107] I rely on this account for a general definition of evil, but I make the following addition. Formosa states that the perpetrator need not directly intend harm. "It is enough that an evildoer acts in a way such that harm is a reasonable foreseeable consequence of the act."[108] I agree. This definition can account for an evildoer acting either intentionally or thoughtlessly. I add that because various motives, intentions, and cares can move one to do evil, the perpetrator element reduces to the claim that human agency is necessary for bringing about an evil act. [109] Thus, I define the perpetrator element as "brought about by human agency."

Formosa rightfully argues that a victim element is necessary in an act of evil. He says that the kind of harm that a victim suffers from an evil act must be life-wrecking, but he does not elaborate on what this means. What do we mean by "life" and what do we mean by "wreck"? I will discuss in the next chapter that Kekes believes that the harm of evil acts is that it creates obstacles to attaining a good life. This is probably correct, but is still vague. Death is certainly life-wrecking, but only if we define life merely biologically as the condition that separates animals and plants from inorganic matter. But not all death is evil, and not all evil harm includes death. Both Formosa's and Kekes's accounts implicitly include the idea that a good life is one in which one can pursue the goals one sets for oneself. To wreck this pursuit is to destroy or severely injure one's ability or

---

[107] Ibid., 223.
[108] Ibid., 225.
[109] In the next section I argue that an agent can be moved simply by what she cares about to bring about evil. This argument will further support the idea that an evildoer need not have the intention to do evil.

capacity to do so. Hence, a life-wrecking harm is a harm that destroys or

jeopardizes the person's ability to live a good life. The victim cannot without

severe difficulty pursue her conception of the good. I would also add that the

notion of "wreck" includes the idea that the effects of the harm are often

irreparable. One cannot compensate the harm done with a fungible good. For

example, when a store keeper loses some inventory to theft, she can be financially

compensated. When a victim loses her dignity or hope, she cannot be. She may in

time recover her dignity and hope, but the harm is irreparable. This irreparability

is what often jeopardizes an agent's ability to live a good life, and often strikes us

as an evil-making component of the act. When we examine the cases presented at

the beginning of this chapter, the harms were all brought about by human agency

and they all destroyed or greatly injured the victims' abilities to attain good lives.

Therefore, "irreparable life-wrecking harm" names the kind of harm brought

about by an evil act. In section 6 below, I add more content to this notion by

examining the vulnerability of the victim of evil.

Formosa's complete definition is:

> An evil act is an act of wrongdoing in which the perpetrator of that
> act is at least partly responsible for other individuals suffering what
> would at least normally be a life-wrecking or ending harm, and
> where in so acting we judge the perpetrator, in the light of all the
> relevant factors, to be deserving of our very strongest
> condemnations. The relevant factors include intention, motive,
> effect, degree of harm, and the perpetrator's situation and
> circumstances.[110]

---

[110] Ibid., 230.

This combination approach allows for many roots of evil and includes different elements that are relevant to an act's evilness. Now we have some essential elements of an evil act. But not all combination approaches define evil in the right way.

Marcus Singer also presents a combination theory of evil. He defines evil as:

> *Evil acts*…are acts that are horrendously wrong, that cause immense suffering and are done with an evil intention or from an evil motive, the intention or motive to do something horrendously wrong causing immense unwarranted suffering. And malevolence, the doing or willing of what is wrong because it is wrong is what malignant evil, evil in its most extreme form, consists in. If an action is thought of as so wrong or bad that one cannot conceive of oneself as performing it, or conceive of any reasonably decent person as doing it, then that action is evil.[111]

Singer has incorporated those features that Formosa has argued are necessary for an act to be evil. There is a perpetrator component because the act must issue from an evil intention or from an evil motive. There is a victim component because the act must cause immense unwarranted suffering. There is a bystander component because one cannot conceive of any reasonably decent person as doing it. Singer admits that there are gradations of evil such that different evil acts may be evil in slightly different ways, but he believes he has shed light upon the nature of evil.

---

[111] Marcus G. Singer, "The Concept of Evil." *Philosophy* 79, no. 308 (2004): 205.

There are, however, a number of problems with his definition. First, he argues that an evil act must issue from an evil intention or motive. In fact Singer explicitly states "[o]ne cannot do something evil by accident or through thoughtlessness."[112] But this claim is false. If an individual commits an act out of negligence, and the kind and degree of harm necessary for evil obtains, then the act might be evil.[113] As J.L. Austin points out we have differing standards of what we deem acceptable behavior depending on the circumstances surrounding the action.

> The extent of the supervision we exercise over the execution of any act can never be quite unlimited, and is expected to fall within fairly definite limits ("due care and attention") in the case of acts of some general kind, though of course we set very different limits in different cases. We may plead that we trod on the snail inadvertently: but not the baby—you ought to look where you are putting your great feet. Of course it was (really), if you like inadvertence: but that word constitutes a plea, which is not going to be allowed, because of standards.[114]

Particular circumstances give rise to specific standards of behavior. These standards dictate whether an individual's excuse is acceptable or not. If one walks into a room and trips and falls we might accept her appeal to clumsiness. If, however, she walks into a room filled with babies lying on the floor, we expect that she take due care in her steps. In such a situation an appeal to simple clumsiness or thoughtlessness would be inadequate to exculpate. The point is that

---

[112] Ibid., 190.
[113] As previously mentioned, in the next chapter I explain, based on Frankfurt's conception of cares, how it is possible for an agent to bring about evil without choosing to do so
[114] J.L. Austin, "A Plea for Excuses" in *Philosophical Papers* (Oxford: Oxford University Press, 1961), 142-143.

there may very well be circumstances where "I did it by accident" or "I wasn't thinking" not only do not excuse a perpetrator from blame, but are sufficient for an act to be evil. Therefore, malicious intent is not a necessary element of evil.

Singer also includes a bystander component that I think is unnecessary for an act to be evil. He says that a test for whether an act is evil is that one cannot conceive of a reasonably decent person as doing it. I understand Singer's intuition here. When most of us think of genocide, war rape, torture, as well as a host of other atrocious acts, we cannot conceive of how a reasonably decent person could participate in such acts. The problem with this criterion is that it is too subjective to be reliable as a definition for evil. Conceptions of "acts beyond the pale" of wrongdoing have changed throughout history. There are numerous examples: public torture used to be a common occurrence in Western countries that now shun it as cruel and evil.[115] It is doubtful that no one involved in these practices was reasonably decent. Also, the oppression of women and minorities was the status quo until relatively recently. It seems possible that the perpetrators of this oppression were "reasonably decent" despite their horrible treatment of others. Further, the interest in the moral status of animals has grown recently. If more and more people become convinced that animals should not be imprisoned and eaten, it does not follow that those who do not understand this now are not reasonably

---

[115] See also Friedrich Nietzsche, *On the Genealogy of Morality* trans. Maudemarie Clark and Alan J. Swensen (Indianapolis: Hackett Publishing Company Inc., 1998), especially the Second Treatise. Nietzsche's explanation of the genealogy of guilt and bad conscience relies upon the historical practice of accepted violence against others as a means to repay a debt. Making others suffer was a universal currency and a practice that preceded the invention of moral concepts such as guilt, duty, or conscience. Even if Nietzsche's conclusions are questionable, he bases them the historical fact that festivals, noble weddings, and other public events frequently included the spectacle of making someone suffer.

decent individuals. What these examples illustrate is that as our understanding of the moral community develops, and as we understand who deserves moral consideration and which acts should be condemned, we improve our ethical behavior. However, it does not follow that we are not "reasonably decent" now even if is true that we can be more ethical in the future.

Furthermore, reasonable decency is not an incorruptible trait. Just because someone might be considered reasonably decent, it does not follow that she is incapable of committing evil acts. In fact, ordinary decent people have throughout history and literature shown themselves to be more than willing to commit evil deeds. One of the tragedies of evil is that it does not belong exclusively to monsters. For example, most ordinary Germans did not resist when Hitler asked them to help "cleanse" their society of Jews. More recently, Serbs turned against their Bosnian neighbors without much opposition. Or, in *Heart of Darkness* Kurtz commits atrocious acts once freed from the social constraints of Victorian England. It did not take much for these otherwise "reasonably decent" individuals to turn into perpetrators of evil. One might argue that these individuals were never reasonably decent, but were so perhaps only in appearance. This argument only further supports my point that we cannot trust whom we judge as reasonably decent. Therefore, the test of whether one can conceive of a reasonably decent person doing an act, fails as a test for evil.

Singer says that he is only interested in the nature of what he calls "EVIL,"[116] and not particular acts of evil, or evils. However, I think it is precisely

---

[116] Singer, 186.

the particular acts of evil that can tell us something about the nature of evil. For example, it is one thing to understand the general claim that evil results in immense suffering or a life-wrecking harm. It is a more developed moral understanding, however, that can *see that* a particular kind of act will always have such miserable consequences. It follows from understanding what it means to be a person that the destruction or abuse of that which is necessary for personhood would constitute an evil act. Harry Frankfurt's conception of personhood allows us to name a particular act of evil.

5. Defining Evil through Frankfurt's Conception of the Person

In the preceding sections I have outlined some of the necessary elements of an evil act. I also appealed to Rachels's analysis of the different ways a being can be harmed to show that we may need to examine particular circumstances and specific capacities to decide whether an act is evil. I now present one particular evil act.

One way that an act can be evil is that it severely damages or destroys the capacity or capacities necessary for personhood. Frankfurt provides us with a conception of a person that points to what this capacity is. He begins his argument by noting that it is not helpful to look at what is unique about persons as prior theories of personhood have done. Rather, to understand what is significant about being a person we need to examine what is most important to us as persons. In Frankfurt's words:

> [t]he criteria for being a person do not serve primarily to
> distinguish the members of our species from the members of other

species. Rather, they are designed to capture those attributes which are the subject of our most human concern with ourselves and the source of what we regard as most important and most problematical in our lives.[117]

One essential difference between persons and other creatures is to be found in the structure of a person's will. Humans are not alone in forming desires or even in making decisions based upon deliberation. However, it does seem to be a peculiar characteristic of humans that they can form second-order desires. In addition to wanting to be moved to do a particular act, humans can also want, or not want, to have certain desires and motives. They are capable of wanting their will to be a certain way. Many animals have the capacity to form first-order desires, which are simply the desire to do or not do this or that. But humans also have the capacity to want or not want the desires they have.

An agent might want to act without the desire being the agent's will. Consider statements of the form "*A* wants to *X*" where "*X*" is some action. Statements of this form identify first-order desires. But the mere fact that an agent has a first-order desire does not mean that the agent is motivated to act on the desire. That is, the existence of a first-order desire does not necessitate that the desire plays a decisive role in what the agent actually does or tries to do. *A* wants to *X* even if *A* also has other desires that are stronger or more motivating, or when *A* wants to do something else instead. Only when *A* wants to *X* and the desire for *X* is moving *A* to do what she is actually doing, does the statement "*A* wants to *X*" identify *A*'s will. An agent's will is the desire by which she is motivated in some

action or the desire by which she will be motivated when or if she acts. "An agent's will, then, is identical with one or more of her first-order desires."[118] Thus, the notion of will as Frankfurt uses it is the notion of an effective desire.

Now let's consider statements of the form "*A* wants to want to *X*" which identify second-order desires. These kinds of desires can be either effective or not. To illustrate the latter possibility, Frankfurt relates the story of a physician researching narcotics addiction who wants to help his patients by knowing what it is like to desire the drug as an addict. It is a genuine desire insofar as he does not merely desire to want the drug, but rather to be moved to some extent to take the drug. That is, he wants to be moved by the desire to take the drug. However, it is entirely possible that he does not want this desire to be effective. He knows the dangers of addiction quite well, and does not want to actually become addicted. He wants to want to take it, but he wants not to take it. His second-order desire to take the drug does not entail that he has a first-order desire to take it.

Frankfurt admits that such an individual does not represent the ordinary case. He "stands at the margin of preciosity, and the fact that he wants to want to *X* is not pertinent to the identification of his will."[119] There is, however, the kind of situation such that the statement "*A* wants to want to *X*" does pertain to what *A* wants her will to be. Here we see the distinction between second-order desires and what Frankfurt calls "second-order volitions." In these situations the statement means that *A* wants the desire to *X* to be the desire that moves her effectively to act. She wants the desire to provide the motive for what she actually does. In such

---

[118] Ibid., 14.
[119] Ibid., 15.

cases, when we say that "*A* wants to want to *X*," then we also mean that she already has the desire to *X*. It is only by having the first-order desire to *X* that *A* can coherently have the second-order volition that *X* be her will.[120] Not only does *A* have certain desires, but she also desires to desire in a certain way and wants those desires to move her effectively to act. In other words, *A* wants her will to be a certain way. When *A* wants a certain desire to be her will, then she has a second-order volition. "[I]t is having second-order volitions, and not having second-order desires generally, that I regard as essential to being a person."[121]

There is a close relationship between the capacity for forming second-order volitions and another capacity that Frankfurt thinks is essential to being a person—freedom of the will. It is only because a person has second-order volitions that she is capable of enjoying or lacking freedom of the will. To understand what Frankfurt means by "freedom of the will" he contrasts it with other kinds of freedom that are sometimes confused with freedom of the will. Having a free will does not mean that one is free to do what one wants to do. This freedom is freedom of action. Freedom of the will is the freedom to want what one wants to want, or in other words, to will as one wants to will.

> When we ask whether a person's will is free we are not asking whether he is in a position to translate his first-order desires into actions. That is the question of whether he is free to do as he pleases. The question of the freedom of the will does not concern the relation between what he does and what he wants to do....the

---

[120] It is possible to have a second-order desire without a corresponding first-order desire. It is not possible to have a second-order volition without having a corresponding first-order desire. After all, the content of the second-order volition is a first-order desire.
[121] Ibid., 16.

question about the freedom of his will has to do with whether it is the will he wants to have. [122]

An agent enjoys freedom of the will when she conforms her will to her second-order volitions. That is, her will is free when she can have the will she wants to have. Frankfurt provides us with the example of the unwilling addict to illustrate an unfree will. This addict hates her addiction and always struggles unavailingly to resist its force. She tries to overcome the desire for the drug, but it proves too powerful to withstand. She has conflicting first-order desires, both to take the drug and to refrain from taking it. She also has a second-order volition to refrain from taking the drug. She wants this desire to be effective in her action. It is this desire she wants to constitute her will. However, she is in fact moved to act by another desire – the desire to take the drug. Therefore, she does not have the will she wants. Given the power of the desire to take the drug, she is not free to will as she wants to will.[123]

> The enjoyment of a free will means the satisfaction of certain desires – desires of the second or of higher orders – whereas its absence means their frustration. The satisfactions at stake are those which accrue to a person of whom it may be said his will is his own. The corresponding frustrations are those suffered by a person of whom it may be said that he is estranged from himself, or that he finds himself a helpless or passive bystander to the forces that move him.[124]

---

[122] Ibid., 20.
[123] It does not follow that she takes the drug unwillingly. She takes the drug willingly, but it is not the will she wants to have.
[124] Ibid., 22

The capacity to exercise freedom of the will, that is, the capacity to conform one's will to one's second-order volitions, is essential to being a person. Therefore, an individual whose capacity to do so is irrevocably damaged or destroyed is not a person. Because this capacity is what constitutes an individual's personhood, any attack on it is also an attempt to damage or destroy that individual's personhood. Recall that in the previous section I gave a general definition of evil and said that it must be brought about by human agency and must result in a life-wrecking and often irreparable harm. Hence, I define a particular kind of evil act as any agent-caused act that damages or destroys an individual's[125] personhood by damaging or destroying that individual's capacity to form second-order volitions, or capacity to conform one's will to one's second order volitions.

Let's now examine this definition more closely. I include "agent-caused"[126] in the definition of evil because there can be non-agent-caused damage or destruction to one's will that does not constitute evil. Although some participants in the discussion of the religious problem of evil include natural evil as a kind of evil, I do not do so here. Since nature is indifferent to human projects or to one's will, the natural damage or destruction to one's capacity to be a person is accidental. As I argued in section 4 of this chapter, any act of evil must have a perpetrator component, by which I mean it must be brought about by human agency. If, for example, Alzheimer's Disease destroys one's capacity to conform

---

[125] In the next chapter I show that one can do evil both to others and to oneself.
[126] As I showed earlier, there must be a perpetrator component in any sufficient account of evil. Because it is important to explain how agents might unintentionally do evil, the perpetrator component boils down to an event brought about by human agency.

one's will to one's second-order volitions, or even to form second-order volitions, it is certainly sad for that individual and her family, but this event is not evil.

It also does not seem sufficient for evil that a particular second-order volition is thwarted. In fact, the thwarting of one's attempt to conform one's will to one's second-order volitions occurs quite often. Legal and moral prohibitions are aimed precisely at stopping an individual from acting on a specific desire if the action resulting from the desire is against the law or immoral. By doing so, the individual's ability to effectively act from her desire is frustrated. This sort of frustration occurs quite regularly and, furthermore, is necessary for societal and interpersonal interaction. It is clearly not evil.

There is another kind of coercion directed at an agent's will that seemingly fits my definition of evil, is in fact not evil, and therefore might serve as an objection to it. There are two social institutions targeted at altering behavior by changing the individual's second-order volitions: the moral education of children and the behavioral correction of prisoners. Surely these two institutions are good, at least in theory. How a specific society chooses to educate its children or correct its prisoners may in fact be evil. My point here is that it need not be.

Both the moral education of children and the correction of prisoners are focused on altering or eliminating the second-order volitions of the child or prisoner which motivate actions that are immoral and/or illegal. When the little boy opens the birthday girl's presents and is sent to his room, the goal of the punishment is to show him that his agency has an upshot in the world and that he

is an appropriate candidate for the reactive attitudes.[127] These realizations are supposed to have the effect that the boy does not open the birthday girl's presents the following year. Even if the first-order desire arises within him to open the girl's presents, he has a second-order volition not to be the kind of person who opens little girl's presents. More generally, he realizes that all of his actions might have a moral component such that he does not want certain first-order desires that may harm others to be effective in his action. He wants some specific desires to be his will and not others; he wants his will to be a certain way.

With the behavioral correction of prisoners, it seems unrealistic that the goal is the moral correction of the individual, rather than an attempt to awaken the realization that time in prison is not worth satisfying whatever desires motivate illegal acts. This realization might be arrived at through rational reflection or through fear. Either way, if the reformed prisoner has the second-order volition not to be the kind of person who commits the act that would land him in prison, the correction of his behavior has been successful from a legal standpoint.

Here we have two examples where the will of the individual was subject to thwarting and alteration. Do these cases fit my definition of evil? I have defined evil as agent-caused damage or destruction of one's capacity to form second-order volitions or to conform one's will to one's second-order volitions. At first glance, one might think that the moral education of the child and the behavioral correction of the prisoner fit this definition. They are both prohibited to act on that desire they want to be effective in their action. Notice, however, that

---

[127] This example comes from John Martin Fischer and Mark Ravizza, *Responsibility and Control*, 208.

in both examples the capacity of the individual to conform her will to her second-order volitions is not under attack. Rather, only particular second-order volitions are targeted for correction or elimination. The latter is not sufficient for evil and is often good. This is not to say that there may not be some example where the targeting of an individual's particular second-order volition does make up part of the explanation of why the act was evil.[128] But the targeting on its own is not sufficient for evil. However, if the targeting included inhibiting the *capacity* of the individual to want what one wants to want, then this would be evil.

This definition of evil might help answer McKenna's worry, discussed in chapter 1, that French's account of losing innocence for the sake of responsibility may justify child abuse. French states that "there seems to be a moral obligation for the mature members of society to cause the end of innocence, to guide children through the passage to adulthood." [129] Since this passage requires a personal acquaintance with wrongdoing, McKenna worries that French has not said enough to ensure that his argument cannot be used as a license for child abuse.

I have explained above why the moral education of children, although aimed at the alteration of the child's second-order volition(s), does not necessarily constitute an evil act. It is because in education only a particular second-order volition is targeted for correction rather than her capacity to form second-order

---

[128] See Peter French "Torture" in *War and Moral Dissonance* (New York: Cambridge University Press, 2011) for an explanation of how torture is wrong because it targets the victims second-order volitions which constitute the individual's identity.
[129] Peter French, "Losing Innocence for the Sake of Responsibility" in *Responsibility Matters* (Lawrence, KS: University Press of Kansas, 1992).

volitions or her ability to conform her will to her second-order volitions.

Conversely, in child abuse we have empirical evidence that the trauma caused

often significantly damages or destroys the ability of the individual, both as a

child and later as an adult, to conform her will to her second-order volitions.

Many victims of abuse have long-term psychological and behavioral problems.

Several studies have shown that many adults who experienced abuse as children

exhibited symptoms of anxiety, depression, posttraumatic stress disorder, and

reactive attachment disorder among others.[130] Individuals who are depressed or

who have PTSD, do not want to be motivated to act by these disorders. That is

why they are classified as disorders. They interfere with the individual's ability to

be motivated by what she wants to be motivated by. These individuals share

something in common with Frankfurt's drug addict who hates her addiction. They

are unable to pursue those desires they truly prefer to have. Under my definition

of evil, child abuse then easily fits, but moral education does not. Now, the

significant damage and/or destruction to the individual's capacity to will as she

wants to will need not be the only reason that child abuse is evil. Surely it also has

to do with attacking the extreme vulnerability[131] of, or exploiting, the child. But it

explains one way in which it is evil. Namely, the capacity essential for

personhood has been damaged and maybe even destroyed.

---

[130] Child Welfare Information Gateway, "Long-Term Consequences of Child Abuse and Neglect," *U.S. Department of Health and Human Services: Administration for Children and Families*, 2008, http://www.childwelfare.gov/pubs/factsheets/long_term_consequences.cfm#psych, (accessed March 12, 2012).

[131] In the next section I show that extreme vulnerability of the victim is always a feature of an evil act.

One might object that there are situations in which it is evil to target for alteration or elimination a particular second-order volition. Frankfurt tells us that sometimes one cares about something, not so much, but in such a way, that it is impossible for her to refrain from a certain course of action.[132] In a case of volitional necessity, a person can do no other than *A*; she is unable to refrain from doing *A*. Conversely, in a case of unthinkability a person cannot do *A*; she is unable to perform *A*. In this latter case, she cannot do *A* because doing *A* is unthinkable for her. Unthinkability and volitional necessity are modes of necessity which set the limits of a person's will and thereby shape the boundaries of her volitional identity. Because the elements which determine both volitional necessity and unthinkability are outside of a person's direct control, they constitute the stable volitional nature of the person. Since these volitional limits are important to the identity of the individual, if they were altered or eliminated, then the identity of the individual would also be altered or eliminated. Would this be an evil act?

There may be cases where it would be evil. But again this act would not always be sufficient for evil. Imagine that the target of alteration or elimination was the second-order volition of a Nazi who cared about exterminating Jews in such a way that he could not refrain from doing so. There does not seem to by any moral objection to going through with the alteration or elimination. In fact, the Allies instituted a "re-education" program in Germany after the end of WWII for precisely this purpose. If there were still dedicated Nazis who survived the war

---

[132] Harry Frankfurt, "The Importance of What We Care About" in *The Importance of What We Care About* (New York: Cambridge University Press, 1998).

and avoided capture, then the attempt at their re-education was a good thing. Therefore, even if the second-order volitions targeted for alteration and elimination are ones engendered by the agent's deep care, this act is not alone sufficient to define evil.

I do not intend my definition of evil to be exhaustive. There are other examples of wrongdoing that we would call "evil" even if they do not involve the damage and/or destruction of one's capacity to will as one wants to will. There may be physical harm that is evil, and there is certainly harm to animals that is evil. But the present definition explains what is always evil, even if there are other kinds of evil as well.[133] This present account of evil is important for persons because the capacity to have a free will, understood in the special way as the ability to want what one wants to want, is essential to being a person. In other words, evil is that which "de-personifies" the agent.

Finally, Frankfurt's conception of the will can help us understand the distinction between evil acts and evil character. Any moral agent can have a first-order desire to cause harm to another. If the harm damages or destroys the victim's capacity to want what she wants to want, then it is evil. When an agent commits an evil act because she has a first-order desire to cause the victim harm, the act itself does not tell us anything about the perpetrator's character. However, when the perpetrator has a second-order volition to damage or destroy the victim's capacity to will as she wants to will, then the perpetrator has an evil

---

[133] As noted earlier, Rachels argues that we explain different kinds of harm based on the victim's ability to be harmed. Similarly, I think we can explain the different ways acts are evil based on how an individual capacities and interests.

character. In such cases, the perpetrator wants the desire to harm the victim to be her will, that is effectively motivate her action. The perpetrator's desire to cause harm identifies her will and, consequently, identifies her as a person, specifically an evil person.

This particular act of evil, namely agent-caused damage and/or destruction of one's capacity to will as one wants to will, and the more general guidelines discussed in the preceding section, namely immense suffering on the part of the victim and some relevant feature of the perpetrator's motives or intentions, share a common characteristic: they both require that the victim be in a state of extreme vulnerability. I now show that extreme vulnerability is indeed present in cases of evil. Furthermore, I argue that the acquaintance with evil necessary for the shattering of moral illusions is the recognition that one has the ability to exploit, either directly or indirectly, the extreme vulnerability of others.

6. Extreme Vulnerability

I have presented both general characteristics of an evil act and named a particular kind of evil act. However, when I think of the various acts that philosophers call evil, viz. genocide, war rape, torture, the molesting and murdering of children, just to name a few, there is a shared characteristic among the victims that is not mentioned in the analyses of evil above. This shared characteristic is the extreme vulnerability of the victim of an evil act. Card and Formosa believe that immense suffering on the part of the victim is a necessary component of an evil act. I do not doubt their claim. However, what often makes

119

immense suffering possible is extreme vulnerability.[134] Understanding one's

ability to exploit another's extreme vulnerability is a personal acquaintance with

evil. This understanding shatters the illusion that all harm can be adequately

described by "wrongdoing," and leads the way to realizing that non-malicious

agents can bring about evil.[135]

Robert Goodin defines vulnerability as that which "amounts to one

person's having the capacity to produce consequences that matter to another."[136]

Ruth Sample describes this vulnerability as "a relationship between two or more

people, at least one of whom is in a position of causal relevance to the welfare of

the other(s)."[137] Understood in this way, being vulnerable is quite an ordinary

component of relationships and social interaction. We are all capable of producing

consequences that matter to another, and others can produce consequences that

matter to us. Therefore, being vulnerable in itself clearly does not entail being

harmed or being wronged, and certainly not being a victim of evil. But it does

mean that one is in the position of possibly being harmed by another. Being

vulnerable is necessary but not sufficient both for being harmed and also for being

a victim of evil.

---

[134] In "Torture in Ordinary Circumstances" Card does include defenselessness of the victim as a factor that defines torture, and torture as that act that picks out the salient aspect of evil, namely intolerable harm. She goes on to say that defenselessness come in degrees. However, it is difficult to imagine degrees of defenselessness. Either one is defenseless or not. Thus, I think it is better to think of victims of evil in positions of extreme vulnerability.

[135] I discuss this realization in chapter 4.

[136] Robert E. Goodin, *Protecting the Vulnerable: A Reanalysis of Our Social Responsibilities* (Chicago: The University of Chicago Press, 1985), 114.

[137] Ruth Sample*, Exploitation: What is it and Why It's Wrong* (New York: Rowman & Littlefield, 2003), 37. I also draw on Sample's summary of Goodin's argument.

When one is vulnerable, one is dependent on others. Goodin argues that the dependency of others and our ability to help them alone obligates us to provide assistance. He challenges the thesis of voluntarism[138] which argues that responsibilities may only be "voluntarily self-assumed."[139] He contrasts such special moral responsibilities, which are created by specific people with respect to specific people, with general moral rights and duties which are universal and thus the same for everyone. For example, I have a general moral duty not to murder, and others have a general moral duty to refrain from murdering me. I have a special or particular moral duty to love my son. I do not have a duty to love all of the children in the world, nor do others have the duty love my son. We can distinguish between these two kinds of duties, and yet both are still morally binding. Goodin rejects the idea that these special obligations arise due to some choice or decision we have made in respect to another; rather they arise due to the fact that "those persons to whom we owe special obligations are dependent on us for something crucial to them."[140] He thinks so even in the case of one's children. He believes that one is obligated to love and care for one's children, not because one chose to have children, but because ordinarily the parent is in the best position to provide the love and care the children need. The children are dependent on the parents for parental love. It follows then that in situations involving special obligations, I may have done nothing prior to create the relationship of dependency, but if it exists, I am morally obligated to provide the

---

[138] Goodin does not mean the metaphysical view that the moral law is a function of God's will.
[139] Goodin, 13.
[140] Sample, 36.

assistance needed by the other person. The fact of dependency is enough to obligate.

This dependence can be both general and particular. General dependence imposes a negative obligation on others whereas particular dependence imposes a positive one. I am generally dependent upon others in the sense that I rely on others to refrain from harming me. I am vulnerable to the possibility that another may physically assault me as I walk down the street, break into my house at night while I sleep, or kidnap, torture, and murder me. This dependence is negative because I hope that others refrain from acting and harming my well-being. Notice that I do not necessarily have an active relationship with the others upon whom I am generally dependent. This general dependence corresponds to the general moral duties that Goodin argues are universal.

Dependence also has a particular form. A situation of particular dependence is one in which the dependent requires another person, not just to refrain from interfering, but to act in a particular and direct way in order to fulfill some need. The drowning swimmer is particularly dependent on the passerby because the passerby has the ability, and is the only one close enough, to save the life of the drowning swimmer. The drowning swimmer does not require non-interference, but direct aid. In cases of particular dependence the dependent and the other person have an active relationship. This relationship might be activated simply because the other (the non-dependent) has the ability and is in the proximity to offer the required aid. This ability and proximity activate a special

122

duty to provide assistance. Goodin says that we are obligated to "suspend the ordinary rules of behavior in dealing with those particularly vulnerable to us."[141]

I think Goodin's analysis of vulnerability can shed light on a feature present in acts of evil. The dependence in cases of evil comes out of *extreme* vulnerability. That is, the dependent is vulnerable to immense suffering or an irreparable life-wrecking harm. Sample gives us a good summary of Goodin's definition of extreme vulnerability. There are four elements: "an asymmetry of power, the possession by the more powerful person of what another needs, a monopoly by the more powerful person on the thing needed by the weaker person, and control of the needed object by the more powerful person."[142] Goodin's focus here is on why exploitation of the vulnerable is wrong.[143] I think his argument can be extrapolated to provide insight into evil acts as well. If the thing needed by the dependent is her own life or dignity, her capacity to will as she wants to will, that which is crucial to pursuing her life as she wishes, or the life or dignity of her loved ones, then we begin to see why the malicious or thoughtless control and destruction of that needed thing would constitute an evil act.

We can also think of extreme vulnerability simply in terms of the ability to be severely harmed. If immense suffering or an irreparable life-wrecking harm must be suffered for an act to be evil, what allows for the possibility of such a degree of harm is the victim's extreme vulnerability. The more vulnerable one is,

---

[141] Robert E. Goodin, "Exploiting a Situation and Exploiting a Person" in *Modern Theories of Exploitation*, ed. Andrew Reeve (Beverley Hills: Sage Publications, 1987),187.
[142] Sample, 31.
[143] See also Larry May, *War Crimes and Just War* (New York: Cambridge University Press, 2007), especially chapter 4 "Humane Treatment". May argues that compassion and mercy are duties when one person is vulnerable to another.

the greater the potential for a harm to be life-wrecking or result in immense suffering. It follows that the same act done to two different individuals may be evil in one case but not the other. For example, I often walk around my neighborhood with my wife and our infant son. If a stranger were to come up to us and punch me in the face, this act would be wrong, but not evil. I would feel pain, but not immense suffering or a life-wrecking harm. Even if the perpetrator acted maliciously, the harm element is missing to make his act an evil one. On the other hand, if the stranger were to come up to us and punch my infant son in the face, then the act is evil. The very same act done to two different victims is wrong in one case and evil in another. The difference is that my infant son is in a state of extreme vulnerability (due to his infancy). The extreme vulnerability allows for the harm done to him to result in immense suffering or take on life-wrecking proportions. In other words, extreme vulnerability is a necessary condition for the harm element necessary for an act to be evil. Because extreme vulnerability is an essential element of an evil act, then when one understands, or sees that, other individuals are in situations of extreme vulnerability and, moreover, that one has the ability to abuse or exploit this vulnerability either directly or indirectly, then one has had a personal acquaintance with evil.

There is another interesting feature of the vulnerability of victims of evil. In cases of evil the victim is particularly dependent, not just on anyone, but on the perpetrator. We might think it is a truism to say that to avoid harm the victim requires the perpetrator not to inflict harm on her. But there is an interesting facet of this relationship. Goodin tells us that an agent is morally obligated to help

124

those in need, when the person's need will not be satisfied without assistance from others, and the agent has the ability to provide the assistance. That is, the agent has a special obligation. In cases of evil, the victim is not only in need but is in a situation of extreme vulnerability. Therefore, the object needed is not merely desired, but required for survival, well-being, or personhood. The perpetrator stands in closest proximity to the dependent to provide assistance given the fact that it is the perpetrator's actions that will directly deprive the dependent of the object (perhaps her own life, dignity, personhood, etc.) that is crucial to her continued ability to pursue her life as she wishes. The perpetrator does not only withhold assistance, which alone is morally blameworthy, i.e. is wrong, but takes direct action to exploit the extreme vulnerability of the dependent in a malicious or thoughtless way that causes immense suffering or a life-wrecking harm, i.e. is evil. The perpetrator has not only the power and ability to aid but to harm and destroy as well. This power further obligates the perpetrator. The violation of this further obligation in a malicious or thoughtless manner that causes immense suffering or a life-wrecking harm is evil.

For example, it is obvious that Jews during the Holocaust were in an extremely vulnerable situation. They depended on the Nazis to allow them to live. The Jews had a particular dependence because they required the Nazis, that is, each and every individual Nazi, to take action to save them from immense suffering or death. Because the Nazis were causing the harm, they were in the best position to end it. Or, in the report of the girl who hitched a ride with the man who tortured, raped, and left her to die, she had a particular dependence on him

125

not to subject her to these experiences. Often it is the perpetrator who puts the victim into the state of extreme vulnerability making the perpetrator the only person upon whom the victim is particularly dependent. This relationship helps explain why such acts are evil. In situations of genocide, war rape, the molestation and torture of children, the perpetrator is the only person (or persons) who has the ability to come to the aid of the victim for the very reason that the perpetrator has forced the victim into the situation of extreme vulnerability and is directly causing the harm. One reason (but not the only one) that genocide is evil is that the victims are extremely vulnerable to the direct actions of the perpetrators. The perpetrators are morally obligated to aid the dependents because of the relationship of particular dependence. Yet, the perpetrators inflict the very immense suffering and life-wrecking harm only they could have prevented.

This particular and direct dependence is exploited by the perpetrator to commit an evil act. This theory explains why crimes against children are usually evil acts. Children are in a position of perpetual dependence. They are extremely vulnerable in ways that adults are not. But of course adults are often extremely vulnerable as well. On March 31, 1992 a Serbian paramilitary group slaughtered hundreds of Bosniak (Bosnian Muslims) civilians in the town of Bijeljina in northeastern Bosnia. This was the first day of the Bosnian war. The soldiers shot and killed fleeing civilians in the streets. The civilians, mostly women, children, and elderly men, were placed into this situation of extreme vulnerability by the soldiers. It is not merely that the soldiers were simply bypassing those in need and violated a Good Samaritan obligation. We might find such people morally

126

blameworthy for not supplying aid when they had the ability and proximity to do so, but we would not usually consider them to be evil.[144] One feature of the act that makes it evil is that the soldiers created the situation of extreme vulnerability, made the victims particularly and directly dependent on them, and then took positive action to exploit their extreme vulnerability in order to cause them immense suffering and life-wrecking harm.

7. Conclusion

An act is evil is when it is maliciously or thoughtlessly motivated by an agent who exploits someone who is in a situation of extreme vulnerability causing the victim immense suffering or an irreparable life-wrecking harm. We see this both through the analysis of evil through Frankfurt's concept of a person and through the analysis of Goodin's concept of vulnerability. This approach has the advantage of containing both a perpetrator and a victim component. It picks out something peculiar about all victims of evil, namely extreme vulnerability. Finally, it helps explain the acquaintance with evil necessary for a certain moral disillusion: the shattering of the illusion that the category of wrong sufficiently characterizes all of the degrees of badness. When I say that an individual must have a personal acquaintance with evil it would be absurd to claim that an individual is required to accept that she is directly capable of genocide, war rape, or torture. I do not want to argue that such an acquaintance is necessary for this moral disillusion. The acquaintance with evil need not entail the explicit

---

[144] There could be a situation where a passerby maliciously ignores another's desperate situation, and we would call the act evil. The account I give here does not present the only reason an act is evil. Rather, I explain one common reason an act is evil.

acknowledgment of such a concrete proposition. Rather, the acquaintance with evil is the recognition that one is either directly or indirectly capable of maliciously or thoughtlessly exploiting someone in a situation of extreme vulnerability causing them immense suffering or an irreparable life-wrecking harm. This recognition need not be explicit; it may be an inchoate intuition. Nevertheless, it is an awareness of the extreme vulnerability of others and one's direct or indirect ability to maliciously or thoughtlessly exploit it with the direst consequences. This awareness provides the means for taking responsibility for the evil one causes. As I mentioned at the beginning of chapter 2 this awareness need not make the agent morally good. Moral disillusion does not entail moral goodness; it entails moral understanding. This understanding gives one the ability raise ethical questions about moral goodness or the adequacy of certain moral principles and directives. In the next chapter I show that another phase of moral disillusion involves understanding that one can bring about unchosen evil and be held morally responsible for doing so.

CHAPTER 5

UNCHOSEN EVIL

1. Introduction

In the previous chapter I defined evil and asserted that we can do evil

directly or indirectly. An agent can do evil directly either intentionally or

inadvertently. When an agent does evil directly and intentionally, she is

malicious, or diabolical. Recognizing that some people are malicious is important

to accurately evaluating moral character. However, not all evil is malicious.

Moreover, if most of us are not malicious, then we might think that most of us do

not bring about evil. But, this is an illusion. Many of us do bring about evil even

though we think evil ought to be avoided. We can bring about evil inadvertently

or indirectly; we can bring about evil without directly choosing to do so.

In this chapter I address the illusion that non-malicious people do not

bring about evil. There are at least two ways in which non-malicious people do

evil. A person may find herself in certain circumstances such that she does

something dreadful. The Milgram experiments and the Stanford Prison

Experiment provide empirical evidence that many people will commit horrible

acts when they are told to or given permission to do so. But non-malicious evil

does not only occur under circumstances we might happen to find ourselves in. It

is also possible that we create the circumstances that lead to evil. These

circumstances can arise do to our characteristics or our cares. Because it is not

abnormal for non-malicious people to do evil, we are bound to fail if we resolve

never to do evil again. Our goal instead should be to shatter those illusions that may keep us from minimizing the evil we bring about.

In this chapter I will explain how non-malicious people do evil either by finding themselves in situations where they do dreadful things, or by creating those circumstances from their own characters and cares. These evildoers may hold the belief that evil ought to be avoided; nevertheless they do evil. I show that we are powerfully motivated to act by what we care about, not always by what we believe. Because evil can be done by non-malicious agents who hold the belief that evil is morally wrong, moral responsibility ought to track what we care about rather than what we believe.

2. Situations of Evil

The Milgram experiments[145] on obedience to authority were conducted by Yale psychologist Stanley Milgram beginning in 1961. Milgram set out to test the hypothesis that the millions of accomplices to murder during the Holocaust were following orders despite holding the belief that murder (and genocide) is morally wrong.[146] The results of the tests indicate that obedience to authority is a very powerful motivator even when that obedience will result in violating a moral belief.

---

[145] See Stanley Milgram, *Obedience to Authority* (New York: Harper & Row, 1969).
[146] However, I am uncertain that the Germans who obeyed commands to help exterminate Jews, Roma, homosexuals, the handicapped, and others would have so readily obeyed had the command been to exterminate other "Aryans". In the Milgram experiments the perpetrator did not know who the victim was. During the Holocaust, the perpetrators knew who the victims were, and readily participated in their torture and extermination. So, although the experiments do help show that people will do dreadful things in order to please authority figures, they do not confirm the hypothesis Milgram originally set out to test.

In the test, a volunteer subject was given the role of "teacher." Unbeknownst to the test subject, the role of the "learner" was played by an actor who was part of the psychologist's team conducting the test. The "teacher" and "learner" were separated into two rooms where they could still communicate, but no longer see each other. The "teacher" was told that her task was to teach certain word pairs to the "learner." Every time the "learner" made a mistake, the "teacher" was to administer an electric shock to the "learner." The voltage of the shock increased 15 volts after each mistake. The "teacher" was given a test shock to be made aware of what the "learner" would experience. In one version of the test, the "learner" told the "teacher" that she had a heart condition.

The "teacher" was given a list of word pairs. She would read the word pairs to the "learner" along with four possible answers. The "teacher" believed that for every wrong answer the "learner" was receiving a shock. From the room in which the "learner" was supposedly situated, the "teacher" heard pre-recorded sounds for each shock level. After administering several of the increasingly severe shocks, the "teacher" would hear banging on the wall coming from the "learner's" room, complaints about her heart condition, screams, and eventually only silence.

Once all responses from the "learner" had ceased, many of the people playing the role of the "teacher" wanted to stop the experiment. However, they were told to continue and assured that they would not be held responsible for what happened to the "learner." The experiment was ended either when the test subject refused to continue despite being prompted by increasingly authoritative commands, or after she had administered the highest level shock, 450 volts, three

131

times in succession. In the first set of experiments, 65% of the test subjects continued despite the "learner's" complaints and screams, and administered the highest voltage of shock. The experiment was repeated multiple times in different cultures and the results consistently showed that between 60-65% of test subjects were willing to complete the experiment.

In the Stanford prison experiment,[147] conducted by Stanford psychologist Philip Zimbardo in 1971, out of a group of seventy-five student volunteers, twelve were selected to play "prisoners" and twelve selected to play "prison guards." The basement of the psychology building at Stanford was converted into a mock prison. The "prisoners" were given uncomfortable prison clothing, ankle chains, and were assigned numbers. They were "arrested" in their homes, then brought to the prison where they were finger-printed and strip-searched. The "guards" were given uniforms and wooden batons and told that although they should not harm the prisoners, they could use fear tactics to communicate their status of authority.

A prison riot broke out on the second day. The "prisoners" blockaded their prison door and refused to follow the orders of the "guards." The "guards" used fire extinguishers on the "prisoners" to quell revolt. To avoid future revolts the "guards" instituted a roll call where "prisoners" had to count off by stating their prisoner number. Some "prisoners" were not allowed to urinate or defecate, mattresses were confiscated forcing many to sleep on the concrete floor, and some were forced to go nude as an act of humiliation. As the days went on, many of the "guards" became increasingly ruthless, and many of the "prisoners" showed signs

---

[147] See Philip Zimbardo, *The Lucifer Effect: Understanding How Good People Turn Evil* (New York: Random House, 2008).

of psychological distress. Although the experiment had been planned to continue over two weeks, Zimbardo shut it down after only six days. In this short time, about one third of the "guards" showed sadistic tendencies. Many of the "guards" were disappointed that the experiment ended early. Of the fifty outside observers who saw the "prison" over the six days, only one person objected on moral grounds to its horrendous conditions.

Both of these experiments provide empirical evidence that people can do dreadful things given certain circumstances. The Milgram experiments show us that people are quite willing to check their moral beliefs at the door when a figure of authority orders them to harm another. The Stanford prison experiment shows us that people can become cruel, brutal, and even sadistic when given authority supported by an institutionalized ideology. The test subjects who participated in these experiments were not malicious individuals. They gave no indication, prior to the experiments, of desiring to exploit the extreme vulnerability of others. They believed that it was morally wrong to do evil to others. Yet, in both cases, it did not take much prodding for them to become evildoers. These experiments provide strong evidence that, given the right conditions, people can easily become perpetrators of evil. To understand that such a possibility exists, one must give up the illusion that only malicious people do evil things. Non-malicious, even ordinarily good, people are capable of atrocious acts.

But it is not always the case that we happen to find ourselves in situations that move us to act in evil ways. We might also create such situations given our characteristics and what we care about.

3. Kekes on Unchosen Evil

The notion of unchosen[148] evil plays a major role in John Kekes's analysis of evil in *Facing Evil*. He shows how the individual's own characteristics are sometimes the source of the life-wrecking harm that keeps one from attaining a good life. The individual does not choose to have the characteristics that cause her harm, and is thus an agent of unchosen evil. When the agent realizes this fact about herself, the experience can be tragic. I will bolster Kekes's position by examining Frankfurt's conception of how cares motivate action often more powerfully than purely cognitive beliefs and deliberate choices. Bringing these approaches together allows us to see that moral disillusion is more than realizing that one is an appropriate candidate for the reactive attitudes. There are multiple phases of disillusion. One loses moral innocence by seeing that one can cause harm to others. Eventually, one might also see that one can exploit the extreme vulnerability of others and cause them life-wrecking harm. The more complex phase that I am presenting in this chapter includes seeing that not all evil is malicious. We can cause both others and oneself life–wrecking harm even without choosing to do so, and we can be held morally responsible for doing so. Recognizing this moral fact shatters the illusion that good people cannot do evil.

---

[148] The philosophical literature on the nature of choice is immense. However, Kekes uses "choice" as we ordinarily use it. When we choose to perform an action we decide to perform it in order to bring about a certain result. It follows that an "unchosen" action is one that we do not decide to perform in order to bring about a certain result. Unchosen evil is evil we bring about without deciding to bring it about. In the next section I follow Frankfurt's argument that we do not need to have alternative possibilities available to us in order to be held morally responsible for how we guide ourselves to act. In this sense, we can be said to "choose" even without deciding from among alternative possibilities.

Kekes's argument begins with the assertion that we all aspire to live good lives. He does not mean good in the sense of morally good, but in the sense that we all hope to pursue our own conception of the good whatever it may be. The problem is that our aspirations to live good lives are often frustrated by what Kekes calls "essential conditions of life."[149] By examining three tragic stories, viz. *Oedipus the King*, *King Lear*, and *Heart of Darkness*, he illustrates that the essential conditions of life are the contingency of life, the indifference of nature, and human destructiveness. One's capacity to live a good life is vulnerable to these conditions.

These conditions, however, are not purely external to human agency. They also manifest themselves through human characteristics. Because we cannot choose our characteristics, we are sources of unchosen evil. In the following section I will explain unchosen evil, not so much in terms of characteristics as Kekes does, but in terms of what one cares about. In this section, however, I follow Kekes's argument.

The obstacles that keep us from living good lives arise regardless of how morally decent or deserving we are. In fact, as I will show below, they can even arise out of cares that we would ordinarily consider morally good or praiseworthy. The temptation is to explain away these obstacles by positing a supernatural harmony to the world; in Kekes's words "by succumbing to the transcendental

---

[149] John Kekes, *Facing Evil* (Princeton, NJ: Princeton University Press, 1990), especially 21-27 and 66-83. I do not think of the following conditions as "essential" as Kekes does. It is sufficient for my account that these conditions are prevalent, or simply just possible. It is the acknowledgement of, and reflection upon, these conditions that comprises a more complex phase of moral disillusion.

temptation."[150] That is, I might think to myself that even if I have had bad fortune in life, as long as I remain morally pure of heart, I will get my just deserts in the afterlife. Or, one often hears people explain others' undeserved suffering by claiming that God has His reasons. What seems horrible to us is really part of a hidden harmony that we are incapable of understanding as finite beings. Or, I am not an agent of evil because I do not intend to exploit the extreme vulnerability of others. But such thoughts are illusions.

Although we are vulnerable to constitutive and circumstantial luck such that we may not be able to attain the life we want for ourselves, obstacles to the good life do not only originate from circumstances external to human agency. Kekes argues that these obstacles often arise due to one's own characteristics. I will later argue that many of the obstacles come from what we care about. To illustrate this vulnerability, Kekes considers three tragic situations.

The first is depicted in *Oedipus the King* by Sophocles. Before Oedipus was born it was foretold that he would kill his father and marry his mother. Although his parents arranged to have him killed, Oedipus avoided death and grew into adulthood. He believed himself to be the son of the king and queen of Corinth. Oedipus learned of the prophesy and left Corinth so that the foretold events would not obtain. However, it is precisely these reasonable actions that brought about the occurrence of the prophesied events. His self-imposed exile brought him to Thebes where his biological father ruled as king. Oedipus met a band of men along the road, was provoked into a fight, and killed them.

---

[150] Ibid., especially 27-30.

Unbeknownst to him, one of the men was his father. He entered the city and solved the riddle of the Sphinx liberating the city from its oppression. As a reward he was made king and took the widowed queen, his unknown mother, as his wife.

He unknowingly committed both parricide and incest, two acts that he and his society thought were deeply immoral. Yet, his actions throughout his life were as reasonable and decent as one could expect. He was only led to kill his father and marry his mother due to his devotion to avoid performing these very acts. After discovering the truth of what he had done, he famously blinded himself, his wife/mother committed suicide, and his children/siblings were disgraced. The effects of his actions were life-wrecking for himself, his mother, his father, and his children. Thus, he was an agent of evil. But he was not an intentional agent of evil. He was motivated by his cares; cares we would ordinarily consider praiseworthy. Although he was seemingly reasonable and virtuous, he led a tragic life.

This tragedy picks out what Kekes considers to be one of the essential conditions of life: contingency.

> We can take, then, as the suggestion of the play that human life is vulnerable to *contingency*. Our sensibility allows for contingency because it allows that there are vast areas of our lives in which we lack understanding and control…As a result we may be reasonable and decent, and we may still find ourselves forced to do evil that we abhor in circumstances we are not responsible for.[151]

---

[151] Kekes, 23.

Oedipus was an agent of evil despite the fact that he actively tried to avoid the very acts he eventually performed. Circumstances beyond his control arose such that his reasonableness and virtue led him into tragedy.

However, the mere fact that we are vulnerable to the vicissitudes of life is hardly a new insight. We are all painfully aware that we cannot always achieve what we intend to given circumstances we do not control. Kekes tries to tie the fact of contingency into Oedipus's character. If he had not been so assertive, he would not have admitted to killing his (at the time of the killing unknown) father and his parricide would have remained in the past. But I do not think his argument is entirely convincing, nor is it necessary. Moreover, Aristotle admits as much when he says that "one swallow does not make a summer,"[152] meaning that we cannot judge a person to be flourishing until her life is complete. Thus, being a victim of bad luck is hardly sufficient to explain unchosen evil for which we are morally responsible.

The main part of my argument has to wait until the next section where I discuss the importance of what we care about. However, I can already say now that what makes the story of Oedipus significant for my present purposes is that he is brought to do evil by what he cares about. His hope for a good life is not destroyed by another exploiting his extreme vulnerability. Rather, it is destroyed by his own cares. He was motivated to act because he cared about acting virtuously; specifically, avoiding incest and parricide. Yet, these cares led to commit the very acts that directly opposed what he cared about. Not only was he

---

[152] Aristotle, *Nicomachean Ethics*, 1098a18-19.

an agent of unchosen evil, but the evil he did was to himself. He realized that his own cares brought about a life-wrecking harm visited upon himself. It is this tragic realization that moved him to blind himself. He learned that reasonableness and virtue do not always produce a good life. Although the lesson was tragic, he learned something about the nature of moral life; we may cause evil without choosing to do so. Recognition of this possibility is a moral disillusion. In the next two sections, I fill out this phase of disillusion by explaining both how our cares can motivate unchosen acts, and also how we can be held morally responsible for performing them.

The second essential condition of life is indifference and is depicted in *King Lear*. Lear divided his kingdom among his two wicked daughters, and disowned Cordelia, who loved him. He aced foolishly and came to pay for his foolishness though his own suffering. He learned his lesson and took responsibility for the events that occurred. He understood that his suffering came about due to his actions. After losing everything he began to cultivate those virtues we ordinarily consider praiseworthy: pity, compassion, and remorse for what he had done. He was reunited with Cordelia who had forgiven him. But we do not get the happy ending that we might expect. Cordelia was executed and Lear died with a broken heart. "We learn that goodness may be punished, that suffering and moral growth need not be compensated for, and that people come to undeserved harm. This situation is caused by the *indifference* of the scheme of things towards human merit."[153]

---

[153] Kekes, 23.

The story of Lear illustrates that there is no cosmic justice. Good people may suffer and the wicked may flourish. Being good is not always its own reward and one can easily profit by acting unjustly. There is no moral harmony to the universe compensating those who suffered undeservedly. A virtuous life does not necessarily lead to a good life. The universe does not reward rationality or goodness. But the world is not evil. It is simply indifferent to human agency. Being moral may not end in reward at all.[154] The universe is indifferent to moral merit.

The third essential condition of life is human destructiveness. In *Heart of Darkness*, Kurtz traveled to Africa to civilize the natives. "He had a strong sense of moral and cultural values."[155] He intended to overcome their barbarism not by waging war, but through the excellence of his character and intelligence. He traveled to a distant outpost in the African jungle and lived there among what he considered to be primitive tribes. The natives come to regard him with awe such that he seemed a supernatural being to them. He was given such power that nothing stood in his way. There was nothing to restrain him. He ordered midnight rituals of sex, violence and cruelty. Finding himself there in a position of absolute power he discovered within himself the barbarism he had set out to conquer. "The horror! The horror!" was the realization that the heart of darkness was inside himself.

---

[154] One might argue that one is not morally praiseworthy at all if one acts in accordance with moral guidelines only because of an expectation of some reward. It seems that one is truly a moral person when one acts in accordance with morality simply because it is the right or good thing to do.

[155] Ibid., 24.

This story shows that destructiveness is not always explained by appealing to external factors acting upon us, but is often brought about by what moves us to act. It can be a source of evil and create an obstacle to good lives. Therefore, we are often our own obstacle to attaining a good life. This realization is tragic. But understanding this feature of moral life dissolves the illusion that moral purity is attainable.

The essential conditions of life are not responsible on their own for evil in the world. If they were, then we could blame the way the world is for the occurrence of evil. The essential conditions of life manifest themselves through human vices. According to Kekes, contingency, indifference, and destructiveness give rise to specific characteristics. Contingency often appears as insufficiency, indifference as expediency, and destructiveness as malevolence.

Kekes argues that insufficiency is due to an inadequate development of some capacity required for acceptable moral conduct. The lack may be cognitive, emotive, or volitional. If cognitive, then the insufficiency often appears as dogmatism. Dogmatists may have a strong commitment to moral principles but their principles are mistaken. They have not developed the required critical faculty to discover they are wrong. If emotive, then the insufficiency often takes the form of insensitivity. Such people feel contempt for victims of evil and are insensitive to the suffering of others. These people lack the emotive capacity to sympathize with their intimates and realize they are suffering. They cannot see that others are vulnerable where they are not. It is also unchosen since its source may be genetic, hormonal, or environmental. If volitional, then the insufficiency

appears as weakness. These people are too weak to do what they recognize they ought to do. They are incapable of making the required effort. In all cases the evil is characteristic, yet unchosen.

Expediency is due to pursuing the goods of life without regard for the evil that may result from one's pursuit. These agents concentrate on achieving their goals without care for those who get sacrificed along the way. Expediency can take the form of selfishness or fanaticism. When someone is selfish all that matters is one's own goal and everything else is an instrument for reaching it. When someone is a fanatic some external goal dominates her actions and she is fully committed to achieving the goal such that everything is a mere instrument in the pursuit. In both cases one is indifferent to the evil one causes.

Malevolence is the disposition to act contrary to what is good. The emotional source is ill will. The malevolent person rationalizes her deeds and perceives them as righteous or justified. Through malevolence, destructiveness finds expression.

Although I do not see the direct link between contingency as a fact of life and dogmatism, insensitivity, and weakness of will in human character, I think Kekes's general point is correct. Human character is partly influenced by the way the world is. The moral agent does not appear *ex nihilo* with the capacity to choose one action or another. Rather, the agent's character is formed by her tradition and social circumstances. For this reason, it is important for moral evaluation to look at the agent's character rather than merely her ability to choose one act over another. It is not that choice does not play an important role in moral

evaluation; rather, the claim is that it cannot play a *foundational* role. Tradition and education have as their object character. Social norms form one's character. People have commitments because they grow up in a tradition that teaches them how to be. They might adopt or reject that tradition, but they do not choose it initially. Therefore, the formation of character precedes making choices. People find themselves in a position to choose because they already have commitments to the alternatives they confront. This approach supports the conclusion that the domain of morality is wider than the domain of choice. I will explain in the following section that our commitments come about by virtue of what we care about.

The contingency of human existence, the moral indifference of nature, and the presence of destructiveness in human motivation sometimes give rise to tragic situations. Part of the tragedy is that overcoming these obstacles is not within our control, and that we create the obstacles often because of what we care about. Although the tragic figures mentioned above did not choose the evil they caused, each brought about its occurrence due to his character. Each story illustrates in a different way that we are often the source of unchosen evil. Recognizing this fact eradicates the false hope that the universe coincides with reason, or that happiness will reward moral merit. If one falsely and naïvely believes that cosmic justice will compensate the good and punish the evil, one is blind to certain moral facts. One of these moral facts is that we may bring about unchosen acts of evil.

In this section I have established that unchosen evil is an undeniable feature of moral life. In the next section I explain how it is possible for an agent to

143

perform an action without choosing to do so. We will see that we are often

motivated to act based on what we care about and not what we believe. The

following section will then address how an agent can be held morally responsible

for actions that that she does not choose to perform. These pieces fit together to

reveal that an agent can bring about unchosen evil and be held morally

responsible for doing so. Awareness of this possibility shatters the illusion that

only malicious individuals bring about evil.

4.   The Importance of What We Care About

Our vulnerability to unchosen evil, to life-wrecking harm we visit upon

ourselves and others, is best understood by appealing to Frankfurt's conception of

cares. It is important to examine cares in addition to explicitly held beliefs

because cares often motivate without our choosing that they do so; that is, we

often have them without choosing them and, hence, they motivate unchosenly.

Beliefs, on the other hand, are held for reasons.[156] The content of our beliefs have

a truth value; they are propositions that we hold to be true or false. Because we

hold beliefs based on reasons, we can either accept or reject them based on either

good or insufficient reasons. When a belief motivates me to act, it usually does so

because I have already accepted the belief to be true. More generally, I can reflect

upon my beliefs about the world, retain those which are coherent and have

grounds, and discard those which are unfounded or result in a conflict or

---

[156] I am using a fairly loose definition of "belief" here. I remain neutral regarding the
disagreements between defenders of epistemic foundationalism and coherentism concerning the
architecture of knowledge. It seems uncontroversial that beliefs are cognitive and can be made
explicit upon self-reflection.

contradiction.[157] But cares are held not only cognitively, but affectively and volitionally. They can motivate us to act with or without our explicit cognitive consent. Moreover, in situations where what we care about conflicts with what we believe, the care often proves to be the more powerful motivator. Frankfurt gives us the example of a teenage mother who truly believes it would be best to give up her baby for adoption, but cannot bring herself to go through with it. What she cares about guides her action despite believing that an alternative action would be best. What we see in this example is that one's cares often guide an agent's action and show us what the agent's will really is. Because we are often guided to act by what we care about even when we hold a belief that it would be best to do otherwise, we can see how one might visit an unchosen evil upon oneself and others.

Frankfurt tells us that caring consists in guiding oneself along a distinctive course or in a particular manner. The agent's actions are thereby guided by one's cares.

> Caring, insofar as it consists in guiding oneself along a distinctive course or in a particular manner, presupposes both agency and self-consciousness. It is a matter of being active in a certain way, and the activity is essentially a reflexive one. This is not because the agent, in guiding his own behavior, necessarily does something *to* himself. Rather, it is more nearly because he purposefully does something *with* himself. [158]

What it means to care about something is to be guided in a certain way. But one is

---

[157] For a general discussion of the different kinds of grounds of belief see, Robert Audi, *Epistemology: A Contemporary Introduction to the Theory of Knowledge* (New York: Routledge, 1998), 7-8.

[158] Harry Frankfurt, *The Importance of What We Care About* , 83.

not guided through life like an empty raft on a river. Rather, one plays an active

role in how one is guided. The activity is reflexive insofar as the agent cares about

what (or who) she is. In other words, it is important to the agent what she cares

about. Frankfurt adds,

> [w]e are particularly concerned with our own motives. It matter
> greatly to us whether the desires by which we are moved to act as
> we do motivate us because we want them to be effective in moving
> us or whether they move us regardless of ourselves in or even
> despite ourselves.[159]

When we are moved to act by something we care about, we are not completely

passive. The cares move us to act because we want them to be effective in moving

us to act. We may not always choose which cares move us to act, but we reflect

upon the cares we have and have a second-order volition that they be effective in

our action.[160]

But caring about something should not be confused with liking or wanting

something or considering something to be of value. The difference between these

beliefs and attitudes and caring can be seen by examining their temporal

characteristics. Caring can easily be distinguished from wanting, liking, or

valuing something because "the outlook of a person who cares about something is

inherently prospective."[161] Someone who wants, likes, and values something can

---

[159] Ibid., 163.

[160] Frankfurt brings in his conception of identification to explain in what way the agent is active in her being moved to act by her desires and cares. I think that Frankfurt's argument is problematic, and further, I believe it is unnecessary to explain the phenomenon. Although there may be some endorsement on the part of the agent as she reflects upon her cares and has a second-order volition that certain cares are effective in moving her to act, this endorsement need not be wholehearted as Frankfurt argues, nor must it be cognitive as others have claimed. It can be affective or volitional.

[161] Frankfurt, 83.

do so in a temporally discrete moment without considering whether she has a future. Moments of wanting, liking, and valuing, can occur as isolated or discrete moments whereas the moments in the life of someone who cares about something are bound together by a "continuing concern with what he does with himself and with what goes on his life."[162] Frankfurt adds, "[c]onsiderations of a similar kind indicate that a person can care about something only over some more or less extended period of time."[163] Thus care can neither be momentary nor instantaneous. Because cares guide our actions and lives, they are much more intimately connected to who we are and what goes on in our lives.

For this reason cares are more important to who we are than our decisions and choices. The making of a decision takes only a moment. If one decides to care about something, it does not follow that the person does in fact care about it. It only follows that the person has formed the intention to care about it. Whether the intention gets fulfilled is another matter. Decisions indicate what an agent intends to be her will, but not what her will actually is. "If we consider that a person's will is that by which he moves himself, then what he cares about is far more germane to the character of his will than the decisions or choices he makes."[164] But cares are not more important than decisions only because the agent can change her mind after a moment, but because despite having decided on a certain course of action she may be unable to carry out her intention. She may not be able to bring herself to follow the course of action upon which she has decided.

---

[162] Ibid., 84.
[163] Ibid.
[164] Ibid.

The fact that someone cares about a certain thing is constituted by a complex set of cognitive, affective, and volitional dispositions and states. It may sometimes be possible for a person, by making a choice or decision, effectively to bring it about that he cares about a certain thing or that he cares about one thing more than another. But that depends upon conditions which do not always prevail. It certainly cannot be assumed that what a person cares about is generally under his immediate voluntary control.[165]

We often see cases where what one cares about conflicts with what one intends. Whether it is the man who has decided to leave his wife, but cannot bring himself to leave, or the teenage mother who has decided it would be best to give up her newborn baby for adoption but cannot bring herself to do so.[166] What each person cares about, and not only what one has decided would be best for her to do, tells us about the character of her will. It is possible then, that when an agent is guided by her cares she may act contrary to how she intends to act. Thus, we see that insofar as an agent is guided by what she cares about, it is possible that she acts unchosenly.[167] Moreover, if one's care-guided actions result in evil, to oneself or to others, then one has become a source of unchosen evil.

With this framework of cares in place, we can understand that unchosen evil can arise in two different ways. The first way is that one is guided by one's cares and evil results as a contingent consequence of one's actions. In such cases, the cares do not entail the performance of an evil act, rather evil occurs given the contingent circumstances. This kind of unchosen evil is depicted by the story of

---

[165] Ibid., 85.

[166] We tend to think of these kinds of dramatic examples, but they need not be.

[167] "Unchosenly" only in the ordinary sense of what it means to choose. For Frankfurt, because our cares guide us to act, we can be said to "choose" those actions.

Oedipus. He cared about avoiding parricide and incest, and he cared about taking responsibility for killing his father. Clearly, there is nothing essential about these cares that would entail any occurrences of evil. In fact, we would ordinarily consider such cares to be praiseworthy. And yet, evil occurred. Oedipus's cares guided him in such a way that he committed parricide and incest. Take another example. Imagine a new mother who cares immensely for her new baby. It is her first and, as first-time parents are wont to do, she is excessively worried about every aspect of his eating and sleeping. Every cry the baby utters sends the mother into a near panic worried that the baby is in desperate need of something she is not providing. She knows that she does not need to worry so much and repeatedly tells herself as much. In other words, she holds the belief that there is nothing wrong with the baby. Yet, she cannot help herself. She cannot bring herself to stop worrying. The constant worrying creates a tense atmosphere that the baby senses. The baby's eating and sleeping habits worsen because of the tension and the worry he senses. As his eating and sleeping become more irregular, the mother's worry intensifies. And the cycle continues. Ordinarily, we praise a caring mother and would castigate a mother who is indifferent to her baby's needs. Yet, in this case the mother's cares bring about a negative situation. If we add to the story that the baby's health deteriorates to the point of a life-wrecking injury or death, then the mother would be a source of unchosen evil. Her ordinarily praiseworthy cares guided her such that she brought about evil without choosing to do so.

The second way that one's cares can guide actions that bring about unchosen evil is when certain cares necessarily, or at least ordinarily, have an evil upshot in the world. These are situations where one cares about harming others, or cares about acquiring something for oneself such that one is indifferent to the vulnerability of others. This kind of unchosen evil is depicted by Kurtz in *Heart of Darkness*. He cared about having absolute power and about visiting violence upon others. What is interesting about Kurtz's situation is that he was unaware that these cares characterized his will until he was guided by them. His lack of awareness is what makes the story tragic. He realizes only after the violence and destruction had been done, that the source of the violence and destruction was his will. He discovers that he could not bring himself to treat the natives with kindness or respect. We can see that becoming aware that one actually cares about things that one believes are morally abhorrent, and that one has acted upon these cares despite holding the belief that such acts should be condemned, is a kind of tragedy. Caring about attaining absolute power and being violent towards others are cares that either necessarily, or at least ordinarily,[168] cause harm to others. Often this harm will be life-wrecking and, hence, evil. In such cases, one has not chosen to do evil. Before Kurtz journeyed to Africa he truly believed in the moral and cultural mores of Victorian England. Yet, he could not act upon his beliefs. He could not bring himself to act according to his beliefs, and instead acted under the guidance of his cares. By being motivated to act by what he cared about rather than what he believed, he became a source of unchosen evil.

---

[168] It does not matter for my purposes.

So far I have argued that unchosen evil is an undeniable feature of moral life, and on the basis of Frankfurt's conception of cares I have shown that unchosen evil can come about due to how we are guided by what we care about. But Frankfurt is not alone in seeing how our cares are often more powerful motivators than our beliefs. In her essay "Alief and Belief" Tamar Szabó Gendler explains that beliefs are not alone in guiding our action. Alongside beliefs, we have "aliefs" which guide our actions often in ways contrary to how our beliefs might guide us. For example, when some people visit Hulapai Skywalk, the 70 foot glass walkway extending outward from the rim of the Grand Canyon, they often cling to the side rail or to the security guard, not wanting to move to the center of the walkway or even to look down. But why else did they travel to the Grand Canyon, purchase the expensive ticket, take the shuttle to the entry point, and walk out onto the glass? They must have held the belief that it would be safe and enjoyable to do so. If they had even the slightest doubt about the safety of Skywalk, they would not have gone to the trouble to be there. And yet, once they walk out onto the glass, something else was going on alongside their belief that it was safe. They had the alief that it was not safe and that they should get off.

Gendler cites other examples provided by psychological studies. Test subjects were reluctant to drink a glass of juice in which a completely sterilized cockroach had been stirred, hesitant to wear a washed shirt that had been previously worn by someone they disliked, and disinclined to eat fudge that had been formed into the shape of dog feces. In each case, the test subjects believed that the juice was sterile, the shirt was clean, and the feces-shaped fudge had not

151

somehow changed its chemical composition.[169] However, they also held the alief that the items should be avoided. Clearly, these examples show that there is something alongside the content of beliefs that motivates how we act.

But the content of the alief does not necessarily replace the content of the belief. If we had asked the visitors to the Skywalk if they thought the glass platform was safe even as they clung in fear to the side, they would have probably replied in the affirmative.

> In each of the cases presented above, it seems clear what the subject believes: that the walkway is safe, that the substance is edible or potable…Ask the subject directly and she will show no hesitation in endorsing such claims as true. Ask her to bet, and this is where she will place her money. Ask her to think about what her other beliefs imply and this is what she will conclude…. At the same time, the belief fails to be accompanied by certain belief-appropriate behaviors and attitudes: something is awry.[170]

Perhaps the belief has been temporarily forgotten. However, if the visitors to the Skywalk truly thought the walkway was unsafe, they would probably do more than simply hold the attendant in fear or only reluctantly move toward the center. They would probably scream in terror! Such behavior seems quite strange if we hold a belief-desire-intention type approach to understanding action. However this behavior is not so strange when we realize that there is something going on alongside the belief that motivates behavior; namely, what one cares about. In the case of the Hulapai Skywalk, one's natural care to avoid danger is activated by

---

[169] Tamar Szabo Gendler, "Alief and Belief," *The Journal of Philosophy*, Volume CV, no. 10 (October 2008), 635-636.
[170] Ibid., 638.

the visual experience of looking beneath one's feet and seeing only empty space. The person experiences an alief that runs counter to her explicitly held belief. This sort of phenomenon is not rare. Therefore, any theory of action that excludes the possibility of aliefs ignores a fundamental aspect of human behavior. Because our behavior can also have moral import, it is crucial to moral disillusion to understand that one is guided by more than one's explicitly held beliefs. That is, not everything one does comes from a deliberate conclusion or even from a choice. We are often guided by primitive desires or what we care about. Just because we hold the morally right belief, it does not follow that we will perform the morally right action.

In "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction" Uriah Kriegel argues that moral judgments come in two varieties, moral aliefs and moral beliefs. Only the former are inherently motivating. The belief that one ought to perform some act, and the desire to perform that same act, could each exist in each other's absence. This modal separability shows that they do not make up a unitary mental state. To say that a belief is not inherently motivating is to say that there is no internal necessary connection between belief and motivation.

Kriegel shows that we have two complementary faculties. Thought is both rationalist and associationist; not just one or the other. We use the associationist mechanism for the purpose of everyday life and revert to the rationalist rule-based mechanism when greater accuracy is required. The former allows us to think and react quickly to our environment. It is unconscious, inflexible, but fast, efficient

153

and directly tied to action. We employ the latter when we encounter conflicts that we have to reason out. It is explicitly conscious, flexible, but slow and inefficient. Just as there is a duality of processes in our thinking, so too is there a duality of products. The associationist mechanism produces aliefs, whereas the rationalist mechanism produces beliefs. This dual process approach thus theorizes two ways in which we can explain our behavior. Some of our acts are motivated by what we alieve and some by what we believe.

This structure further explains the behavior of the agents in Gendler's psychological examples. Kriegel states it thusly: "an alief is a mental state whose occurrence causally explains our behavior in cases where our behavior does not match our beliefs, but may be operative as well when our behavior does match our beliefs."[171] The question remains in what way this distinction plays in our moral judgment.

It is quite a familiar occurrence that a person's honest moral proclamations are not mirrored by her moral practice. Kriegel refers to the psychological study[172] that shows that many people who would never assent to a racist proposition like "Black men are more dangerous than white men," nevertheless show traces of racist dispositions, a tendency registered in differential activation of the amygdala (associated with threat detection) in the presence of white and black faces. Although the moral belief is explicit and conscious, the behavior of

[171] Uriah Kriegel, "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction," *Australasian Journal of Philosophy* (forthcoming), 7

[172] Amodio, D., E. Harmon-Jones, and P. Devine, "Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report," *Journal of Personality and Social Psychology* 84, no. 4 (2003): 738–753.

the subjects is explained by their aliefs. Aliefs are inherently action-guiding. This inherent motivational aspect is what constitutes an alief. Moral beliefs, on the other hand, are characterized by their conscious accessibility and their dissociation from ongoing action. They are only indirectly connected to action. Beliefs may lead to action and may shape aliefs over the long-term, but they are not constitutively motivational states as are aliefs.

When moral aliefs and moral beliefs diverge, as in the implicit racist mentioned above, we have an individual who has not aligned her aliefs and beliefs. Either there is a lack of goodness or a weakness of will. The implicit racist either does not want to align her aliefs and beliefs or is not volitionally strong enough to do so. Certainly, it is conceivable that when an agent realizes she ought to perform some action, but lacks the motivation to do so, this lack of motivation can be explained by her not caring enough.

This framework shows why holding morally good beliefs is not sufficient to make one morally good. The amoralist can reason to the conclusion that one ought to perform such and such an act, that is, can reason to the right thing to do, but completely lack the motivation to do the right thing. The amoralist's moral beliefs are disconnected from her motivation, and probably from what she cares about. If by "practice" we mean how we behave, then moral beliefs are only one part of moral practice, and seemingly not the main part. Moral aliefs are more essential to moral practice than moral beliefs.

Moreover, the primacy of aliefs over beliefs in moral practice is reflected in our moral evaluation. We are usually quite bothered when the self-proclaimed

155

anti-racist shows racist tendencies in her behavior. In fact, we find her more

blameworthy for her racist behavior than praiseworthy for her anti-racist beliefs.

Conversely, we are often enamored with the self-proclaimed egoist whose actual

behavior reflects a concern for others. Assuming these individuals truly hold the

beliefs they say they do, our moral evaluation tracks how they behave, not what

they believe.[173] "When aliefs and beliefs diverge, the true moral character of a

person is reflected more accurately in her *aliefs*."[174] The reason for this is that real

moral commitments and cares are constitutively tied to the motivational states that

govern behavior.

Because an agent can reflect upon her aliefs and beliefs, notice if her aliefs

and beliefs diverge, and attempt to align her aliefs and beliefs, she can be held

morally responsible for what she alieves. She can either endorse or attempt to

alter the aliefs she recognizes. Moral agents do not exist in a discrete isolated

moment. They have a history; one in which they have the capacity to reflect upon

their behavior and alter it in future instances if it strikes them as inappropriate.

This capacity gives us guidance control over our behavior even if that behavior

emanates from aliefs or cares. Having guidance control over what one alieves

means that we can be held morally responsible for what one alieves. In chapter 6,

I investigate more fully the importance of reflecting upon what we alief and what

we care about.

One reason that aliefs are inherently motivational might be that they can

be produced by what the agent cares about, and cares are inherently motivational.

---

[173] This is not to say that beliefs have no moral importance. Only *lesser* importance than aliefs.
[174] Kriegel, 16.

Because the agent cares about x, she alieves y. It is not the case that because she believes x, she alieves y. The test subjects do not believe, but alieve, that the juice is contaminated because they care about staying healthy. The visitors to the Skywalk do not believe but alieve, that they are going to die because they care about survival. Aliefs do not respond to truth in the way that beliefs do. One can alieve in a certain way despite believing some truth that eliminates any reason for holding the alief. We should not, however, conclude that aliefs and beliefs are essentially opposed. It happens quite often that what one alieves overlaps with what one believes. However, when they are distinct, as in the above examples, the alief tends to motivate behavior despite holding a conflicting belief. One reason that aliefs are such powerful motivators of action might be that they are generated by what the agent cares about.

In this section I have relied on Frankfurt's account of the importance of what we care about to explain that cares might bring about aliefs and unchosen evil. Aliefs are not irrational, but may be arational. They do not always conflict with explicitly held beliefs, but they sometimes do. And when they do, they often guide our actions despite our holding a conflicting belief. When aliefs and cares motivate in ways that bring about evil, then the agent is a source of unchosen evil. Although the agent does not explicitly choose to act in the way she does, she cannot bring herself to do otherwise. In the next section I explain that despite the absence of a choice, one can still be held morally responsible for what one does. Moreover, acknowledging that one can bring about evil without choosing to do so, and yet still be morally responsible, comprises this phase of moral disillusion.

5. Moral Responsibility for Unchosen Evil

In the previous two sections I explained both that unchosen evil is an undeniable feature of moral life, and also that unchosen evil can come about due to the way in which our cares move us to act. In this section, I argue that the agent who acts unchosenly may be held morally responsible for her actions. To do so, I go back to Fischer and Ravizza's explanation of guidance control. Their description of this kind of control clarifies two important elements of my argument. First, it lends a structural explanation to the way in which our cares guide us. Second, it demonstrates how an agent can be held morally responsible while acting under guidance control. Merging Frankfurt's conception of cares with Fischer and Ravizza's analysis of guidance control will justify how an agent of unchosen evil may be held morally responsible.

Frankfurt attacks the Principle of Alternate Possibilities in "Alternate Possibilities and Moral Responsibility." In his example, Jones shoots Smith while Black stands waiting to interfere and ensure that Jones decides or wills to shoot Smith should Jones decide not to shoot Smith. Because Black is a counterfactual intervener, Jones lacks an alternate possibility to shooting Smith. Yet, he shoots Smith without any coercion or intervention by Black, and we hold him morally responsible for it. The example illustrates that an agent may be held morally responsible even if she lacks any alternate possibilities. We hold an agent morally responsible for her actions, not based upon the availability to choose among

various possibilities, but based upon the quality of her will at the time she performs her action.[175]

For Frankfurt, moral responsibility is essentially a time-slice notion.[176] It must be, he claims, because "[u]nderstanding what a person is, either as an entity of a certain generic type or as an individual, differs from understanding how he came to be that way."[177] Thus, one of the themes that connects the various essays collected in *The Importance of What We Care About* is that inquiry into who we are, rather than what has produced us, is the more philosophically interesting enterprise. Concentrating on a causal history of an agent rather than on the present structural features of the person's will cannot reveal what is essential to that person qua person.

In "The Problem of Action" Frankfurt explicitly attacks the causal approach to understanding the nature of action. The problem with appealing to causal histories, both to explain action and also to explain the difference between actions and mere happenings, is that casual theories suppose that actions and mere happenings differ only in their histories and not in themselves. Such an

---

[175] I realize that it is far from settled whether an agent must have alternative possibilities in order to be held morally responsible. Frankfurt's argument did, nonetheless, require libertarians who defend PAP to reformulate their position. For example, van Inwagen attempts to find a principle that retains alternative possibility but is immune to Frankfurt's attack. He formulates the Principle of Possible Action, Principle of Possible Prevention 1, and Principle of Possible Prevention 2. See Peter van Inwagen, *An Essay on Free Will* (Oxford: Oxford University Press, 1983), 162-182.

[176] I do not agree that moral responsibility is essentially a time-slice notion if by "essential" we mean at the exclusion of a historical or causal approach. I agree that moral responsibility is essentially a time-slice notion if by "essential" we mean that when evaluating someone's moral responsibility we take a time-slice analysis of the quality of the will as necessary to the evaluation. However, I think that moral responsibility also requires an historical account. I do not think that these two approaches are mutually exclusive. Fischer and Ravizza disagree with Frankfurt that a time-slice approach is appropriate for understanding moral responsibility. They favor a historical approach. See Fischer and Ravizza, *Responsibility and Control*, 170-206.

[177] Frankfurt, viii.

explanation is implausible since it directs attention away from the events at issue and from the times in which they occur. These theories tell us nothing about the agent since they look only to prior causal events. It is no wonder, according to Frankfurt, that such theories frequently run up against counterexamples.

To find out about the agent, whether our focus is understanding the nature of action, moral responsibility, or personhood, we must take a time-slice approach. This approach is necessary given the nature of what it is to be a person. A person's identity is constituted by her second-order volitions which are motivated by her cares. Those desires that a person wants to have and wants to be effective are those which comprise her will and her identity, and which move her act. Thus what is important is not how a person came to perform an action, but whether this action is presently under her guidance. Guidance does not require the ability to do otherwise as we saw in the case of Jones, Smith, and Black. Rather, there are two ways we can understand guidance. In Frankfurt's terms, as discussed in the previous section, we are guided when we have second-order volitions that certain desires and cares are effective in our action. In Fischer and Ravizza's terms, we have guidance control when we are moderate reasons-responsive. I now briefly repeat the argument I made in chapter 1 for why having guidance control, rather than choice among alternative possibilities, is necessary for ascribing moral responsibility to an agent.

In chapter 1, I unpacked the notion of guidance and how it relates to moral responsibility, but I will briefly repeat the argument. Inspired by Frankfurt-style cases, Fischer and Ravizza distinguish two different kinds of control, only one of

160

which is necessary for moral responsibility. Regulative control is the kind of control where the agent has alternative possibilities. But what is important for ascriptions of moral responsibility is not what the agent could do, but what she actually does. Therefore, we examine the actual scenario rather than looking to the presence of alternative scenarios.

The absence of regulative control does not entail that a morally responsible agent does not have any control over what she does. Although she may not have regulative control, i.e. alternative possibilities, she must have guidance control. "Guidance control of an action involves an agent's freely performing that action."[178] In a situation where an agent has guidance control, she freely does whatever she does. It does not matter whether she could have freely done something in an alternative sequence. One way to think of regulative control is as a dual guidance control. In these situations the agent has guidance control in both the actual sequence and the alternative one. But this kind of control is unnecessary for holding an agent morally responsible. Recall Fischer and Ravizza's example of Sally, who while driving a car, turns the car to the right. This car is a driver instruction automobile with dual controls for the driver instructor. If Sally had not turned or guided the car to the right, the instructor was ready to engage his controls to ensure the car turned to the right. Sally does not have regulative control because she could not have done otherwise than turn the car to the right. However, she does exercise guidance control. She freely turned the car to the right. This unusual case points out something significant regarding

---

[178] Fischer and Ravizza, *Responsibility and Control*, 31.

action and control. One can have a certain sort of control without having the sort of control that involves alternative possibilities.

Having guidance control means being moderate reasons-responsive. Recall that an agent is moderate reasons-responsive when her action-issuing mechanism is reasons-receptive in a strong way and at least reasons-reactive in a weak way. The agent must be able to respond in a regular fashion to reasons so that her responses and actions create a pattern or seem logical. The mechanism is reactive in weak way when it can react to a sufficient reason to do otherwise. It need not actually react to a sufficient reason, but if there is a possible world where it does, then it is weak reasons-reactive. It follows then that an agent need not have alternative possibilities available to her. There must only be some possible world where her action-issuing mechanism reacts to a sufficient reason to do otherwise. When an agent is able to recognize a sufficient reason to do otherwise, and there is some possible world where she acts otherwise based on this sufficient reason, then the agent has guidance control over her actions and can be held morally responsible for them.

What both the Frankfurt-style cases and Fischer and Ravizza point out is that the traditional association of moral responsibility with control is correct. But the relevant sort of control does not involve alternative possibilities. The sort of control associated with moral responsibility is guidance control. When an agent is guided by her cares, she exhibits guidance control even if she could not have volitionally brought herself to do otherwise. Therefore, the agent is morally responsible for her actions even when she does not choose to be guided by her

162

cares from among various alternative possibilities. One can still be moderate reasons-responsive even when being guided by what one cares about. As I stated above, what one cares about helps constitutes one's will. Hence, when one is guided by one's cares, one acts according to one's will. That is, one acts as one really wants to act even if it is unchosen. Therefore, an agent can be held morally responsible for acts of unchosen evil.

For these reasons, one can also be held morally responsible even when acting under volitional necessity. There are occasions when a person realizes that what she cares about matters to her not merely so much, but in such a way, that it is impossible for her to pursue or to refrain from a certain course of action. These cares motivate second-order volitions in the agent and sometimes exhibit a peculiar kind of necessity "in virtue of which his caring is not altogether under his own control."[179] Frankfurt names these two kinds of necessity volitional necessity and its counterpart unthinkability. In a case of volitional necessity, a person can do no other than *A*; she is unable to refrain from doing *A*. Conversely, in a case of unthinkability a person cannot do *A*; she is unable to perform *A*. In this latter case, she cannot do *A* because doing *A* is unthinkable for her. Unthinkability, like volitional necessity, is a mode of necessity which sets the limits of a person's will and thereby shapes the boundaries of her volitional identity. Because the elements which determine both volitional necessity and unthinkability are outside a person's direct control, they constitute the stable volitional nature of the person. These situations differ from those in which a person feels an external compulsion

---

[179] Frankfurt, 86.

too powerful to resist. It is not that the person lacks the power to refrain from acting in a particular way. Rather, volitional necessity is such that the person lacks the will to act in a particular way.

Volitional necessity prevents the person from organizing her will in the appropriate way to will other than she does. Of course the person has the physical capacity to act otherwise, but she cannot *bring herself* to do so. "Not only does he care about following the particular course of action which he is constrained to follow. He also cares about caring about it."[180] She guides herself away from being affected by anything that might dissuade her from following the course of action or from caring as much as she does about following it. She cannot bring herself to do otherwise because she does not want to. The necessity is generated when someone requires herself to avoid being guided in what she does by any forces other than those by which she most deeply wants to be guided. Despite the force of the necessity, the agent is still moderate reasons-responsive, that is, still exhibits guidance control and wills as she wants to will, and is therefore morally responsible for her actions.[181]

6. Conclusion

In this chapter I have argued that unchosen evil is an undeniable feature of moral life. Malicious individuals are not the only agents who bring about evil. I have also shown how our cares guide our actions, and that they often do so

---

[180] Ibid., 87.
[181] For a discussion of how an agent is morally responsible even when moved to act by the volitional necessity motivated by collective memories, see Peter A. French, "The Moral Challenge of Collective Memories" in *War and Moral Dissonance* (New York: Cambridge University Press, 2011).

without our making any choice. It follows then that our cares may guide us to bring about unchosen evil. Finally, I have shown how merging Frankfurt's conception of cares with the notion of guidance control explains that we can still be held morally responsible for acts we do not necessarily choose to perform. It follows then that we can be held morally responsible for bringing about unchosen evil. I have brought these elements together to claim that when one acknowledges these fundamental elements of moral life, one has shattered certain moral illusions. One has cast aside the naïve hope in a cosmic reckoning based on moral worth, and in the possibility of moral purity. More importantly, one understands that responsibility can be attributed, not only to actions one chooses to do, but for the evil one does unchosenly. In order to take responsibility for, and potentially minimize, evil in the world it is important that one recognizes how one can be a non-malicious unintentional agent of evil. That is, given this understanding one can ask questions about the adequacy of moral principles and directives, take responsibility for the evil one brings about, and attempt to minimize said evil.[182]

One might wonder if the agent must experience unchosen evil in the same way that the innocent must confront wrongdoing experientially rather than merely conceptually. Recall that in chapter 2 I explained that the moral imagination and the exercise of hypothetical thinking are involved in these more complex phases of moral disillusion. One can shatter illusions by imagining what it would be like for her cares to bring about evil even if she has never experienced it firsthand. One can understand that one can be an unchosen source of evil even if one has

---

[182] I examine how one can minimize evil in chapter 6.

somehow avoided bringing about unchosen evil. Therefore, although the first phase of moral disillusion requires an acquaintance with wrongdoing, these more complex phases can be experienced through the exercise of the moral imagination and hypothetical thinking. One sees that unchosen evil is an undeniable feature of moral life and that one can be morally responsible for it.

As I mentioned throughout this chapter, understanding that one is a source of unchosen evil is often tragic. One realizes that one's cares can visit evil on others and upon oneself. Understanding moral tragedy comprises the second part of this more complex phase of moral disillusion. In the next chapter, I turn to examine the nature of moral tragedy in more detail.

CHAPTER 6

MORAL TRAGEDY

1. Introduction

In the previous chapter I discussed unchosen evil. I argued both that it is

an undeniable feature of moral life and also that awareness of oneself as a

potential source of unchosen evil is an essential component of moral disillusion.

The unchosen evil one brings about is often tragic. It is tragic precisely because

one has brought it about without choosing to do so. Moral tragedy is a corollary to

unchosen evil. That is, one aspect of this disillusion is understanding that one can

bring about unchosen evil. Another aspect of it is to understand the tragic nature

both of the evil one causes and the realization that one can bring it about

unchosenly. Understanding that moral tragedy is an undeniable feature of moral

life helps comprise another moral disillusion. To deny that moral tragedy can

arise is to deny a feature of moral life. I will show in this chapter that it is to deny

one's moral responsibilities to other persons. In order to come to this conclusion I

investigate the nature of tragedy. First, I examine Aristotle's conception of

tragedy, but find it unsatisfactory because it cannot account for an agent

knowingly bringing the tragedy about. Then, I rely on W.H. Auden's distinction

between the Greek tragedy of necessity and the Elizabethan tragedy of possibility

to set the stage for Alasdair's argument that the contemporary emphasis on the

right over the good has tragically alienated us from the good life. I find

MacIntyre's arguments problematic and thus unable to explain how we should

understand moral tragedy. Finally, I turn to Gowans's theory of inescapable moral

wrongdoing and his account of "responsibility to persons" to explain how moral

tragedy arises. Gowans argues that it can be tragic when one does something

wrong even when performing the morally best action. I argue that one shatters a

moral illusion by recognizing the possibility of moral tragedy and that it can be

brought about by one's own agency.

2.   Kinds of Tragedy

For Aristotle the goal of tragedy as literature is to arouse pity and fear in

the audience. The audience members should pity the character and fear that the

same events could happen to each of them. The manner by which audience

members feel pity and fear is by virtue of the story representing a universal

possibility. That is, what happens to the characters in a tragedy could happen to

anyone. Accordingly, the plot should portray a situation that is possible for each

of us to imagine ourselves in.

> From what we have said it will be seen that the poet's function is
> to describe, not the things that has happened, but a kind of thing
> that might happen, i.e. what is possible as being probable or
> necessary. The distinction between historian and poet is not in the
> one writing prose and the other verse—you might put the work of
> Herodotus into verse, and it would still be a species of history; it
> consists really in this, that the one describes the thing that has
> been, and the other a kind of thing that might be. Hence poetry is
> something more philosophic and of graver import than history,
> since its statements are of the nature rather of universals…. By a

universal statement I mean one as to what such or such a kind of
man will probably or necessarily say or do…[183]

Tragedies make universal statements about human life. As such they make

statements about situations that we all may find ourselves in. Where history

reports how a singular individual acted, tragedy exposes us to universal situations

that may befall any of us.

Pity and fear are most effectively aroused by a reversal of fortune. The

reversal of fortune often accompanies a discovery that moves the character out of

a state of ignorance into one of knowledge. This reversal comes about either due

to some character flaw or to fortune, but not because of depravity.

…pity is occasioned by undeserved misfortune, and fear by that of
one like ourselves…. [T]he change in the subject's fortunes must
be not from bad fortune to good, but on the contrary from good to
bad; and the cause of it must lie not in any depravity, but in some
great fault on his part…[184]

The reason that pity and fear may be aroused in the audience members is that

there is nothing idiosyncratic about the tragic figures. Furthermore, the character

flaw that brings about the tragedy is not malice or depravity. As we saw with

Oedipus it was his desire to avoid parricide and incest, accompanied by an

ordinary amount of assertiveness and pride (pride being a virtue for Aristotle, but

not for later Christians) that gave rise to his reversal of fortune. Oedipus and other

tragic figures do not bring about their tragic downfall by willingly doing evil. The

reason for this is that willingly doing evil is impossible for Aristotle. If such a

---

[183] Aristotle, "Poetics" in *The Complete Works of Aristotle Volume Two*, ed. Jonathan Barnes
(Princeton: Princeton University Press, 1984), 1451a37-1451b9.
[184] Ibid., 1453a4-1453a16.

situation were portrayed in tragedies, there could be no arousal of pity and fear because no one could relate to such an impossibility. The plot of the tragedy must be a universal situation. As with Oedipus, it is ignorance and misfortune accompanied by common character flaws that give rise to tragedy.

Although Aristotle offers us considerable insight into the nature of tragedy, he misses that tragedy may arise due to acts one performs willingly. As I discussed in the previous chapter, although one may perform an action unchosenly, she may still perform it willingly. Moreover, as I discussed in chapter 3, there are various roots of evil including performing an evil act intentionally. In section 4 of this chapter I will argue that we often give rise to tragedy by willingly fulfilling one responsibility at the expense of another. Finally, Aristotle's conception of tragedy is unsatisfying because it allows for the universe to be a player rather than simply a venue. One of the lessons we learn from *King Lear* is that the universe is indifferent to human agency.

King Lear has a different manner of portraying the tragic situation or tragic character. W.H. Auden notes that what characterizes the Elizabethan tragedies are failed opportunities to make a choice or choices that would avert the tragedy. This emphasis on choice is very different from the tragedy of necessity that characterizes Greek tragedy. In the Greek tragedies, the characters cannot avoid their fate no matter what they do. In Elizabethan tragedies the character can always avert the tragedy, but never seizes the opportunity to do so.

> In a Greek tragedy everything that could have been otherwise has already happened before the play begins. It is true that sometimes

170

the chorus may warn the hero against a course of action, but it is unthinkable that he should listen to them, for a Greek hero is what he is and cannot change…. But in an Elizabethan tragedy, in Othello for example, there is no point before he actually murders Desdemona when it would have been impossible for him to control his jealousy, discover the truth, and convert the tragedy into a comedy.[185]

We can think of these two kinds of tragedy as tragedy of necessity and tragedy of possibility. In Greek tragedy, the characters cannot avoid the tragedy no matter what they do. In Elizabethan tragedy the characters have opportunities to avert the tragic conclusion. The tragic conclusion is not inevitable because they are avoidable up to a certain point. The tragedy lies in lost opportunity, rather than in necessity.

We see this distinction in *Oedipus the King* and *King Lear*. Oedipus tried to avoid the prophesied events, but everything he did actually brought about the very state of affairs he sought to avoid. His was a tragedy of necessity. One of the reasons we pity Oedipus is because he could not avoid his fate. Lear, on the other hand, has to make choices throughout the play and always has the opportunity to make the choice that will lead him away from a tragic end. He could always make a better choice, but seals his own fate. We pity Lear because he could have avoided tragedy if only he had made better choices. Although the good-making aspects of life were under his control, we learn from the story that the world is indifferent to moral worth.

---

[185] W.H. Auden, *The Dyer's Hand and Other Essays* (New York: Random House, 1954), 174-175.

The Greek tragedies arise out of an emphasis on character whereas the Elizabethan tragedies focus on choice. These emphases are reflected in the philosophies of the times as well. Aristotle defines virtue and vice through an examination of character, and Christian and Enlightenment philosophers make choice the foundation of morality, forgiveness, and redemption. These two views are not inconsistent, but have different emphases. Under "character-morality" our choices are formed by our characters. Under "choice-morality" our choices form our character. The question is whether character or choice should be the foundation of morality.[186] I will not debate the merits of choice and character morality here. What is significant for the present discussion is to see that there are different ways of understanding tragedy. The question remains how we should understand moral tragedy.

One way to understand the difference between the emphases of Greek and Elizabethan or Christian tragedy is due to a shift from an emphasis on the good to an emphasis on the right and rights. After Hobbes we see philosophers discussing right and rights more than the good. Alasdair MacIntyre argues that this shift in focus has brought about a tragic situation. By subsuming the good to the right, we have fragmented our system of moral values and reduced moral life to the following of rules. Individuals in contemporary society are consequently alienated from the good life. This alienation is a kind of tragedy. I examine MacIntyre's approach to tragedy, but find it problematic. I then turn to an account of moral tragedy that arises from conflicting responsibilities to other persons.

---

[186] See Kekes, *Facing Evil*, chapters 5-8.

3. MacIntyre and the Good Life

In *After Virtue* MacIntyre argues that individuals in contemporary society can no longer cultivate the Aristotelian virtues due to the existence of shallow practices that provide no way of unifying life. Without a return to the kind of connectedness between life and practices that characterized ancient and medieval village life, we are doomed to live episodically and alienated from the good life. The inability to achieve the good life is a kind of tragedy. To understand how he arrives at this conclusion I turn to his analysis of virtue through an examination of practices.

In order to motivate his inquiry into virtues, MacIntyre explains why such an endeavor is required. He notes that philosophers engaged in debates concerning what is "good," "permissible," "obligatory," etc. assume that there is a well-demarcated subject matter for investigation. However, they fail to notice that such terms have significance only within specific historical contexts. How these terms are applied in specific societies depends on the cosmological background and historical context of the specific society. Therefore, to uncover what moral concepts mean one must look to the cosmology and history of the relevant society. When these concepts are extrapolated from their context, they lose their authority.[187] Only history can tell us why certain normative rules were in good order and then fell into disorder. It follows that MacIntyre's intellectual history is crucial to the success of his philosophical argument. Later in this section I will

---

[187] See also, G.E.M. Anscombe, "Modern Moral Philosophy," *Philosophy* 33, no. 124 (1958). She makes a similar argument attacking deontologists for using the word "ought" out of its theological contexts. She claims "ought" loses its meaning when ripped out of its original context.

focus on some of the misleading comments he makes about the history of the

Enlightenment, which will in turn cast doubt upon his philosophical conclusions.

But first, his argument.

MacIntyre claims that Nietzsche was the first philosopher to point out that

we do not know the history, and hence the meaning, of our moral concepts. He

showed us that appeals to objectivity were in fact expressions of subjective will.

Thus, the endeavors of Enlightenment thinkers to ground a rational morality in a

rational human nature were doomed to fail. Staying true to his prescription to

appeal to historical context for clarification, MacIntyre points out that the

Enlightenment project, which Nietzsche eventually pulls the rug out from under,

was only made possible by the rejection of the Aristotelian moral system. It was

only because they rejected the Aristotelian and Scholastic moral tradition that

Enlightenment thinkers needed to discover new rationalist foundations for

morality. The failure of the Enlightenment project, foreseen by Nietzsche, has

brought about a chaos of moral values in contemporary society. Robert Wokler

describes MacIntyre's position the following way:

> That project was in his view centrally concerned with providing
> universal standards by which to justify particular courses of action
> in every sphere of life, and although Enlightenment thinkers
> manifestly did not agree as to exactly which principles might be
> acceptable to rational persons, he claims they nevertheless
> collectively propagated the doctrine that such principles must
> exist…. Many post-Enlightenment philosophers have continued to
> pursue that aim, but in the absence of any prevalent framework of
> values within which moral judgments could be agreed, they have

only shown, according to MacIntyre, that this ideal cannot be attained…. Without already settled moral beliefs, we have come to identify our principles only in terms of abstract notions of the self and individual choice, freed from the contingencies of social roles or historical tradition.[188]

In such a society where individualism is deified, each may formulate her own set of values. We are lost in an alienated society without a cohesive set of values resulting in the fragmentation of the self. But Nietzsche's criticisms are directed only at the moral systems of the Enlightenment philosophers, not at what came before. Therefore, if the Aristotelian moral system can be resurrected and sustained, then Nietzsche's criticism of morality can be rejected. Hence, the key question becomes whether Aristotle's ethics be vindicated.

MacIntyre begins to address this key question by examining the nature of virtues. The core of any account of virtue is practices. He defines a practice as

> any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially derivative of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended.[189]

Internal goods are those goods which can only be acquired by participating in and excelling at a certain practice. They can only be identified and recognized by

[188] Robert Wokler, "Projecting the Enlightenment" in *After MacIntyre* ed. John Horton and Susan Mendus (Cambridge: Polity Press, 1994), 108.
[189] Alasdair MacIntyre, *After Virtue* (Notre Dame: University of Notre Dame Press, 1984), 187

engaging in the practice in question. The standards of excellence by which we judge excellent or poor performance in a practice develop historically via participation in the practice and are raised when performers in a practice excel. For example, by playing football one cultivates those skills peculiar to excelling in football. Those engaged in this practice judge the participant's performance based on the rules of football which have developed over time as the sport has been played. We also judge the player's performance based on how well she executes the skills (i.e. internal goods) necessary to, and peculiar to, being successful in the sport. As the sport (i.e. practice) develops over time, so too do the standards of excellence.

The good related to a practice can be both internal and external to the practice. The former can be acquired only through the specific practice in which they inhere. One acquires them by excelling at the particular practice. External goods are contingent to a given practice and can be acquired through many different practices. Wealth, power, and fame are obvious examples. One might acquire wealth and fame by excelling at playing football, but they do not allow one to excel at playing football nor do they belong essentially to the practice. They are goods of competition not goods we cultivate in order to excel in a specific practice.

With this conceptual framework in place MacIntyre explains the nature of virtues. A virtue is a quality which allows us to achieve those goods internal to practices.

> A virtue is an acquired human quality the possession and exercise of which tends to enable us to achieve those goods which are internal to practices and the lack of which effectively prevents us from achieving any such goods.[190]

There are three virtues necessary for any practice with internal goods and standards of excellence: justice, courage, and honesty. This means that regardless of the practice in question the participant needs these virtues to excel at the practice. Although different cultures define truthfulness, courage, and justice differently, they all define them in some way. MacIntyre concludes from this observation that these three virtues are goods without which practices cannot be sustained.

One might think that there cannot be evil practices if practices require virtues. But MacIntyre claims that we do not have to condone everything that flows from a virtue. It is possible that evil may be brought about by virtues.

> I do have to allow that courage sometimes sustains injustice, that loyalty has been known to strengthen a murderous aggressor and that generosity has sometimes weakened the capacity to do good.[191]

I interpret this position as claiming that when we think, for example, of those Nazis who exhibited qualities such as efficiency or dedication to overcome obstacles to their goals we might say that although they cultivated the virtues, they misused them, or used them in the wrong practice. But to make this claim directly opposes how Aristotle defined virtue. For Aristotle, a Nazi cannot be courageous because the trait is not cultivated in the right place, time, or manner.

---

[190] Ibid., 191.
[191] Ibid., 200.

> For instance, both fear and confidence and appetite and anger and pity and in general pleasure and pain may be felt both too much and too little, and in both cases not well; but to feel them at the right times, with reference to the right objects, towards the right people, with the right aim, and in the right way, is what is both intermediate and best, and this is characteristic of excellence. Similarly with regard to actions…[192]

According to Aristotle there can be no courageous thief or heroic Nazi.[193] Although a certain characteristic might appear like a virtue, if it is not with reference to the right objects, people aim, or way, then it is not a virtue. There is a simple reason for this. If we allowed the Nazi to have courage, then it would follow either that courage is not that important a virtue, or not a virtue at all. Part of the problem for MacIntyre is that he relativizes virtue to practices. This move forces him to conclude that all practices have virtues even if they are not moral virtues. Like Aristotle, MacIntyre wants to equate the development of excellence through a cultivation of internal goods with the general concept of virtue. Unlike Aristotle though, it follows from his argument that one can cultivate virtues even when engaged in evil practices.

Nevertheless, MacIntyre's biggest worry is that contemporary society has fragmented our system of moral value such that we live an episodic life ununified by virtue. He considers the way in which virtues provide a unity to life. He

---

[192] Aristotle, "Nicomachean Ethics" in *The Complete Works of Aristotle Volume Two*, ed. Jonathan Barnes (Princeton: Princeton University Press, 1984), 1106b19-1106b24.

[193] Kant believed that a thief could be courageous which led him to conclude that moral worth could only be located in the good will. But rather than being in opposition to Aristotle, Kant's position can be interpreted as being in agreement. Kant's point is that courage alone does not make a morally good person. It must be guided by a good will. Aristotle agrees that a characteristic is not a virtue if it is held in the wrong way. Although they use different terminology, they are actually in agreement that the virtues are unified or interconnected.

observes that one of the reasons that attention to virtue has fallen out of favor in recent times is that contemporary philosophers think atomistically about human action.

> That particular actions derive their character as parts of larger wholes is a point of view alien to our dominant ways of thinking and yet one which it is necessary to consider if we are to begin to understand how a life may be more than a sequence of individual actions and episodes.[194]

We do not usually think about how individual actions derive their character as parts of larger wholes. Yet it is necessary to think this way in order to understand how life is more than just a sequence of individual actions. Furthermore, the unity of human life becomes invisible to us when we make sharp distinctions between the individual and the roles she plays, so that life appears as unconnected episodes. The self so conceived cannot be the bearer of Aristotelian virtues because a self so disconnected loses the social relationships necessary for the Aristotelian virtues to function. Not only are virtues acquired and developed through social situations, but they also can only be acquired and developed throughout one's life and not in a single isolated action. A virtue is not a disposition that engenders success only in one kind of situation. Someone who possesses a virtue can be expected to manifest it in different types of situations. Therefore, the acquisition and development of the virtues requires a unified life.

MacIntyre argues that possessing virtues provides unity to a life. Because one possesses particular virtues in different situations, these virtues connect the

---

[194] MacIntyre, 204.

situations and unify them in the life of the possessor. However, I also understand his argument as claiming that the virtues cannot be acquired by an episodic self because an episodic self does not have the social relationships or engage in the social practices necessary for virtue. Hence a unity of self must be in place for the virtues to be acquired at all. For the self to have unity, the individual must compose a narratable life.

> To be the subject of a narrative that runs from one's birth to one's death is…to be open to being asked to give a certain kind of account of what one did or what happened to one or what one witnessed at any earlier point in one's life than the time at which the question is posed….Thus personal identity is just that identity presupposed by the unity of the character which the unity of a narrative requires.[195]

The unity of life consists in the unity of a narrative embodied in a single life. The good life for the individual is how the individual might best live out that unity. One can best live out this unity by embarking on a narrative quest. The quest is for the good that will enable us to order other goods and understand the integrity of life.

Virtues then are to be understood as those dispositions which will not only sustain practices and enable us to achieve goods internal to practices, but also which will sustain us in the relevant kind of quest for the good. They allow us to seek the good for ourselves as well as communally or politically. The good life is the life spent seeking the good life for persons, and the virtues necessary for the

---

[195] Ibid., 217-218.

seeking are those which enable us to understand what the good life is. Thus the virtues are situated not only in relation to practices but in relation to the good life.

Once he has established that cultivation of the virtues through practices is necessary for attainment of the good life, MacIntyre explains why virtues have been disappearing from contemporary life. There are a variety of virtue concepts functioning in a variety of ways. But there is a lack of consensus concerning the place of virtue relative to other moral concepts and to which dispositions are to be included among the virtues. The ancients understood not only games but also family and politics as practices with internal goods. The modern conception of family and politics does not. Furthermore, labor has shifted from outside the household and has led to an alienation and a move away from practices with internal goods. The internal goods of working have been transformed into the external good of payment. Practices have been moved to the margins of social and cultural life. The notion of engagement in a practice is no longer socially central. The aesthete has become a central figure of modern society and the narrative understanding of the unity of life has been marginalized. The virtues have been as a result transformed. They have been deprived of their conceptual background.

For these reasons, MacIntyre believes, the Aristotelian notion of virtue has fallen out of favor and a rationalist account of duty and moral obligation has gained prominence. One no longer finds value in practices; morality is now seen primarily as rule following rather than as the acquisition of virtues necessary for the good life. The focus is on what is right and not what is good. The individual is extracted from his context and reduced to an atomistic rule follower. He says that

181

the reason for this attention to rules was based on problems posed by philosophical egoism. A person came to be seen as someone who would do whatever was in his own interest unless constrained. Obedience to rules became the only remedy to the unruly passions, which were seen as dominating a person's existence. This outlook has led us to become alienated from the practices that will lead to a good life. This alienation is tragic.

Although MacIntyre does not explicitly couch his argument in terms of tragedy, it certainly is a kind of tragedy that we are alienated from pursuit of the good life. If he is correct that we are disconnected from those practices and virtues necessary to attain a good life, this would indeed be a tragic situation. I like MacIntyre's account for pointing out that not all virtue comes out of reason. Virtue comes out of what we do. He makes a compelling case for thinking about morality, not only in terms of rational principles, but in terms of behavior. Thinking of morality in this way supports my claims in chapter 4 that one can be morally responsible for acting according to one's cares or aliefs even if one holds a rational belief to the contrary. But there are a few objections to his account: one that is historical and two that are philosophical. Because MacIntyre's interpretation of history provides the foundation for his philosophical claims, the objection from history is just as damaging to his argument as are the philosophical objections.

MacIntyre argues that the emphasis on individualism in the Enlightenment project led to a fracturing of values. But it is doubtful that the notion of individualism arose as he describes it. It is unlikely that a few Enlightenment

philosophers attempting to counter philosophical egoism were responsible for the development in society of the notion of individualism. One might even doubt whether there was any unified project that can be assigned to the Enlightenment.

> It may be thought that the diversity of thinkers linked with that whole assemblage is too great, or the tensions between them too profound, to allow any ascription of a generic identity or common purpose to them, and eighteenth-century scholars who have failed to uncover any such "project" or "movement" or even "the Enlightenment" after a lifetime's research devoted to the subject could be forgiven their exasperation when confronted by so great a leap and quick fix.[196]

But even if we grant that there was indeed an Enlightenment project, there are two historical events that preceded it that support the notion that the Enlightenment took off because the principle of individualism already existed, rather than because it invented the principle. In 1215 the Fourth Lateran Council was called by Pope Innocent III. Canon 21 requires that all Christians confess his or her sins at least once a year to his or her parish priest. Previously, only the nobility were expected to participate in confession. Although this canon only reaffirmed prior legislation requiring confession from all Christians, it is only after this Council that the policy is taken seriously and implemented. It reflects a growing realization in European society that each individual person was important.[197]

---

[196] Wokler, 115.
[197] This same Council also required Jews and Muslims to wear special clothes to distinguish them from Christians. It is important to remember that although there was some evidence that the principle of individualism was considered important, it was only restrictively applied.

The second event that helped bring out the notion of individualism was the Protestant Reformation. Martin Luther introduced the concept of the priesthood of all believers. Each person was to have a direct relationship with God. This idea also helps influence the Enlightenment thinkers who place such high value on individuality. It is odd that although he places such great importance on understanding the historical context of systems of moral value, MacIntyre seems to think that Enlightenment thinkers invented the principle of individualism *ex nihilo*. These historical objections cast doubt on MacIntyre's argument from history that we are presently in a tragic situation due to the failure of the Enlightenment.

The first philosophical objection relies on exposing a false dichotomy that MacIntyre sets up between virtue ethics and the Enlightenment system of morality. MacIntyre is wrong that the Enlightenment emphasis on reason and obligation precludes the pursuit of virtue and the development of character. In fact, virtue plays an important role in Kant's moral philosophy.

In the *Doctrine of Virtue* Kant argues that virtue and the moral law are intimately related**.** He explains the difference between acts and ends as they relate to obligation and virtue. One can be coerced to perform an action but not to set an end. Duties prescribed by external law can only consist in actions. Only internal and ethical lawgiving can prescribe as duties the adoption of ends. Virtue brings inner freedom because compliance with the system of laws comprising it brings one's capacity for free choice in accord with one's will. Only a virtuous

disposition can lead one to adopt ends on the basis of their being duties, and

fulfillment of duties of virtue strengthens one's virtuous disposition.

> Now the capacity and considered resolve to withstand a strong but
> unjust opponent is *fortitude* (*fortitudo*) and, with respect to what
> opposes the moral disposition *within us*, **virtue** (*virtus, fortitudo
> moralis*). So the part of the general doctrine of duties that brings
> inner, rather than outer, freedom under laws is a *doctrine of virtue*.
> The doctrine of right dealt only with the *formal* condition of outer
> freedom (the consistency of outer freedom with itself if its maxim
> were made universal law) that is, with **right**. But ethics goes
> beyond this and provides a *matter* (an object of free choice), an
> **end** of pure reason…an end that…it is a duty to have.[198]

Virtue is the strength of will in overcoming obstacles (vices) that we put in the

way of our moral maxims. Pure practical reason is thus a capacity for ends. It

could not determine maxims for actions if it were indifferent to ends. Ethics does

not give laws for action. *Jus* does that and it is the subject of the *Doctrine of

Right*. Ethics gives *maxims* for actions. Laws that can only be given by one's own

will prescribe not actions, but maxims of adopting objective ends.

Virtue plays the important role that one must be the master of one's own passions.

It allows one to follow the ends of pure practical reason. Although Kant clearly

explains virtue in terms of rule following, duty, and moral obligation, it is not

correct to conclude, as MacIntyre does, that virtue is necessarily excluded by rule-

following ethics. But perhaps MacIntyre would reply that Kant does not employ

---

[198] Immanuel Kant, "The Metaphysics of Morals" in *The Cambridge Edition of the Works of
Immanuel Kant: Practical Philosophy*, ed. and trans. Mary J. Gregor (New York: Cambridge
University Press, 1996), 513.

the correct understanding of virtue. It is more than a good will; it requires the practices previously described. This brings me to the second objection.

The second philosophical objection is from David Miller. Miller argues that the virtues may not be necessary to excel in certain practices. He claims that there is a difference between those practices which serve only to provide internal goods and those which serve some social ends beyond themselves. The first are "self-contained" practices and the second are "purposive" practices.

Games are exemplars of the first category and seem to represent what MacIntyre has in mind when he speaks of practices. Here the distinction between internal and external goods has its clearest application. Clearly the good which consists in playing a game well cannot be understood outside the context of the game itself. The standards of excellence that determine what it is to play the game well are defined by those involved with the game-the players, judges, and spectators. But in the case of practices like architecture or farming there is an external purpose that determines both the goal of the practice and the conditions by which it may be judged. In these cases the good derived from being excellent in the practice is not simply constituted internally. The standards of excellence are related to a wider purpose.

> The good which consists in playing a fine innings at cricket is obviously incomprehensible in the absence of the game itself, and moreover the standard of excellence involved-what it is that makes the innings a fine one-can only be identified by reference to the history of the game…. On the other hand, in the case of a productive activity like architecture or farming, or in the case of an

intellectual activity like physics, there is an external purpose which gives the practice its point and in terms of which it may be judged.[199]

Hence we see that there are (at least) two distinct kinds of practices: those that are self-contained and those that have external purposive ends.

This distinction between self-contained and purposive practices is clear but it may not always be clear into which category a practice should fall. In the case of self-contained practices criticism can only be carried out from within the practice itself. But in the case of purposive practices the whole practice may be reviewed via the end it is meant to serve. MacIntyre assumes that all practices are self-sustaining.

But suppose virtues were understood primarily in terms of purposive practices. It would be impossible to defend any list of virtues without making reference to the social purposes that the practices are meant to serve. We can now only understand virtues based on the needs and values of a certain society.

Furthermore, Miller argues, self-contained practices like games and sports can exist only after basic social functions have been discharged. They are in that sense luxury items. If external goods like money and power allow one the means to pursue internal goods in games which we have just seen are luxury items, then it is not the case that the modern world has seen the complete erosion of the virtues.

---

[199] David Miller, "Virtues, Practices and Justice" in *After MacIntyre*, ed. John Horton and Susan Mendus (Cambridge: Polity Press, 1994), 250.

Finally, Miller raises the question of whether we really consider the qualities that allow one to excel in games to be virtues. The fact that participation in these practices is optional, that is people are not required to participate in them by physical necessity or moral obligation, counts against application of moral terms such courage and temperance. Justice in sports, that is abiding by the rules and having the rules apply equally to all, is really good sportsmanship rather than justice as we think of in the courtroom, in war, and before the law.

If Miller is correct that self-contained practices do not require virtues and only purposive practices do, then the practices that require virtues are few and far between. Furthermore, I think Miller lumps two kinds of practices under the category of "purposive," and that if we distinguish two kinds of purposive practices, those that require virtues will be even fewer. Miller says that all purposive practices serve social ends. But I think we can also distinguish between kinds of social ends. There are those social ends necessary for the continued existence and flourishing of the members of the society and those social ends that are useful or helpful, but not necessary. For example, the practice of medicine is necessary to the continued existence of the members of a society. People need to remain healthy to exist and flourish. On the other hand, art is purposive in the sense that it serves a purpose outside itself and creates objects of beauty for people to enjoy. However, the end it serves is not necessary for the continued existence of the members of society. Even if we can argue that art does serve some necessary social end, it is not clear that it requires the same kind or degree of virtues required by medicine or other practices which clearly serve necessary

188

social ends. If most practices do not require virtues at all, or do not require fully cultivated virtues in the robust sense which MacIntyre intends, then virtues do not provide the conditions necessary for moral personhood.

For these reasons, I do not think that MacIntyre provides a good foundation for thinking about moral tragedy. Certainly being alienated from the goods that make a good life would be tragic. However, MacIntyre's argument that we are currently in such a predicament is unconvincing. By his own argument, if we have practices, then we have virtues. But we do in fact have practices. Therefore, we have virtues. Perhaps the virtues we have are different than the ones he would like us to have. Or if he insists that we do not have virtues, then he is wrong that virtues are embedded in practices. There is a more compelling way to think about tragedy. It is to think about moral tragedy in terms of how we fail others rather than how we fail ourselves. It is a moral illusion to think that the locus of moral behavior is the cultivation of virtue for ourselves or for its own sake. Rather, it is fulfilling our responsibilities to others. When we shatter the illusion that moral life is "all about me" we can see that we sometimes transgress a moral value even when we follow the morally best course of action.

4. Inescapable Moral Wrongdoing

Christopher Gowans argues that we are sometimes faced with situations in which moral wrongdoing is inescapable. By inescapable wrongdoing he does *not* mean that we find ourselves in situations in which we have conflicting deliberative conclusions. It is not the case that an agent can be obligated to both perform and refrain from performing the same action. He also distinguishes his

discussion of inescapable moral wrongdoing from the traditional discussion of

moral dilemmas. A moral dilemma obtains when an agent is given two

alternatives A and B, and the agent can conclude that there are appropriate moral

reasons both to do A and to do B, but there are no overriding reasons to perform

one instead of the other. But these situations do not bring about the kind of

inescapable wrongdoing Gowans has in mind. He points out that situations may

arise in which an agent has overriding reasons to pursue one course of action over

another, and yet by pursuing this course of action one transgresses a moral value.

"[T]hough deliberative conclusions may not conflict, there are conflicts in which

infringement of a genuine moral value is unavoidable."[200] In such situations,

"whatever we do we will transgress a genuine moral value."[201] The traditional

discussion of moral dilemmas has focused on the deontic possibility of conflicting

deliberative conclusions. Although this discussion is philosophically interesting, it

ignores the most morally salient feature of the conflict; a moral value is

transgressed no matter what one does. Thus, we should investigate moral conflicts

as raising a normative issue, not only a metaethical one. Gowans shows that

inescapable wrongdoing arises due to our "moral responsibilities to particular

persons in virtue of our appreciation of the intrinsic and unique value of each of

these persons, and of our connections with them."[202] This account of moral

responsibility is based on understanding wrongdoing in terms of one's

responsibility to others, but not simply in terms of producing optimific states of

---

[200] Christopher Gowans, *Innocence Lost* (New York: Oxford University Press, 1994), viii,
[201] Ibid., viii.
[202] Ibid., ix

affairs or a duty to the moral law. I claim that understanding the possibility of inescapable wrongdoing characterizes the shattering of a moral illusion. It is the illusion that one can avoid all moral wrongdoing simply by following the moral law.

Most of us have faced a moral conflict at some point in our lives. We have been in situations where some moral reasons support one course of action and other moral reasons support another course of action. In such cases one usually chooses "the lesser of two evils." Choosing the lesser of two evils is the morally best choice. However, choosing the lesser of two evils still entails choosing an evil. That is, we may act for the best and yet still choose an evil. Although we may make the choice through no fault of our own, we will do something morally wrong no matter what we do.

Many philosophers believe that inescapable wrongdoing is impossible. Although it may be sometimes difficult or impossible to know what the best course of action is, there is a course of action that is completely free of wrongdoing. For example, Kant argues that negative duties cannot conflict. A negative duty is a duty that forbids any act whose maxim cannot be made into a universal law. These are acts we should never perform such as, "do not lie." There is one way to fulfill a negative duty since it is simply a case of not performing the forbidden action. If there is an apparent conflict between negative duties, then conflict is just that, apparent. Reasoning correctly will resolve any apparent conflict between negative duties. Duties cannot conflict due to the practical necessity that characterizes them.

191

> A *conflict of duties* (*collisio officiourum s. obligationum*) would be a relation between them in which one of them would cancel the other (wholly or in part). –But since duty and obligation are concepts that express the objective practical *necessity* of certain actions and two rules opposed to each other cannot be necessary at the same time, if it is a duty to act in accordance with one rule, to act in accordance with the opposite rule is not a duty but even contrary to duty; so a collision of duties and obligations is inconceivable (*obligationes non colliduntur*).[203]

Because a duty expresses a practical necessity, two conflicting rules cannot both be necessary at the same time. Practical necessity differs from natural necessity. It is not necessarily what is the case, rather it is necessarily what ought to be done commanded by the moral law. The moral law is discovered by pure reason. Autonomous agents, as noumena, are free from natural necessity but are subject to the law of freedom which requires that every autonomous act be in accordance with the universal laws of reason. That is, an autonomous agent may choose not to act in accordance with these rules, but as far as she acts rationally she acts in accordance with them. In this sense the "ought" of the moral law is one of practical necessity. Given this necessity, duties cannot conflict. It is inconceivable that two contradictory states of affairs are both necessary. Just as it is impossible according to natural law that two objects are located in the same place at the same time, so is it impossible according to the moral law that two inconsistent acts be obligatory.

---

[203] Kant, "The Metaphysics of Morals," in *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*, ed. and trans. Mary J. Gregor (New York: Cambridge University Press, 1996), 378-379.

It is true that Kant's moral philosophy precludes the possibility of inescapable moral wrongdoing. Therefore, it is inconsistent to accept both Kant's conclusion and the notion of inescapable wrongdoing. But rather than showing inescapable wrongdoing to be impossible, the inconsistency may give us reason to reject Kant's thesis that obligations cannot conflict. In order to have good reason to do so, we need good evidence that inescapable wrongdoing is a fact of our moral experience. We gather this evidence through empirical examples and through Gowans's theory of moral responsibility that he calls, "responsibility to persons."

We gather empirical evidence for inescapable wrongdoing through our own experience of being faced with moral conflicts, but also through vivid examples in literature. Gowans presents inescapable wrongdoing by examining Melville's *Billy Budd*. It is important to the story that it takes place historically at a time when a number of mutinies in the English navy had recently occurred. The crew of the HMS *Bellipotent* was made up of some sailors who had participated in these mutinies and some who had been brought into the navy by impressment. Billy was one who had recently been impressed. His character was such that he was quite simple and incapable of malicious intent. Billy suffered from a speech impediment that was exacerbated by stressful situations. The first-mate, Claggart, was a mean-spirited person who did not like Billy's simple demeanor. He continually picked on Billy knowing that Billy was an easy target for his machinations. Well aware of the King's worry over rebellion, and thus the importance for the Captain of the ship to oppress any thoughts of mutiny,

193

Claggart accused Billy of mutiny. Billy's speech impediment kept him from explaining himself and in frustration he struck Claggart who fell dead to the deck. Not only did Captain Vere and the crew know that Billy was incapable of intentionally murdering Claggart, Vere also knew that he was incapable of organizing a mutiny. They knew that Claggart's accusations were false. But Vere was a man to whom maintaining order was very important. He ordered an immediate trial so that the idea of mutiny would not resonate with any of the other crew members.

The court believed that Billy held no malice towards Claggart and would not have struck him if he could have spoken. The court also accepted that Claggart targeted Billy because of Billy's simple innocence. Claggart was envious of Billy and hated him because of his simplicity and innocence, and for these reasons falsely accused him of mutiny. Captain Vere knew these facts as well. He freely admitted that there were moral grounds for showing Billy leniency. And yet he also cited naval law that dictated the execution of any sailor who strikes his superior officer. Vere argued that as members of the King's military, they owe allegiance to the King's laws. Although the law may be imperfect in this case, they had an obligation to enforce it. We can see that Vere and the officers of the court will perform a morally wrong act no matter what they do. On the one hand, they have moral reasons of compassion and justice to let Billy go free. On the other hand, they have reasons to execute Billy based on their responsibility to enforce the law. Vere decided that their responsibility to enforce the law preempted all other moral obligations. Billy was executed.

We might immediately respond to the story by saying that although Vere had conflicting responsibilities, it is clear what he ought to have done. The moral reasons for letting Billy go free override the moral reasons for fulfilling one's responsibility to enforce the law. Gowans agrees. Recall that inescapable moral wrongdoing is not a situation where reasons for two courses of action are equal. It may indeed be quite clear which action is supported by overriding reasons. Nevertheless, even if Vere had let Billy go free, he would have neglected his duty to the king to enforce the law. Regardless of which course of action he pursues, one of his responsibilities must be ignored. Even if letting Billy go free is the morally best course of action, he can only perform it by ignoring his responsibility to the king. If doing something wrong is understood as not fulfilling one's responsibilities, then Vere would have done something wrong even if he performed the morally best action. He would have done something wrong in some sense, no matter what he did. Hence, wrongdoing is inescapable in some circumstances.

Many philosophers would object and argue that Vere would not have done anything wrong if he had deliberated correctly and realized that the reasons for letting Billy go free were more compelling than the reasons for enforcing the law. But this objection is based on a peculiar position.

> [T]he view that wrongdoing may always be avoided is committed to something like the following position. In a given situation, there is only one actual moral responsibility, and that is to do what moral deliberation determines in the final analysis ought to be done. Hence, the conclusion that Vere morally ought to take a particular

195

course of action completely eliminates the validity of those moral reasons favoring incompatible courses of action. [204]

If the correct conclusion is that Vere ought to have let Billy go, then it follows that he did not also have any responsibility to enforce the law. That is, if moral obligations cannot conflict, then it only appeared that he had another moral responsibility. But this conclusion rests on ignoring some features of our moral responsibility to others.

One reason many philosophers believe that inescapable wrongdoing is impossible is because they believe that "the primary concern of morality and moral theory is the determination and justification of conclusions of moral deliberation."[205] For utilitarians and Kantians, sound deliberation taking a first moral principle as a starting point, cannot produce conflicting moral judgments. If one reasons correctly, and then acts based upon the derived moral conclusions, then one has done what one ought to do. It would be incoherent to argue that there are situations where a correct conclusion, all things considered, of an agent's moral deliberation about what to do would conflict with itself such that the agent ought and ought not perform a specific action. However, there are situations where although an agent does what she ought to do, some moral value is transgressed. In this sense, moral wrongdoing is sometimes inescapable.

If we understand wrongdoing as transgressing some moral value based on some responsibility to someone, then there may be cases where an agent will do something morally wrong no matter what. It is in the specific sense of

---

[204] Gowans, 18.
[205] Ibid., 119.

transgressing moral responsibilities that inescapable wrongdoing arises. Rather than beginning with an abstract moral principle, the "responsibility to persons" account begins by finding the locus of moral responsibility in our relationships with others.

> We are to start with reflection upon what is involved in our particular, concrete relationships with persons with whom we are to a greater or lesser extent, and in various ways, intimate, especially relations of kinship, friendship, and love. It is mainly in the context of such relationships that we first come to employ and understand moral considerations.[206]

Although all of our responsibilities will not be found in intimate relationships, these relationships serve as paradigms on the basis of which responsibilities in other contexts may also be understood. We all suppose that we have specific responsibilities to those persons with whom we are so related. For example, I have several responsibilities to my son, such as the responsibility to nurture and protect him. This responsibility differs from those I have toward my wife, such as the responsibility to be supportive and loyal. The specific responsibilities arise due to the specific nature of the relationships I have. The responsibilities are rooted in two kinds of consideration. "The first is the perception that each of these persons has intrinsic and unique value. The second is the recognition that some connection or another obtains between oneself and these intimates."[207]

To say that a person has intrinsic value means that we find others to be valuable in themselves. However, we need not do so on the basis of their being

---

[206] Ibid., 122.
[207] Ibid., 123.

free and rational beings. For Kant only the noumenal person should be regarded as an end in itself. Respect for persons is derived from the moral law dictated by pure practical reason. But on Gowans's approach we do not have to value a person via the application of an *a priori* law; rather, we come to think this about people due to our interaction with them. Our relationships with others are built upon the experience of their intrinsic value.

To say that each person is uniquely valuable means that each person's intrinsic value is different from that of everyone else. For Kant, a person is intrinsically valuable due to characteristics shared by everyone: persons are rational and autonomous. It follows that nothing that distinguishes persons from one another is of moral significance. Kant is not alone in approaching morality from this kind of standpoint. Utilitarianism is also founded on a principle of abstract moral equality. Bentham asserts that everyone is to count for one and no more than one. However, even if individuals are equal to one another because of these shared characteristics, and although it is true that the concept or moral equality plays an important role in questions of morality, it does not follow that how persons differ from one another does not have any moral significance. Moreover, it does not follow that every alternative to the principle of moral equality of persons results in regarding some people as more valuable than others. It may be that each person is uniquely valuable and that the value of each person is incommensurable.[208]

---

[208] Kekes would disagree that all individuals have equal value. For Kekes, human worth depends on moral merit. Moral merit varies, so human worth is unequal. Evil people do not deserve the same reaction as good ones, because they have less worth. See Kekes, *Facing Evil*, 106-123. This

We experience this incommensurability most poignantly when a loved one dies. We think that the person cannot be replaced because of his or her unique value. Gowans points out that when Creon kills Antigone, Haimon commits suicide because of his inability to go on living without her. Creon thinks that there are other women whom Haimon could love and take as a wife. But Antigone's unique value cannot be replaced. Haimon is not distraught because a free and rational being is now gone. If that were the case, then she could easily be replaced. Although Haimon could have found another wife whom he presumably would have loved dearly, she could not replace the person who was Antigone. We can understand this story only by seeing that each person has unique value.

One might think that focusing on the unique value of persons will lead to regarding some people as more important than others. There are indeed some situations where consideration of unique value should be put aside, such as before the law. But there are many situations where it would be odd not to consider another's unique value and the specific relationship one has. Further, just because one does consider another's unique value and the specific relationship they have with one another when deciding what to do, it does not follow that one degrades the intrinsic and unique value of others.

> The woman who saves her husband from drowning before saving another passenger on the ship is making a comparative assessment with respect to that action. She is judging that it is more important

to her to save her husband than to save the other passenger. But she need not be taken as judging that her husband is more valuable than the passenger simply speaking. She can acknowledge that that person is also an intrinsically and uniquely valuable being, albeit one for whom she does not have the same responsibilities as for her husband.[209]

In this situation we would not morally criticize the woman for choosing to save her husband before the other passenger. In fact, we would probably morally criticize her if she chose to save the other passenger before her husband. The locus of the woman's moral responsibility is found in her relationship to her husband.

Although we understand that persons are intrinsically and uniquely valuable by examining our more intimate relationships, we can then extend this recognition to strangers as well. I recognize that every person has intrinsic and unique value even if I do not know specifically what is uniquely valuable about that person. We are in less of a position to understand what is uniquely valuable about strangers. Yet, we recognize that they are uniquely valuable. Thus, there is a sense of moral equality. The fact that persons are intrinsically and uniquely valuable means they are deserving of a certain kind of treatment. The treatment depends on the kind of relationship one has with another. In contemporary society relationships are established both through choice and also through unchosen circumstances. I chose to get married and hence chose to assume moral responsibilities to my wife. But merely encountering others can be sufficient to establish a relationship. A relationship is a state of affairs relating one individual

---

[209] Ibid., 126.

to another. Awareness of another's plight is sufficient to relate one person to another. Therefore, we need not conclude that the "responsibility to persons" account would justify ignoring a drowning stranger or starving children in a developing country. Rather, it reflects our intuition that our responsibilities are often strongest to those with whom we share an intimate relationship. Nevertheless, we also have moral responsibilities to others we encounter.

These differences in the kinds of relationships we have produce different kinds of moral responsibilities. My responsibilities to my son include nurturing and protecting him. My responsibilities to my students include fairness in grading and ensuring the assignments are worthwhile. Recall in chapter 3 that I argued on the basis of Rachels's analysis of moral considerability that our moral responsibilities differ depending on the person (or animal) with whom we are interacting and on the situation we are in. For example, it is wrong to hit you with a stick, not because you are autonomous, but because you feel pain. It is wrong to treat you like a child, not because you feel pain, but because you are autonomous. We can now add to this account that various responsibilities also arise due to the kind of relationship one has with another.

Now it is possible to see that sometimes moral responsibilities may conflict with one another. Even if I deliberate correctly and conclude what the morally best action is, I still may do something wrong. Because I have various relationships with others, and consequently various moral responsibilities to others, a situation may arise where I can fulfill one responsibility but not another. In this sense a moral value is transgressed. Not fulfilling a moral responsibility

201

constitutes a wrong. When a moral value is transgressed, a wrong has been committed. Hence, it is possible that wrongdoing is sometimes inescapable. If the wrongdoing fulfills the conditions I laid out in chapter 3 for evildoing, then it follows that evildoing is also sometimes inescapable.[210]

Sometimes the wrongdoing or evildoing can take on tragic dimensions. There are numerous kinds of tragic-making characteristics. It can be tragic when

> the morally best action seriously harms or allows to be harmed a person or social entity to whom the agent is morally responsible,…the morally best action results in a harm that is either irreversible or extremely difficult to repair,…the morally best action results in a harm that is far-reaching in its consequences,…the morally best action not only fails to fulfill a moral responsibility, but actively works against that responsibility,…the morally best action harms or neglects a person whom the agent especially values,…the morally best action harms or neglects a person who is especially undeserving of this harm or neglect,…the morally best action renders the agent a tool in the evil projects of others,…the morally best action involves doing something that is degrading to the agent,…finally, the moral conflict is one in which the moral reasons for two conflicting actions do not override each other, and yet each overrides the reasons for all other alternative actions.[211]

If moral wrongdoing or evildoing are sometimes inescapable, and they sometimes take on tragic dimensions, then it follows that moral tragedy is sometimes inescapable. The tragic situations are ones where the agent makes the morally best

---

[210] Although Gowans retells the story of *Billy Budd* to illustrate inescapable wrongdoing, it is not simply that no matter what Vere does he commits a wrong. By executing innocent Billy he actually commits an evil act.
[211] Ibid., 226-227.

choice, a wrong is committed, and the wrong is tragic. One must understand that one can bring about moral tragedy unchosenly and often inescapably to have a full understanding of moral interaction among persons. Hence, understanding inescapable wrongdoing and moral tragedy shatter the illusion that following the moral law will keep one morally pure. One discovers that wrongdoing is sometimes inescapable and also tragic. The shattering of this illusion does not result in the avoidance of future moral tragedies; because they are sometimes inescapable they cannot always be avoided. To believe that they can be is a moral illusion. Rather, it is a moral disillusion to recognize undeniable features of moral life and sees that she can be morally responsible for what is unchosen, inescapable, and tragic while still doing the right thing. The disillusion gives one the ability to question moral principles and directives.

That one does not choose the circumstances that might lead to the inescapability of wrongdoing might be unfair, but what is unfair is not unreal. To deny that inescapable moral tragedy is a feature of moral life is to deny some of our moral responsibilities. Recognizing the nature of inescapable moral wrongdoing and moral tragedy is to take responsibility for oneself as a moral individual.

5. Regret

Taking full responsibility for oneself as a moral individual often involves feelings of regret. If one commits an evil act despite having made the best moral choice, it is natural for the agent to regret what she has done. The agent feels regret because the state of affairs she brought about do not reflect her values. As I

discussed in chapter 1, feelings of regret, remorse, and indignation play an integral role in being a moral agent.

When we value something or somebody, we adopt and pursue ends with respect to it. Kant and Frankfurt both endorse this view, but in different ways. For Kant, when we value the rational capacity of another we adopt and pursue the end to respect that rational capacity and treat the individual accordingly. For Frankfurt, when we care about something we make ourselves vulnerable to the fate of what it is we care about. In a situation of inescapable moral wrongdoing, the ends we have adopted cannot be pursued. Our responsibilities come into conflict creating an obstacle to our fulfilling at least one of them. This inability to pursue the ends we have adopted produces feelings of regret.

Gowans argues that it is appropriate for an agent to feel moral distress in situations where responsibilities conflict. In fact, he claims that the existence of feelings of moral distress constitute a phenomenological argument in support of the thesis that inescapable wrongdoing occurs. Those who claim that conflicts of obligation are only apparent must also conclude that there is no good reason for an agent to feel moral distress in these situations. If an agent cannot do anything wrong as long as she does what is morally best, then there would be no good reason for the agent to feel regret for doing something wrong. But moral distress does occur and we should try to explain it rather than dismissing it. Faced with a conflict of responsibilities, I do not feel regret or remorse because two actions, *A* and *B*, cannot both be done. Rather, I feel regret or remorse because *I* cannot do both *A* and *B*.

204

In chapter 1, I argued that the individual losing moral innocence does not lose innocence by realizing that bad things happen. Rather, the individual loses innocence by realizing that she can bring about bad things. That is, one must locate wrongdoing in terms of one's own agency. At this more complex phase of moral disillusion, one also locates the transgression of value due to inescapable wrongdoing in terms of one's own agency. And one feels regret.

Gowans's phenomenological argument is based on the original formulation of moral distress made by Bernard Williams. Williams defines regret as the feeling that things would have been much better had they been otherwise. He defines agent-regret as the feeling one has towards one's own past actions whether intentional or not.

> The constitutive thought of regret in general is something like "how much better if it had been otherwise", and the feeling can in principle apply to anything of which one can form some conception of how it might have been otherwise, together with consciousness of how things would then have been better. In this general sense of regret, what are regretted are states of affairs, and they can be regretted, in principle, by anyone who knows of them. But there is a particularly important species of regret, which I shall call "agent-regret", which a person can feel only towards his own past actions…[212]

When one feels agent-regret, one wishes that things had been otherwise, but not necessarily that one had acted otherwise. Rather, the agent has a negative evaluation both of the circumstances that constrained her choices and of that she

---

[212] Bernard Williams, *Moral Luck: Philosophical Papers 1973-1980* (New York: Cambridge University Press, 1981), 27.

was the one who brought about the state of affairs. Even if an agent's action is justified by utilitarianism or Kantian ethics, she still might justifiably feel agent-regret that she had to be the one to do whatever she did. For example, Captain Vere could justifiably feel agent-regret even if he had deliberated correctly and let Billy go free. He still would have let his responsibility to the king go unfulfilled. That his action would have been morally the best thing to do would not make feeling agent-regret irrational. Because the feeling is a natural part of moral life it would be irrational to dismiss its validity. The agent regrets not just that a state of affairs obtained, but that she was the one who brought it about. Williams and Gowans help show that an unrealistic optimism towards maintaining one's moral purity reflects a lack of understanding concerning one's moral responsibilities. The morally disillusioned individual sees that she lives in a morally tragic world and regrets that she is sometimes complicit in bringing tragedy about even when she does whatever is morally best.

6. Conclusion

In this chapter I have shown that understanding inescapable moral wrongdoing and moral tragedy describe a moral disillusion. Understanding tragedy as inescapable due to one's own agency differs from the tragedy of necessity that characterized Greek tragedy and Aristotle's conception of tragedy. It also differs from the Elizabethan form of tragedy which was characterized by the possibility of averting the tragic conclusion. MacIntyre's conception of tragedy proved to be untenable due to his misleading interpretation of history and problems with his philosophical account. The Greek and Elizabethan approaches

to tragedy are problematic because they maintain the illusion of moral purity. The Greeks believed that one cannot willingly do wrong. Hence, by being virtuous one could be morally pure. Kant and utilitarians believed that one can avoid wrongdoing by following the categorical imperative or the greatest happiness principle, respectively. In other words, one can be morally pure. The focus is on maintaining one's own moral goodness. But the accounts of unchosen evil and inescapable moral wrongdoing shatter these moral illusions. Because understanding these features of moral life shatter naïve illusions of maintaining moral purity, it follows that *not* recognizing these features supports an illusory morality ignoring facts about moral life. Moral disillusion is not characterized in part by the ability to know the morally right way to act; rather it is characterized by understanding that even when we do the morally right thing, we may still bring about evil or transgress moral obligations. One takes responsibility for the unchosen or inescapable evil one brings about. The proper response to unchosen or inescapable evil is not the fabrication of an abstract moral principle that denies their existence, but inquiry into their nature, feelings of regret or remorse, and acts of confession, contrition, apology, restitution, or reparation. In these ways one takes responsibility for the evil one causes.

CHAPTER 7

KIERKEGAARD AND THE IMPORTANCE OF WHAT WE CARE ABOUT

1. Introduction

In the previous chapters I described the different phases of moral disillusion. I showed that moral life is richer than barely being the target of the reactive attitudes. Acquiring a more nuanced understanding of moral life gives one the ability to raise questions concerning the adequacy of moral principles and directives. In this chapter, I bring in Kierkegaard's conception of the self, and I do so for two reasons. First, Kierkegaard gives us a compelling account of moral disillusion that is compatible with my approach. Second, he *prescribes* shattering moral illusions whereas I have heretofore only described it. Prescribing moral disillusion both raises an important question and also clarifies a possible confusion. One component of the process of shattering moral illusions is becoming interested in how we relate to ourselves and to others. To prescribe the process is to say that we ought to become interested in ourselves and how we relate to others. I consider this interest to be consistent with Frankfurt's description of our second order cares. Frankfurt argues convincingly that people not only care about things, but that they also care about what they care about. Although Frankfurt maintains that caring about what we care about is simply descriptive of what is it to be human, he adds that we can affect what we care about, and that some cares are worthier than others. Hence, like Kierkegaard, he argues we ought to care about some things and not others. But in previous chapters I argued that our cares are not entirely within our control. Thus, the

208

question is: How can we affect our cares if they are not entirely within our control? The answer to this question lies in the increased understanding of moral life that accompanies moral disillusion. The clarification that reemerges from this discussion is that moral disillusion and moral goodness are conceptually distinct. As I have repeated throughout the previous chapters, we need to be careful not to think that moral disillusion entails moral goodness. I show that moral disillusion generates increased moral responsibility, not moral goodness.

Finally, a word of warning. I will focus on some of the main themes found in approaching Kierkegaard's authorship in its entirety. But giving an account of Kierkegaard's entire philosophy would be a monumental task. He is intentionally indirect and seems to enjoy presenting the reader with apparent inconsistencies and contradictions. However, this indirect presentation of apparent contradictions plays an important role in the message he communicates. I will explain this methodology as I explain his philosophical themes. However, it is both impossible and unnecessary to include in this chapter every philosophical theme deemed significant by Kierkegaard. It is impossible because one could write an entire dissertation on each of Kierkegaard's works. It is unnecessary because the absence of some of his philosophical themes in the present discussion will not unfairly skew the conclusions I draw. In this chapter I will explain Kierkegaard's theory of moral disillusion, relate it to Frankfurt's cares, explain how we can affect our cares, and explain why moral disillusion does not entail moral goodness.

2. Kierkegaard's Methodology

Before I give an overview of Kierkegaard's conception of the self and how it gives us an account of moral disillusion, I need to say a few words about the methodology he uses to present his philosophy. Understanding the manner in which he does so is essential to understanding how and why the reader should analyze his authorship in its entirety in order to correctly uncover the message he communicates. That is, one cannot understand the philosophy without understanding the methodology.

It is well-known that Kierkegaard uses pseudonyms for many of his works. Understanding why he does so gives us a clue to understanding his philosophy of moral disillusion. The point of view espoused by each pseudonym should not always be considered to be Kierkegaard's complete view. That is, each pseudonym is like a character in a play. Only by seeing the play in its entirety can one see how the story ends. It follows that it is premature to take, for example, what Judge Wilhelm in *Either/Or* says about the ethical stage as Kierkegaard's final statement of the ethical. Judge Wilhelm is only one character and, therefore, one point of view. For this reason, one cannot identify Kierkegaard's point of view until one considers the themes that continue throughout his various works. His authorship as a whole represents his point of view. Each single work is only one component of the entire philosophy.

The presentation of his philosophy through pseudonyms is not arbitrary. It is quite significant that he presents his views in this way. As I will show below, shattering moral illusions is an undertaking that each individual must do on her

own. Shattering moral illusions requires self-reflection. One must realize that one's own agency is often responsible for wrongdoing or evil in the world. This realization cannot be acquired passively. That is, one cannot be taught this fact of moral life; as we saw with the morally innocent, the agent must confront moral facts on her own. "Kierkegaard is not a didactic author, he offers no catechisms, he never speaks *ex cathedra* or with an *imprimatur*. He teaches by setting the individual free."[213] Kierkegaard employs indirect communication to teach us about shattering moral illusions. He does not explicitly tell us what one needs to do. Rather, he uses pseudonyms and indirect communication to provoke the reader into self-reflection. The reader must come to an understanding about moral disillusion on her own. Thus, reading Kierkegaard's works reflects the process of moral disillusion itself. He presents his philosophy indirectly, requiring an interpretation of his authorship in its entirety because one cannot shatter moral illusions by being explicitly told about individual responsibility. One of the illusions one shatters is that moral responsibility is something abstract and impersonal.

Now I turn to examine his philosophy of moral disillusion. In order to do so, I will explain Kierkegaard's analysis of social ethics that stands in opposition to individual responsibility, the roles of both Socrates and Abraham as exemplifying the shattering of moral illusions, and the meaning of despair. These elements come together to establish that what is important for a person is to be passionate about reflecting upon one's individuality and responsibilities. As I

---

[213] Martin Beck Matuštík, "Becoming Human, Becoming Sober," *Continental Philosophy Review* 42 (2009): 261.

mentioned above, I do not claim to give an exhaustive account of Kierkegaard's philosophy; rather, I present here significant elements of his view of moral disillusion. After presenting Kierkegaard's account, I will show that the passionate inwardness he describes (and prescribes) corresponds to Frankfurt's description of caring about what we care about and affecting our cares. Finally, from this discussion it will become apparent that moral disillusion does not entail moral goodness.

3. Kierkegaard's Conception of the Self

To best understand Kierkegaard's view of individual responsibility and moral disillusion it is helpful to situate his philosophy in its historical context. To do so, however, is not to claim that his philosophy has only historical significance. As I will show in this chapter, his thought has contemporary import. The reason it is helpful to situate him historically is because doing so confronts and resolves a widespread misunderstanding about Kierkegaard's views of ethics. Many philosophers believe that Kierkegaard presents his view of ethics in *Either/Or* and then argues in *Fear and Trembling* that ethics should be "suspended" and transcended in favor of religious life. However, interpreting Kierkegaard in such a way ignores that both of these works are written by pseudonyms. As I previously mentioned, the use of pseudonyms is Kierkegaard's subtle way of telling the reader that the view presented in each pseudonymous work is not the final word on the matter at hand. Again, the pseudonym is only a single character in a larger and more complicated play.

The kind of ethics described by Judge Wilhelm in *Either/Or* represents a

social ethic that Kierkegaard will later contrast with individual responsibility.[214]

At the time that Kierkegaard was writing, academic philosophy was dominated by

Hegelian thought. Many people agreed with Hegel that Reason had realized itself

through his writing and their reading *Phenomenology of Mind*. Hegel had resolved

the most pressing issues of philosophy. Hence, for many, studying philosophy had

become studying Hegelian philosophy. But Kierkegaard rejected Hegel's

endorsement of social ethics, i.e. the ethics of the state, because it subsumed

individual responsibility to an abstract universal. The ethics that is suspended by

Abraham in *Fear and Trembling* is not ethics per se, but Hegelian social ethics.

Hegel claimed that the world-historical idea had resulted in the ethics of

the state. This ethics was the result of a divine rational process. Because of its

divinity and rationality, one must accept the social ethic and act according to its

precepts. Hegel claimed that rejecting the social ethic and relying upon individual

conscience allowed for the possibility of doing evil. Following one's own

conscience could not guarantee ethical behavior, but following the social ethic

could. For this reason Hegel rejected Kantian morality as too general and lacking

concrete guidelines by which one could act. If one followed, however, the ethics

that have come to be as a result of the world-historical idea, one could act

ethically. Only by acting in this way does the individual recognize herself as a

---

[214] Alexander Pushkin's novel *Eugene Onegin* presents the reader with two fitting examples of a social ethic. First. Onegin's friend Vladimir Lensky feels that Onegin insulted the honor of his fiancée Olga. As a result Lensky feels obligated to challenge Onegin to a duel. Onegin feels compelled by social convention and reluctantly agrees to the duel in which he kills his friend. Second, scorned by Onegin's rejection of her love, Tatyana marries an aged prince out of social obligation. Although Onegin later falls in love with her, and she admits she still loves him as well, she is determined to remain faithful to her husband.

subject among other subjects and acquire freedom. It is through the social order

that subjects can relate to each other harmoniously under the law.

> But the subjective will also has a substantial life—a reality—in
> which it moves in the region of *essential* being, and has the
> essential itself as the object of its existence. This essential being is
> the union of the *subjective* with the *rational* Will: it is the moral
> Whole, the *State*, which is that form of reality in which the
> individual has and enjoys his freedom; but on the condition of his
> recognizing, believing in, and willing that which is common to the
> Whole…. The laws of morality are not accidental, but are
> essentially Rational. It is the very object of the State that what is
> essential in the practical activity of men, and in their dispositions,
> should be duly recognized…. It must further be understood that all
> the worth which the human being possesses—all spiritual reality,
> he possesses only through the State.[215]

For Hegel, the individual realizes her individuality only through being with others

in the community of the state. It follows that following the social ethic of the state

was moral since it allowed each individual to fully actualize her worth as a

human. Because the existing state was a result of divine Reason realizing itself,

then "[i]t is certainly the case that the individual person is a subordinate entity

who must dedicate himself to the ethical whole."[216]

Judge Wilhelm in *Either/Or* presents a congruent view of ethics. He

presents this view in one of his letters to the aesthete. He attempts to convince the

---

[215] G.W.F. Hegel, *Philosophy of History*, trans. J. Sibree (New York: Dover, 1956), 38-39.
[216] Hegel, *Elements of the Philosophy of Right*, ed. Allen Wood, trans. H.B. Nisbet (New York: Cambridge University Press, 1991), 102.

aesthete that choosing to live out one's social role is the ethical way to live. By living out one's social role, one becomes a concrete, stable individual.

> The self that is the objective is not only a personal self but a social, a civic self.… He transfers himself from personal life to civic life, from this to personal life. Personal life as such was an isolation and therefore imperfect, but when he turns back into his personality through civic life, the personal life appears in a higher form.[217]

For this reason the Judge recommends marriage to the aesthete. Marriage represents the stable social life that is characteristic of the social ethic. The aesthetic way of life cannot provide stability to the self. One ought to choose the ethical way of life. The way to do so is to choose to live out one's role in society.

But Kierkegaard realizes that simple rule following does not make one morally responsible. "He takes Hegel's well socialized citizen of the nation-state and argues that to become ethical, it is insufficient to join a membership in a conventional community."[218] Rule following gives us no insight into an individual's motives or intentions. Perhaps it serves as a good training device, but it does not make for a fully responsible individual. Under this model people only relate to each other in a herd-like manner. They do not individually reflect upon what it means to be a responsible individual. The people only want to adhere to the demands of the law. They fail to consider whether observing these demands is sufficient for taking responsibility. As I will continue to show, Kierkegaard calls

[217] Soren Kierkegaard, "Either/Or, A Fragment of Life, II," in *The Essential Kierkegaard*, ed. Howard V. Hong and Edna H. Hong (Princeton: Princeton University Press, 2000) 82.
[218] Martin J. Matuštík, "Kierkegaard as Socio-political Thinker and Activist," *Man and World* 27 (1994): 216.

for the individual to question this rule-following social ethic. This questioning is a

shattering of moral illusions.

> Kierkegaard suggests that one way of responding to the dilemmas
> of the age is to cultivate "passionate inwardness".… This
> disposition requires that the individual not assume that what the
> established order promulgates in the law is the best way to live a
> life, but rather adopt a critical posture in order to verify and
> legitimate the methods of justification the state appropriates to
> sustain itself.[219]

The process by which one questions the social ethic is difficult. One must

cultivate a passionate inwardness that allows one to separate the internal and the

external so as not to be completely defined by the external, i.e. society.

Kierkegaard indicates that when the state and one's social role in the state define

who one is, then one is not yet a "self." In my terms, one holds fast to moral

illusions. One must become a self, or shatter moral illusions, through the difficult

process of self-reflection.

This passionate inwardness requires a detachment from one's social role

and the social ethic which defines it. This detachment requires "sacrificing," or

questioning, even those relationships which have the greatest significance. It is

only by questioning these relationships that one can relate to them in a personal or

subjective way. Hence the ethics that is suspended in *Fear and Trembling* is not

ethics wholesale. Rather, Abraham suspends the social ethic in order to cultivate a

---

[219] Mark Dooley, *The Politics of Exodus: Kierkegaard's Ethics of Responsibility* (New York: Fordham University Press, 2001), 45.

more personal responsibility devoid of illusions. To better understand this questioning of the social ethic, I now turn to Kierkegaard's treatment of Socrates.

To help illustrate what is involved in this process of self-reflection and questioning the status quo, Kierkegaard, through his pseudonyms, describes Socrates as a paradigm of how to first react to the social ethic. Socrates questioned the social ethic of his time. He devoted, and lost, his life to this endeavor. He asked his interlocutors to think about whether the "truths" they had been taught were coherent. Socrates forced his interlocutors to question what is good and what is just, without relying on the social ethic. He encouraged his listeners to suspend belief in the established laws and codes of ethics.

> Socrates had the courage and self-collectedness to be sufficient unto himself, but in his relations to others he also had the courage and self-collectedness to be merely an occasion even for the most stupid person. What rare magnanimity—rare in our day, when the pastor is little more than the deacon, when every second person is an authority, while all these distinctions and all this considerable authority are mediated in common lunacy and in a *commune naufragium* [common shipwreck], because since no human being has ever truly been an authority…it never fails that one fool going his own way takes several others along with him.[220]

Socrates did not trust the social ethic to define moral responsibility. However, he did not reject society; that is, he did not question the social ethic in isolation. He detached himself from the society in which he lived in order to carry out the difficult process of self-reflection and questioning. This is not a total alienation or withdrawal from the cultural context, but the adoption of a critical stance towards

---

[220] Kierkegaard, "Philosophical Fragments" in *The Essential Kierkegaard*, 118.

the given socio-political structure. His goal was not necessarily to reject the social ethic, but to put it to question.[221] Most importantly, it was essential that each individual interlocutor raised his own questions concerning the status quo. Socrates accomplished this task by employing a very specific method: the use of irony.

Irony gives one the detachment necessary to question the status quo. As a device it gives one the conceptual space needed to reflect upon one's own responsibilities and the roles given to one by society. Furthermore, the Socratic Method does not give the interlocutors any determinate answer. Socrates only questions. The individual must come to her own conclusions regarding what is good, pious, right, and true. If Socrates were to give his interlocutors the answers, then his answers would have served the same role as the social ethic. The individual would not reflect and come to her own conclusions. She would take the easy route of relying upon what is given to her, rather than the difficult route of self-reflection and self-becoming (i.e. becoming a self, or morally disillusioned).

Socrates, through irony, serves as a gadfly. He provokes reflection among his interlocutors without giving them answers.[222] Kierkegaard, through his pseudonymous authors, also serves as a gadfly. He wants to tease the unreflective out of their complacency. He wants to provoke the aesthete and the ethical person (i.e. the person like Judge William in *Either/Or* who wholly subscribes to the

---

[221] Nietzsche gives *Twilight of the Idols* the subtitle *Or How to Philosophize with a Hammer* alluding to a similar notion of sounding out ideas, i.e. tapping them lightly, to see if they are hollow or ring true.

[222] For this reason we might say that he really did corrupt the youth of Athens by motivating them to question the social ethic and arrive at their own definition of individual responsibility.

social ethic) to reflect upon themselves as individuals. If Kierkegaard would have told us directly what was required to shatter moral illusions, his guidelines would have served the same role as the social ethic. Each individual must undergo the process of questioning and self-reflection. The ironist does not reject the existing order, she rejects its deification. She raises the question what is ethical. The shattering of moral illusions gives one the ability to raise ethical questions. The aim is not to reject social responsibility, but to raise the level of reflection upon it.

But Kierkegaard is not satisfied with the Socratic paradigm. He thinks there is still more to say about moral disillusion. The figure who represents a further step in the process of shattering moral illusions is Abraham. *Fear and Trembling* is often misinterpreted as a call to abandon ethics for faith in God. But, as I mentioned above, the fact that some works are by pseudonyms is telling. *Fear and Trembling* is authored by Johannes de Silentio rather than Kierkegaard. Silentio gives us important information about moral disillusion, but he does not give us the full picture. We can interpret Silentio's interpretation based on how he relates to Abraham. Silentio describes Abraham as the knight of faith, but tells the reader that he himself is not the knight of faith.

> I cannot make the movement of faith. I cannot shut my eyes and plunge confidently into the absurd; it is for me an impossibility, but I do not praise myself for that…. [F]or the marvelous I cannot do—I can only be amazed at it.[223]

Silentio is not the knight of faith, but describes the knight of faith as someone he admires yet cannot emulate. In the "Exordium", Silentio presents different

---

[223] Kierkegaard, *Fear and Trembling*, ed. and trans. Howard V. Hong and Edna H. Hong (Princeton, Princeton University Press, 1983), 34-37.

versions of the story to see which way coheres with his admiration of Abraham. Kierkegaard is trying to show the reader through Silentio's admiration of Abraham and his various interpretations that the story need not be understood literally. Rather than taking the story literally, it serves as an allegory describing the process of shattering moral illusions.

In *Fear and Trembling* we become acquainted with two kinds of knights: the knight of infinite resignation and the knight of faith. We can think of each knight as a stage in the process of moral disillusion. Before one can become the knight of faith one must become the knight of infinite resignation.[224] This resignation is the teleological suspension of the ethical. By "ethical" Silentio means the social ethic. It is symbolized by Abraham's attempted sacrifice of Isaac. The resignation then is the questioning (or sacrifice) of everything one has been told or taught is good and right. For example, one might question the ethical validity of the golden rule, the categorical imperative, or the greatest happiness principle and realize that such abstract universal principles do not adequately cover the nuances of moral life. One must question everything. But this process is not easy. As Silentio says, "Only the one who works gets bread."[225] This old adage is supposed to reflect the notion that shattering long-held moral illusions is quite difficult work. After one questions the social ethic in its entirety, one experiences "fear and trembling" due to the realization that the grounds for one's moral responsibility are now in one's own hands. The individual can no longer rely on the social ethic to define moral responsibility. The peace and security of

---

[224] Ibid., 37.
[225] Ibid., 27.

the social ethic has been shattered and one finds oneself utterly alone in defining responsibility. One shatters one's moral illusions. This shattering through self-reflection is the teleological suspension of the ethical.

The taking of Isaac to Mount Moriah to sacrifice him is a metaphor that shows one must be willing to question or "suspend" all aspects of the social ethic even something as basic as a father's love for his son. The story is not intended to support actually killing one's son, but to symbolize the thoroughness with which one must question the social ethic. "Faith" does not mean simple obedience to God. In three of the "Exordium" in *Fear and Trembling*, Silentio has Abraham obey the command to kill Isaac, but then mocks these tellings as false. He also rejects the person who takes Isaac too hastily to Mount Moriah. Recommending obedience in itself is not the point of Silentio's version of the story. If Abraham is the knight of faith, and his faith is not exemplified by obedience, then "faith" means something different than simple obedience for Kierkegaard. "Faith is not a doctrine but awakening…. [F]aith…means moving over the abyss without attachment to what I know or will as good."[226] The focus of Silentio's telling of the story is not only what Abraham does or does not do, but the quality of his will. Abraham turns inward to reflect upon what responsibility means to him. He raises the knife to Isaac's throat with the "faith" that he will get him back. This part of the story relates metaphorically to the notion of questioning one's inherited sense of responsibility, not with the hope of losing all responsibility, but with the

---

[226] Matuštík, "Becoming Human, Becoming Sober," 265.

intention of deepening it and understanding it more fully; that is, with the hope of getting it back.

To become a self and to become individually responsible, Abraham must relate on a personal level to what is morally required of him. By suspending the ethical and defining responsibility through one's personal commitments, one becomes a responsible individual rather than a simple rule-follower. "Johannes de silentio shakes us from the idea that the key to moral or religious earnestness is in the public sphere where we discover and can then conform to a lucid list of rules."[227] Abraham has responsibilities to his son, not because the role of father has been defined for him by society, but because he questioned this definition and "returned" with Isaac. The father/son relationship has not been destroyed. The message of the allegory is that the relationship has been redefined.

One might think that if the allegory encourages a suspension of the social ethic, then it also encourages its rejection. But suspension and rejection are distinct. The suspension of the ethical (i.e. of the social ethic) is not undertaken with the goal of being a hermit outside of society. Through resignation one always has the intention of further existing in society, but with a renewed sense of self. Socrates did not leave Athens, but remained there to serve as a gadfly to his fellow citizens. Abraham returned with Isaac from Mount Moriah. As Johannes Climacus tells us:

> The [ironic/negatively free] individual does not cease to be a
> human being…. His life, like that of another, has the diverse

---

[227] Edward F. Mooney, *Selves in Discord and Resolve: Kierkegaard's Moral-Religious Psychology from "Either/Or" to "Sickness Unto Death"* (New York: Routledge, 1996), 51.

predicates of a human existence, but he is within them like the person who walks in a stranger's borrowed clothes. He is a stranger in the world of finitude, but he does not define his difference from worldliness by foreign dress (this is a contradiction, since with that he defines himself in a worldly way); he is incognito, but his incognito consists in looking just like everyone else.[228]

For Kierkegaard it is important that the person who shatters moral illusions does not leave society to live as a hermit, either literally or metaphorically. Some interpreters have accused him of arguing as much. But this is a misunderstanding. Socrates and Abraham look like everyone else, but they are inwardly different because they have defined responsibility in terms of their own agency, rather than in terms of an abstract moral law or the social ethic. The story of Abraham does not, therefore, support religious fanaticism. To dress in monastic robes or to become a religious fanatic is to further rely on external doctrine to define one's understanding of morality and responsibility. The morally disillusioned individual takes on a personal and passionate approach to responsibility.

This return to society is why Silentio says that the knight of faith resigns everything with the knowledge he will get it back. The goal of the knight of faith is not solitude but a more harmonious and more reflective society.

> [F]or Kierkegaard, inwardness requires that the individual become more passionately engaged with others…. [W]ith this greater level of self-awareness the subject does not see his or her objectives only in terms of one's social or state role.[229]

---

[228] Kierkegaard, "Concluding Unscientific Postscript" in *The Essential Kierkegaard*, 410.
[229] Dooley, 49.

The knight of faith does not resign, or suspend, the ethical with the goal of isolated individualism. His goal is heightened sense of self and a more passionate and responsible point of view.

This personal understanding of one's responsibilities leads Johannes Climacus in *Concluding Unscientific Postscript* to say that "truth is subjectivity."[230] But this claim should not be interpreted as espousing some sort of relativism. What Climacus means is that what is important is not only what is true, but also how one relates to what is true. It might be true that a father should love his son, but if the father has not reflected upon this role, questioned it, and then accepted it personally as a responsibility, then the fact itself is not important. The responsibilities one has must be significant to the individual. In this way what would otherwise be disinterested objective truth has become interested subjective truth. I will show below that Kierkegaard's description of being interested in ourselves and his prescription that we ought to be interested in our responsibilities to others coincide with Frankfurt's description of second-order cares and his prescription that we ought to care about some things and not others.

By seeing the Abraham story as an allegory, we do not have to make the absurd conclusion that Kierkegaard advocated following a simple divine command ethics. The "faith" that characterizes the knight of faith need not be religious faith as we ordinarily understand it. By "faith" Kierkegaard means the process I have here described. "Faith" is reflecting upon oneself, questioning the social ethic, and coming back to the society not as an isolated individual, but one

---

[230] Kierkegaard, "CUP" in *The Essential Kierkegaard*, 157-175.

who has more passionate, reflective relationships with others. To define faith as belief in a specific doctrine would contradict the need for each individual to undergo the ordeal of self-reflection. Getting ethical answers from religious doctrine is similar to getting them from the state or social ethic. Therefore, when Silentio speaks of the knight of faith, he refers to the individual who has risked the social ethic for a more complete, personal, and individual moral responsibility.

Kierkegaard's theory of selfhood requires the individual to question or suspend the social ethic in which she lives, accept the fear and trembling that this suspension entails, and reflect upon how one should relate to others in society. One must shatter the illusion that moral responsibility is something given rather than taken. His is not a theory of irrational subjectivism as some have accused him of, but rather one of radical individual responsibility. Understood within the context of his entire authorship we can understand Kierkegaard's philosophy as one which hopes to motivate the reader to reflect upon what is involved with being a self devoid of moral illusions.

So far I have shown that Kierkegaard urges us to develop a passionate inwardness through self-reflection. One essential component of this reflection is despair. Despair not only plays a central role in his explanation of moral disillusion, it connects up with my argument in chapter 4 that the morally disillusioned agent recognize that she can be a source of unchosen evil.

In *The Sickness Unto Death* Anti-Climacus tells us that the self is the relation of itself to itself. "The self is a relation that relates itself to itself or is the

225

relation relating itself to itself in the relation".[231] Although this sentence seems at

first glance to be nonsense, Anti-Climacus is trying to tell the reader that a self,

that is, one who shatters moral illusions, is one who becomes a self through self-

reflection. The individual relates itself to itself, or, in other words, engages in self-

reflection. A "human" is not yet a "self." The use of these terms is Kierkegaard's

way of saying that an agent can hold fast to moral illusions or shatter them. When

a human shatters her moral illusions, she becomes a self. It follows that moral

disillusion is not a primary state of existence. It must be realized by striving and

transcending one's context. The individual must relate itself to itself. One must

become self-reflective. In the previous chapters I have also argued that one

shatters moral illusions by reflecting upon oneself and one's relationships with

others.

But just because one begins the process of reflecting upon one's moral

responsibilities, it does not follow that one completes the process fully or that one

actually becomes more responsible. For example, one might suspend the social

ethic only to seek answers in religious fanaticism. Or one might suspend the

social ethic and not return to take up one's responsibility in a more dedicated

fashion. Or one might accept that performing evil is often unchosen and

inescapable, and thus conclude that there is no use fighting against it; we should

embrace our tendencies to do evil. That is, one might become demonic by

shattering moral illusions.[232] Finally, one might realize that the universe does not

---

[231] Kierkegaard, "The Sickness Unto Death" in *The Essential Kierkegaard*, 127
[232] Below I address this possibility directly.

reward moral merit or have an underlying rational harmony and collapse in despair.

Kierkegaard, through Anti-Climacus, gives us a way of understanding one's radical responsibility and dealing with the despair that may accompany the understanding. It involves some reflection upon the nature of the fallibility of one's will. This reflection matches up with my comments in chapters 4 and 5 on the recognition of one's agency as a source of unchosen evil and moral tragedy. Understanding these moral facts in terms of one's own agency can lead to despair. The one way of rooting out despair is to accept the fact that one's will creates obstacles to one's willing. That is, the only cure for despair is to learn to despair properly. "To despair soberly is to know oneself as despairing not over my failed body or mind or over this or another loss of things and persons but as despairing of myself."[233] To despair in this way means to become fully aware of one's potential for self-destruction. One of the reasons that Oedipus blinds himself is because he is surprised by the realization that he was the cause of his own downfall. However, if the morally disillusioned individual understands that the capacity for this destruction resides within oneself, then the experience of it is less tragic. I will discuss below that Kekes makes a similar argument in *Facing Evil*. When one understands that one is a source of unchosen evil, one is less likely to experience unchosen evil as a tragedy. Similarly for Kierkegaard, when one acknowledges that the capacity to cause evil to others and to oneself lies within one's own will, then one has shattered illusions and despairs correctly.

---

[233] Matuštík, "Becoming Human, Becoming Sober", 267.

4.   The Importance of What We Care About

In the previous section I presented Kierkegaard's philosophy of moral disillusion. Along the way I mentioned some of the ways in which his conception of moral disillusion is consistent with mine. But Kierkegaard's philosophy of moral disillusion is not merely descriptive. It also prescribes that one ought to undergo the steps to shatter moral illusions, or in his words to become a self, and to take responsibility for who one is. Kierkegaard tells us that this process involves a passionate inwardness. Becoming a self involves being interested in who one is. "When I care for myself not just aesthetically, as if from a poetic distance or through an idea of myself, but consider the entirety of my life, then I am transformed by existential pathos."[234] Again, this is why Climacus says that truth is subjective. What is important is how one relates to the truth and what facts mean to the individual. In the present section I show that his account of passionate inwardness is congruent with Frankfurt's account of caring about what we care about. Kierkegaard argues that we ought to develop passionate inwardness and Frankfurt argues that we ought to care about some things instead of others. Throughout the previous chapters I have argued that what we care about is not always entirely under our control. This raises the question of how we can affect our cares if they are not entirely within our control. The answer to this question lies in the increased understanding of moral life that constitutes moral disillusion. Finally, out of this discussion will reemerge the claim that moral disillusion does

---

[234] Matuštík, "Becoming Human, Becoming Sober," 252-253.

not entail moral goodness. However, it does entail an increased moral responsibility.

Frankfurt says that the formation of a person's will is primarily a matter of her coming to care about things and of her coming to care about certain things more than others. These processes may not be entirely under one's control. Although they may not be entirely under one's control, Frankfurt claims that it is often possible to affect them. Further, we can distinguish between things that are worth caring about and those that are not. However, in his early work, he is unclear about how to go about affecting one's cares. If we can discern the kinds of things a person ought to care about, but it remains unclear how a person ought to affect her cares, then it follows that it is difficult to prescribe practically that one ought to care about certain things.

First, we need to establish that it is logically coherent to claim both that one's cares are not entirely within one's control and also that one ought to care about certain things. Although one's cares are partially constituted by external circumstances one cannot control such as place of birth, upbringing, social milieu, etc., it does not follow that anything goes.

> It may still be possible to distinguish between things that are worth caring about to one degree or another and things that are not. Accordingly, it may be useful to inquire into what makes something worth caring about – that is, what conditions must be satisfied if something is to be suitable or worthy as an ideal or as an object of love – and into how a person is to decide, from among the various things worth caring about, which to care about. Although people may justifiably care about different things, or care

differently about the same things, this surely does not mean that their loves and their ideals are entirely unsusceptible to significant criticism of any sort or that no general analytical principles of discrimination can be found.[235]

I am in complete agreement with Frankfurt here. We should be able to evaluate the content of a person's cares. The normative evaluation of what a person cares about is especially important. Kierkegaard likewise holds the position that one ought to cultivate passionate inwardness; that is, one ought to care about reflecting on one's responsibilities to others. But how can we prescribe what one ought to care about, if one's cares are not entirely under one's control?

Frankfurt makes a distinction that can provide grounds for evaluating the worthiness of a person's cares. He notes that it is important to the person what she cares about, and there are two distinct ways in which something may be important to a person. First, its importance may be due to considerations that are independent of whether or not she cares about the thing in question. Second, the thing may become important to the person because she cares about it.

> Correspondingly, there are two distinct sorts of ground on which a person who thinks it worthwhile to care about a certain thing might attempt to justify his view. He might claim that the thing is independently important to him and that it is worth caring about for this reason. Or he might maintain without supposing that the thing is antecedently important to him at all, that he is justified in caring about it because caring about it is itself something that is important to him.[236]

---

[235] Frankfurt, *The Importance of What We Care About*, 91.
[236] Ibid., 93.

A person can reflect upon things that affect her and care about them because of the thing's capacity to affect her. She may also cease caring about those things that do not have this capacity. But this reflection and beginning or ceasing to care relates only to the first way of caring. The person recognizes the thing as independently important (or not important) and changes her cares accordingly. However, if the thing is important to the person because she cares about it, one can only justify the caring in terms of the activity of caring as such. Caring about what one cares about is important.

The different ways of caring are important to us. They are important to us because these serve "to connect us actively to our lives in ways which are creative of ourselves and which expose us to distinctive possibilities necessary for freedom"[237]. We are devoted to what we care about. The cares we have provide constraints on our choices. These constraints are liberating because they guide our action and establish our identity. Without these constraints (or necessities) we would have too many choices all of which would be meaningless. There would be no order or cohesion to our lives. Such chaos is detrimental to autonomy. Hence, caring about things is constitutive of autonomy.

In "The Importance of What We Care About" Frankfurt argues that because the act of caring is valuable in itself, it follows that even if what one cares about cannot be justified, the very act of caring is justified. Yet, we also want to be able to criticize someone for caring about the wrong sort of thing. He says we want to be able to criticize someone for caring about something that is not worth

---

[237] Ibid., 93.

caring about. I would add that we want to able to criticize someone for caring about the morally wrong thing.

> Suppose, for example, that what a person cares about is avoiding stepping on the cracks in a sidewalk. No doubt he is committing an error of some kind in caring about this. But his error is not that he cares about something which is not really important to him. Rather, his error consists in caring about, and thereby imbuing with genuine importance, something which is not worth caring about. The reason it is not worth caring about seems clear: it is not important to the person to make avoiding cracks in the sidewalk important to himself. But we need to understand better than we do just why this is so…[238]

Frankfurt ends the essay with some vague and unsatisfying comments about divine love. The question still remains what criteria we have for criticizing someone's cares. What principles do we have for saying that what someone cares about is not worth caring about? If the reason were clear, then we would not need to understand better than we do.

As I will now make clear, Frankfurt readdresses this issue in some of his later work, and helps orient our thinking in terms of how we are justified in criticizing what one cares about. He argues that a person should try to know as much as possible about the things she cares about. Knowing about what one cares about is motivated by caring about what one cares about.

> Once we have learned as much as possible about the natural characteristics of the things we care about, and as much as possible about ourselves, there are no further substantive corrections that

---

[238] Ibid., 93-94.

can be made. There is really nothing else to look for so far as the normativity of final ends is concerned. There is nothing else to get right.[239]

Because [p]eople naturally want the things they care about to coincide,"[240] if they learn as much as possible about them and themselves, then they can affect their cares. They might decide that they actually do not care about something as much as they thought they did. Or they might realize that something they cared about is not important to them after all. The things one cares about should coincide up to a point. By caring about what one cares about, and subsequently learning as much as possible about what one cares about and about oneself, one can sometimes affect one's cares. This ability to affect our cares comes from understanding. Understanding should influence what we care about. As I discussed in chapter 2, this kind of understanding is an ability. Here we see this ability as an ability to make-up one's will by affecting one's cares. This ability is partially comprised by seeing that moral life is richer than the minimum conditions of moral responsibility. That is, the understanding I have described throughout the previous chapters can affect what one cares about.

Kekes makes a similar argument as well. He suggests that we ought to develop "the reflective temper." The first step in doing so is epistemic. We must understand the essential conditions of life. As I presented in chapter 4, our projects are vulnerable to evil regardless of our personal moral merit. Understanding this vulnerability is not merely descriptive; we also care about our

---

[239] Harry Frankfurt, *Taking Ourselves Seriously and Getting it Right*, ed. Debra Satz, (Stanford: Stanford University Press, 2006), 50.
[240] Frankfurt, *The Importance of What We Care About*, 93.

lives. This understanding is the ground of true hope. Kekes illustrates his point that we need to cultivate a deeper understanding about life by noting that what is tragic about Oedipus, Lear, and Kurtz is the expectation that morally praiseworthy actions will succeed and not come into conflict with the essential conditions of life. We expect that contingency, indifference, and destructiveness will be overcome by reason and decency. But the intrusion of evil will not come as so devastating if we expand our understanding and realize that evil is often inevitable.[241] This depth of understanding improves our lives because we will not collapse under the grief of tragedy. If we can cultivate this understanding before tragedy occurs in our lives, then we need not undergo immense anguish in the face of tragedy.

We cannot control the contingent, indifferent, and destructive nature of the universe and of humans. We can, however, control our response to these factors.[242] By understanding that evil is inevitable we abandon the unrealistic expectation that our lives will not conflict with the essential conditions of life. This understanding has emotional and motivational effects. By acquiring depth, we gain more control over our own reactions to calamity.

> This understanding has both an individual and a general dimension. Individually, it is understanding that our personal aspirations are subject to the vicissitudes of contingency, indifference, and destructiveness, quite independently of our merits as agents or of the merits of our projects. Generally, it is understanding that what is true of ourselves is also true of

---

[241] Or, to despair correctly, as Kierkegaard prescribes.
[242] Kekes's argument owes much to the Stoics.

humanity, that all human projects are in jeopardy because of the existence of evil.[243]

Once we have deepened our understanding of evil, we are prepared to respond appropriately to it. Kekes calls this response the "reflective temper." Enlarged understanding gives us the truth, but the reflective temper gives us the response. Through reflection we can become aware of when and how the essential conditions of life express themselves through our various psychological states and try to diminish their expression when possible. The act of reflection alters the psychological states by reflecting on them. That is, self-reflection alters the self that is reflected upon because the self has become reflective. This reflection and understanding do not free us from the essential conditions of life, but increase control. By understanding how we bring about unchosen evil, we may be able to minimize it. Increased control expands the area in which we can make choices. Thus we have some control (it may be quite minimal in certain people) over how much evil we cause in the world. We can never eradicate evil in the world, but we can minimize it.

However, there is no guarantee that the person who cares about what they care about, who cultivates passionate inwardness, who reflects upon her own character, will care about moral goodness or be morally good. Caring, understanding, and reflection do not entail moral goodness. Becoming a self or shattering illusions of moral purity do not always bring about a morally good individual or morally good acts. What do we say about the person who

---

[243] Kekes, *Facing Evil*, 184.

understands (i.e. has shattered moral illusions) and does not care that she is an agent of unchosen evil or that she may be in situations of inescapable evildoing?

All we can do is urge people to understand what they care about, to reflect upon their responsibilities to others, and to understand the richness of moral life. One can be urged to do so by reasoned argument, as Frankfurt, Kekes, and Kierkegaard have done. We have seen that there are good reasons to understand what you care about or to reflect upon your life. But what one ends up caring about cannot always be affected by reasons. If it could be, then we would have far less evildoing in the world. We could simply present an evildoer with reasons why she should not do evil, and she would be persuaded. However, people do not act solely based on reasons. An individual might be rational, but still unreasonable. Because the root of this unreasonableness is the will, Frankfurt calls it "volitional irrationality".

> What is at stake here is not a matter of avoiding mistakes and getting things right. The volitionally irrational lover of death or disability or suffering has not overlooked something, or misunderstood something, miscalculated, or committed any sort of error. From our point of view, his will is not so much in error as it is deformed. His attitudes do not depend upon beliefs that might be demonstrated by cogent evidence or argument to be false. It is impossible to reason with him meaningfully concerning his ends, any more than we could reason with someone who refuses to accept any proposition unless it is self-contradictory. Many philosophers believe that an act is right only if it can be justified to other rational beings. For this to be plausible, it is not enough that the rationality of the others be merely of the formal variety. Those

> whom we seek to convince must be volitionally rational as well. If
> they are not, then their practical reasoning—however firmly
> correct it may be—builds upon a foundation that is radically
> opposed to ours…. We can therefore do no more with them than to
> express the bewilderment and revulsion that are inspired by us by
> the grotesque ends and ideals that they love. [244]

Understanding the things we care about may reveal conflicts among them. Our

carings may be shown to be misguided. In this way our cares may need

correction. Furthermore, people are often wrong about what they care about.

Understanding themselves can correct misguided cares. Sometimes what one

person cares about conflicts with what another person cares about and reason

cannot resolve the conflict. This is sad, but true. Although sometimes cares can be

affected by reasons, at some point an appeal to reasons fails to affect what one

cares about because one cares about it. If conclusions about moral behavior could

be agreed upon solely through formal cognitive rational grounds, then many

ethical debates could be resolved. We would no longer need to debate, for

example, whether abortion is morally justified because everyone would be

convinced by reasons for or against it. The reason that many disagreements exist

is due to people's differing volitional commitments. People who are pro-life

usually cannot be persuaded that abortion is morally justified because they are

volitionally opposed to accepting the rational reasons supporting the practice.

More generally, it is not that we should not use reasons to try and convince others

to be morally good; it is that reasons are not always sufficient to convince others

---

[244] Frankfurt, *Taking Ourselves Seriously and Getting it Right,* 39-40.

to be morally good. The person must also care about being morally good. Hence, we must shatter the illusion that everyone has the ability to be good.

For these reasons, moral disillusion does not entail moral goodness. An individual can understand the nuances of moral life, but either not care about performing morally good acts, or care about performing evil ones. There is no guarantee that the morally disillusioned individual will be morally good in deed or character. Moreover, as I discussed in chapter 4, a person can even hold the belief that one ought to be morally good, but still not care about being so. That is, one can acquire the understanding without having the corresponding care. As I concluded in chapter 4, because actions are often more powerfully motivated by aliefs and cares than beliefs, we ought to track moral responsibility based on cares rather than beliefs.

And yet if one is morally disillusioned, one is capable of caring about moral goodness in a more realistic manner than those who still retain some illusions. The morally disillusioned individual has all of the cognitive tools needed to affect her cares. As Frankfurt, Kierkegaard, and Kekes all point out, a deeper understanding of moral life can affect what one cares about. Questioning the social ethic, or the nature of one's responsibilities to others, can affect what we care about. We can best care about how our actions affect others when we understand all of the ways in which they do so. This is one of the reasons that moral disillusion is an ability. What should we say about those who possess this ability but act immorally?

238

When one understands the richness of moral life but does not care about it, then one is more morally blameworthy than someone whose lack of care is based on a lack of understanding. The morally disillusioned person who performs evil has the ability to minimize evildoing by caring about it. Because the will she wants to have is one that does not care about moral goodness, we consider her morally worse than someone one does not care because she does not understand. We see this difference between those who should have known better (i.e. did not understand, but could have), and those who did know better but still brought about evil (i.e. those who did understand and did not care). I argued in chapters 3 and 4 that evil does not require the intention to do evil. One can also do evil thoughtlessly, inadvertently, or indirectly. But we consider an agent more morally blameworthy for maliciously and intentionally bringing about evil. Similarly, moral responsibility does not require moral disillusion. However, a person is more blameworthy for performing wrong and evil acts when morally disillusioned. Moral disillusion is an understanding that is an ability. Those who have the ability to minimize evil, but do not guide themselves in a way to do so are more blameworthy than those who lack the ability because they lack the understanding.

Recall the story of *Billy Budd*. There are two characters in the story who commit immoral acts. Claggart targets Billy for his schemes because he does not like Billy's innocent demeanor. He is morally blameworthy for falsely accusing Billy of mutiny. Captain Vere sentences Billy to death. But he does so fully aware, not only of Billy's innocence, but that he is in a situation of inescapable wrongdoing. He fully understands that he has sentenced an innocent person to

239

death. Despite understanding his responsibility to Billy, he did not care enough about it. The lack of care guides Vere's actions. If he had cared in a certain way, he would not have been able to go through with sentencing Billy to death. He would have been strongly motivated to save Billy's life. Captain Vere is more morally blameworthy than Claggart because he was morally disillusioned but still brought about evil. Claggart was surely a bad person, but he lacked a nuanced understanding of moral life. Vere had this understanding, and thus more ability than Claggart to minimize evil. He is morally responsible for not doing so.

When such a moral belief/behavior mismatch is present in an agent, we morally evaluate the agent not just on her behavior, but on the mismatch. This framework allows us to see why someone like Thomas Jefferson deserves our moral censure. Jefferson wrote the inspiring words in the preamble to the Declaration of Independence proclaiming that "all men are created equal." Yet, he owned hundred of slaves, unwilling to free them even upon his death. He is considered an American hero, and based upon his beliefs, this status is warranted. But based upon his actions, motivated by what he cared about, he deserves our moral condemnation. What makes Jefferson's slave owning so horrible, more horrible than merely owning another human being, is that he believed that owning slaves was morally wrong. He had the understanding and the ability to behave in a morally praiseworthy manner. But he did not care to do so. He is different from slave owners who did not hold the belief that all men are created equal. These slave owners are certainly morally blameworthy for owning slaves, but Jefferson is more blameworthy. It is not merely that he should have known better. Rather,

he did know better, but he did not care. If moral evaluation were only to track beliefs, then Jefferson does not deserve our criticism. But he does deserve our criticism. Therefore, moral evaluation ought to track what a person cares about.

5. Conclusion

In this chapter I have shown that Kierkegaard's philosophy is an urging to shatter moral illusions. I also argued that the passionate inwardness he describes is congruent with Frankfurt's account of caring about what we care about. Both Kierkegaard and Frankfurt endorse deepening our understanding about what is important to us. This more nuanced understanding is moral disillusion. Although moral disillusion gives one the ability to be morally good, it does not entail it. The morally disillusioned can perform evil acts. However, because they have a deeper understanding of the nuances of their moral agency and its effects on others, they have an ability to take responsibility for, and potentially minimize, evil. When the morally disillusioned have this ability, but do not care to exercise it, they are more morally blameworthy than those who have yet to shatter their moral illusions.

CHAPTER 8

CONCLUSION

In the preceding chapters I have described important phases of moral disillusion. I described the loss of moral innocence, the kind of understanding that accompanies disillusion, the nature of evil, unchosen evil, inescapable wrongdoing, and the importance of our cares and self-reflection. I concluded that moral disillusion develops the ability to take a critical stance towards ethics and raise questions as to the adequacy and suitability of moral directives. Holding fast to moral illusions can lead one to do dreadful things. By acting, not only contrary to a moral system, but in accordance with one, an agent can fail to see the evil she causes. Hence, following moral directives is only one aspect of moral behavior. Putting moral principles to question is equally, if not more, important. When one recognizes how evil comes about, one has the possibility of taking responsibility for it, and even minimizing it.

Moral disillusion is a scalar process. We enter the moral community by recognizing that we are capable of wrongdoing and are appropriate candidates of the reactive attitudes. We lose moral innocence, but gain the ability to participate in the moral community. We see that our actions have upshot and moral significance. But there is more to moral life than satisfying the minimum conditions of moral responsibility. We can also shatter the illusion that only malicious individuals do evil deeds. We see that we can exploit, either directly or indirectly, another's extreme vulnerability. We can be agents of unchosen evil, or bring about evil despite pursuing the morally best course of action, and be morally

242

responsible for doing so. Understanding moral responsibility for unchosen evil and for inescapable moral wrongdoing is based on understanding the importance of what we care about. Our cares are powerful motivators of our action. They can motivate us to act with or without our explicit cognitive consent. Moreover, in situations where what we care about conflicts with what we believe, the care often proves to be the more powerful motivator. We can affect our cares by understanding what we care about. Conversely, we can also come to care about things because we understand them. Moral disillusion is a shattering of moral illusions; an understanding of the nuances of moral interaction and moral responsibility. This understanding is an ability to raise questions as to the (in)adequacy of moral principles and directives, and to take responsibility for the evil one causes.

Degrees of this understanding correspond to degrees of moral responsibility. What characterizes the morally innocent is a certain kind of ignorance, or inability to participate in the moral community. The morally innocent cross the threshold into moral responsibility when they understand that their actions have moral significance. The more one understands about the moral significance of one's actions, the more morally accountable one is. It is an illusion to believe that one can cultivate moral purity by following a moral directive. One can be an agent of unchosen evil or transgress a moral value even when pursuing the morally best course of action. Reflecting upon oneself and one's relationships with others allows one to see the complexities of moral life. One understands that one cannot always avoid wrongdoing or evildoing. Ironically, it is the recognition

243

that evil and tragedy are inevitable that provides the means to minimize them both. The morally disillusioned person understands this. Whether she cares about it or not is a separate issue.

Taking responsibility for the evil one causes is important. If we only take responsibility for wrongs that we bring about as a result of deliberate choice, then we ignore large portions of what we are morally responsible for. Moreover, understanding that one's actions have upshot in the world and that one is an appropriate candidate of the reactive attitudes is a scalar process as we discover the many different kinds of upshot, and the varied attitudes that are appropriate in different circumstances. Taking responsibility for unchosen evil and inescapable wrongdoing, along with questioning moral principles, allows one to respond appropriately to the evil one causes. By understanding how one causes evil, one can potentially minimize it. We cannot do anything about that of which we are ignorant. Hence, shattering the illusion of moral purity is essential to minimizing the harm one causes. When the harm one causes is unchosen and inescapable, then it cannot always be minimized. In such circumstances the appropriate response is regret, remorse, confession, contrition, apology, restitution, or reparation. Holding fast to illusions of moral purity keeps one blind to the ways in which we cause evil. Hence, we must shatter moral illusions for the sake of taking responsibility.

REFERENCES

Amodio, D., E. Harmon-Jones, and P. Devine, "Individual Differences in the Activation and Control of Affective Race Bias as Assessed by Startle Eyeblink Response and Self-Report." *Journal of Personality and Social Psychology* 84, no. 4 (2003): 738–753.

Anscombe, G.E.M. "Modern Moral Philosophy." *Philosophy* 33, no. 124 (1958): 1-19.

Aristotle. *The Complete Works of Aristotle Volume Two*. Edited by Jonathan Barnes. Princeton: Princeton University Press, 1984.

Auden, W.H. *The Dyer's Hand and Other Essays*. New York: Random House, 1954.

Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. New York: Routledge, 1998.

Austin, J.L. "A Plea for Excuses." In *Philosophical Papers*. Oxford: Oxford University Press, 1961.

Card, Claudia. *The Atrocity Paradigm: A Theory of Evil*. Oxford: Oxford University Press, 2002.

___."Torture in Ordinary Circumstances." In *Moral Psychology: Feminist Ethics and Social Theory*, edited by Peggy DesAutels and Margaret Urban Walker, 141-161. New York: Rowman and Littlefield, 2004.

Dooley. Mark. *The Politics of Exodus: Kierkegaard's Ethics of Responsibility*. New York: Fordham University Press, 2001.

Feinberg, Joel. *Problems at the Roots of Law: Essays in Legal and Political Theory*. Oxford: Oxford University Press, 2003.

Fischer, John Martin and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 1998.

Fischer, John Martin. *The Metaphysics of Free Will: An Essay on Control*, Aristotelian Society Monograph Series, vol. 14. Cambridge Mass.: Blackwell Publishers, 1994.

Formosa, Paul. "A Conception of Evil." *The Journal of Value Inquiry* 42 (2008): 217-239.

Frankfurt, Harry. *The Importance of What We Care about: Philosophical Essays*. Cambridge: Cambridge University Press, 1988.

___. *Taking Ourselves Seriously and Getting it Right: The Tanner Lectures in Moral Philosophy*. Edited by Debra Satz. Stanford: Stanford University Press, 2006

French, Peter A. *Responsibility Matters*. Lawrence: University Press of Kansas, 1992.

___. "'Seeing' and 'Seeing That', 'Observing' and 'Observing That'." *American Philosophical Quarterly* Monograph Series: Studies in Epistemology no. 9 (1975): 89-97.

___ *War and Moral Dissonance*. Cambridge: Cambridge University Press, 2010.

Gendler, Tamar Szabo. "Alief and Belief." *The Journal of Philosophy* CV, no. 10 (October 2008): 634-663.

Goodin, Robert E. "Exploiting a Situation and Exploiting a Person." In *Modern Theories of Exploitation*, edited by Andrew Reeve, 166-200. Beverly Hills: Sage Publications, 1987.

___. *Protecting the Vulnerable: A Reanalysis of Our Social Responsibilities*. Chicago: The University of Chicago Press, 1985.

Gowans, Christopher W. *Innocence Lost.* Oxford: Oxford University Press, 1994.

Hanson, Norwood Russell. *Patterns of Discovery*. New York: Cambridge University Press, 1965.

Hegel, G.W.F. *Elements of the Philosophy of Right*. Edited by Allen Wood and translated by H.B. Nisbet. New York: Cambridge University Press, 1991.

___. *Philosophy of History*. Translated by J. Sibree. New York: Dover, 1956.

Hong, Howard and Edna Hong, ed. *The Essential Kierkegaard*. Princeton: Princeton University Press, 1980.

Kant, Immanuel. *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Edited and translated by Mary J. Gregor. New York: Cambridge University Press, 1996.

___. *Religion within the Boundaries of Mere Reason*. Edited and Translated by Allen Wood and George Di Giovanni. Cambridge: Cambridge University Press, 1998.

Kekes, John. *Facing Evil*. Princeton: Princeton University Press, 1990.

Kierkegaard, Soren. *Fear and Trembling*. Edited and Translated by Howard V. and Edna H. Hong. Princeton: Princeton University Press, 1983.

Kriegel, Uriah. "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction." *Australasian Journal of Philosophy* (forthcoming).

MacIntyre, Alasdair. *After Virtue*. Notre Dame: University of Notre Dame Press, 1984.

Matuštík, Martin Beck. "Becoming Human, Becoming Sober." *Continental Philosophy Review* 42 (2009): 249-274.

___. "Kierkegaard as Socio-political Thinker and Activist." *Man and World* 27 (1994): 211-224.

___. *Radical Evil and the Scarcity of Hope: Postsecular Meditations*. Bloomington, IN: Indiana University Press, 2008.

May, Larry. *War Crimes and Just War*. New York: Cambridge University Press, 2007.

McKenna, Michael. "The Loss of Innocence and the Things that Remain". *APA Newsletter* 7, no. 2 (Spring 2008): 5-9.

Midgley, Mary. *Wickedness: A Philosophical Essay*. Boston: Routledge and Kegan Paul, 1984.

Milgram, Stanley. *Obedience to Authority*. New York: Harper & Row, 1969.

Miller, David. "Virtues, Practices and Justice." In *After MacIntyre*, edited by John Horton and Susan Mendus, 245-264. Cambridge: Polity Press, 1994.

Mooney, Edward F. *Selves in Discord and Resolve: Kierkegaard's Moral-Religious Psychology from "Either/Or" to "Sickness Unto Death"*. New York: Routledge, 1996.

Nagel, Thomas. *Mortal Questions*. New York: Cambridge University Press, 1979.

Nietzsche, Friedrich. *On the Genealogy of Morality*. Translated by Maudemarie Clark and Alan J. Swensen. Indianapolis: Hackett, 1998.

Putnam, Hilary. *Meaning and the Moral Sciences*. Boston: Routledge and Kegal Paul, 1978.

Rachels, James "Drawing Lines." In *Animal Rights*, edited by Cass R. Sunstein and Martha C. Nussbaum, 162-174. New York: Oxford University Press, 2004.

Richards, Norvin. "Innocence." *American Philosophical Quarterly* 31, no. 2 (Apr., 1994): 157-167.

Sample, Ruth. *Exploitation: What is it and Why It's Wrong*. New York: Rowman & Littlefield, 2003.

Singer, Marcus G. "The Concept of Evil." *Philosophy* 79, no. 308 (2004): 185-214.

Strawson, P.F. *Freedom and Resentment and Other Essays*. London: Methuen and Co., 1974.

Urmson, J.O. "Saints and Heroes." In *Essays in Moral Philosophy*, edited by A.I. Melden, 198-216. Seattle: University of Washington Press, 1958.

van Inwagen Peter. *An Essay on Free Will*. Oxford: Oxford University Press, 1983.

Walker, Margaret Urban. *Moral Understandings*. New York: Oxford University Press, 2007.

Watson, Gary. "Two Faces of Responsibility." In *Philosophical Topics* 24 (1996): 227-248.

Widerker, David and Michael McKenna, ed. *Moral Responsibility Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Burlington, VT: Ashgate, 2003.

Williams, Bernard. *Moral Luck: Philosophical Papers 1973-1980*. New York: Cambridge University Press, 1981.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Upper Saddle River, NJ: Prentice Hall, 1958.

Wokler, Robert. "Projecting the Enlightenment." In *After MacIntyre*, edited by John Horton and Susan Mendus, 108-126. Cambridge: Polity Press, 1994.

Wolf, Susan. "Moral Saints." In *Virtue Ethics*, edited by Roger Crisp and Michael Slote, 79-98. Oxford: Oxford University Press, 1997.

Wolgast, Elizabeth. "Innocence." *Philosophy* 68, no. 265 (Jul., 1993): 297-307.

Zimbardo, Philip. *The Lucifer Effect: Understanding How Good People Turn Evil*. New York: Random House, 2008.