

Empathy, Enhancement, and Responsibility

by

David Gurney

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

Approved April 2016 by the
Graduate Supervisory Committee:

Joan McGregor, Chair
Elizabeth Brake
Gary Marchant
Rebecca Tsosie

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

This dissertation engages with the philosophical, psychological, and scientific literature on two important topics: empathy and human enhancement. My two broad goals are to clarify the role of empathy in ascriptions of responsibility and to consider how enhanced empathy might alter those ascriptions. First, I argue that empathy is best thought of as a two-component process. The first component is what I call the rational component of empathy (RCE). RCE is necessary for moral responsibility as it allows us to put ourselves in another's shoes and to realize that we would want help (or not to be harmed) if we were in the other's place. The second component is what I call the emotive component of empathy (ECE). ECE is usually an automatic response to witnessing others in distress. Expanding on Michael Slote's view that moral distinctions track degrees of empathy, I argue that it is ECE that varies in strength depending on our relationship to specific people.

Second, I argue that in order to achieve Peter Singer's goal an "expanding circle" of care for all human beings, it will be necessary to use some form of artificial empathy enhancement. Within this context, I try to show that empathy enhancement is 1) a reasonably foreseeable possibility within the next decade or so, and 2) morally defensible.

Third, I argue that philosophers who argue that psychopaths are not morally responsible for their actions are mistaken. As I see it, these philosophers have erred in treating empathy as a singular concept and concluding that because psychopaths lack empathy they cannot be held morally responsible for their actions. The distinction between RCE and ECE allows us to say that psychopaths lack one component of

empathy, ECE, but are still responsible for their actions because they clearly have a functional RCE.

Fourth, I paint a portrait of the landscape of responsibility with respect to the enhanced empath. I argue that the enhanced empath would be subject to an expanded sphere of special obligations such that acts that were previously supererogatory become, *prima facie*, morally obligatory.

DEDICATION

For Dave Lund and Kit Christensen, who taught me the value of empathy in philosophy.

ACKNOWLEDGMENTS

I would like to thank my committee for the encouragement they have provided to me throughout the writing process. Not only has their feedback made this dissertation much better than it would otherwise have been, but their overall support has made the road to completing my PhD a bit less rocky.

TABLE OF CONTENTS

	Page
INTRODUCTION.....	1
CHAPTER	
1 HUMAN ENHANCEMENT	11
2 EMPATHY AND THE EXPANDING CIRCLE	58
3 DIMINISHED EMPATHY: AN OVERVIEW OF PSYCHOPATHY	85
4 DIMINISHED EMPATHY: DEFINING THE PLACE OF EMPATHY	118
Part I	119
Part II.....	129
5 ENHANCED EMPATHY: THE POSSIBILITY OF ENHANCING EMPATHY.....	167
6 ENHANCED EMPATHY: MORAL AND LEGAL RESPONSIBILITY	201
Part I	201
Part II.....	217
CONCLUSION	244
REFERENCES.....	254

INTRODUCTION

I think it's safe to say that talk of empathy has occupied a prominent space in public discourse in recent years. Indeed, the President of the United States has consistently referred to empathy as a cornerstone of morality throughout his presidency. In a university commencement speech in 2006, President Obama exhorted students to “cultivate a sense of empathy—to put yourself in other people’s shoes—to see the world from their eyes.”¹ In that same speech, Obama declared, “Empathy is a quality of character that can change the world.”² He went so far to encourage the American people to empathize with our enemies. In an interview with Thomas Friedman, Obama took the politically controversial stance that “even with your adversaries, I do think that you have the capacity to put yourself occasionally in their shoes[.]”³ Obama even coined a term for the lack of empathy that he sees as contributing to so many problems throughout the United States and the world: the “empathy deficit.”⁴

And Obama is not alone. In 2014, Bill and Melinda Gates presented a commencement speech that revolved around empathy. Melinda Gates stated, “empathy []

¹ President Barack Obama, Commencement Address at the University of Massachusetts at Boston (June 2, 2006).

² *Id.*

³ Thomas L. Friedman, *Obama Makes His Case on Iran Nuclear Deal*, N.Y. TIMES (July 14, 2015), http://www.nytimes.com/2015/07/15/opinion/thomas-friedman-obama-makes-his-case-on-iran-nuclear-deal.html?_r=0.

⁴ See Mark Honigsbaum, *Barack Obama and the ‘Empathy Deficit,’* THE GUARDIAN (Jan. 4, 2013), <https://www.theguardian.com/science/2013/jan/04/barack-obama-empathy-deficit>.

tears down barriers and opens up new frontiers for optimism.”⁵ In talking about a poor group of high school students from a bad neighborhood, Gates intoned that “my empathy intensifies if I admit to myself: ‘That could be me.’”⁶ When we consider the enormous wealth (and health) gap between poor and developed nations, in conjunction with technology that allows us to be more and more *aware* of these inequities, the emphasis of empathy makes sense. Peter Singer, in his book *The Expanding Circle*,⁷ asks us to make use of our reason to extend our concern for others from those in our families, towns, states, and countries, to those in far-off places. Yet, it might very well seem that empathy, as opposed to reason, is the best tool for achieving Singer’s goal of an “expanding circle.” I will make this point at several points throughout this dissertation.

Of course, the emphasis on empathy in recent years has not been without its critics. Historian Jeffrey Herf criticized Obama’s stance on empathizing with our enemies. According to Herf, “[empathizing] . . . [by] projecting one’s own feelings onto others . . . is problematic when dealing with the foreignness of foreign affairs and the otherness of other countries because foreigners and others often think about the world very differently than we do.”⁸ Responding to talk of empathy from Obama and other

⁵ Bill and Melinda Gates, Commencement Address at Stanford University (June 15, 2014).

⁶ *Id.*

⁷ PETER SINGER, *THE EXPANDING CIRCLE* (2011).

⁸ Jeffrey Herf, *Empathy Does Not Equal Understanding: Obama on Iran*, TIMES OF ISR. BLOG (April 14, 2015), <http://blogs.timesofisrael.com/empathy-does-not-equal-understanding-obama-on-iran/>.

public figures, Yale psychologist Paul Bloom went so far as to write an article for the Boston Review titled “Against Empathy.”⁹

Complementing the public discourse around empathy, philosophers have taken up the cause of empathy as well. Michael Slote, whom I will discuss at various points throughout this dissertation, has taken upon himself the onerous task of creating an entire moral system based on empathy. As Slote sees it, “all, or almost all, the moral distinctions we intuitively or commonsensically want to make can be understood in terms of—or at least correlated with—distinctions of empathy.”¹⁰

Perhaps even more so than with empathy, recent years have seen extended discussions of human enhancement technologies. In 2003, the President’s Council on Bioethics under George W. Bush presented a report titled *Beyond Therapy: Biotechnology and the Pursuit of Happiness*.¹¹ Written in no small part by anti-enhancement researchers such as Francis Fukuyama and Michael Sandel, the report’s conclusion warned that the prospect of radical human enhancement brings with it “the risk of attacking human limitation altogether, seeking to produce a more-than-human being, one not only without illnesses, but also without foibles, fatigue, or foolishness.”¹²

In this dissertation, I hope to add something of value to the discussion of empathy and enhancement by considering them together, in other words, by considering the

⁹ Paul Bloom, *Against Empathy*, BOS. REV. F. (Sept. 10, 2014), <https://bostonreview.net/forum/paul-bloom-against-empathy>.

¹⁰ MICHAEL SLOTE, *THE ETHICS OF CARE AND EMPATHY* 4 (2007).

¹¹ THE PRESIDENT’S COUNCIL ON BIOETHICS, *BEYOND THERAPY: BIOTECHNOLOGY AND THE PURSUIT OF HAPPINESS* (2003).

¹² *Id.* at 307.

possibility of empathy enhancement. Let me outline the broad goals I have in this dissertation. I aim to show that the best means of achieving Singer's (and others') goal of an "expanding circle" of concern for all human beings is through artificial empathy enhancement. To reach this point, I argue that empathy is best understood as a two-part concept consisting of a *rational* component and an *emotive* component. Generally, I follow Michael Slote's analysis of empathy as a constitutive component of our ascriptions of moral responsibility, but I will argue that conceptualizing empathy in two components shores up some of the concerns with Slote's argument. This is a very (very) broad outline of my dissertation. Let me now break it down into individual chapters so that the reader will have a better idea of the goals and means of my dissertation.

In Chapter One I discuss the concept of human enhancement generally, with emphasis on the philosophical arguments for and against enhancement. I have decided to place this discussion at the beginning of my dissertation because in latter chapters I will be assuming that human enhancement, at least with respect specifically to empathy enhancement, is morally acceptable—indeed, perhaps morally required. Thus, Chapter One attempts to address the concerns raised by opponents of human enhancement technologies such that the latter chapters of my dissertation do not feel ungrounded. While I address many of the arguments that have been hoisted against human enhancement, my overall focus in Chapter One is on the theme presented by many philosophers, that there is something inherently valuable about *the natural* that cannot be replaced—and, indeed, can be undermined—by *the artificial*. I argue that these two concepts cannot be well defined, and even insofar as they can be defined they are inherently intertwined. Thus, simply labeling some human endeavor—say, life

extension—as artificial can never be a sufficient argument against it. In that vein, I follow Allen Buchanan in concluding that with human enhancement technologies we must always consider the gains, in quality of life, that they might provide, in addition to considerations of the damage they might cause.

In Chapter Two I consider Peter Singer’s famous argument that we have the same duty to help those in far-off places as we do to help those in our own neighborhoods, if doing so requires the same amount of effort. In line with Singer, I argue that evolution has placed a limit on the extent to which we *feel* concern for people outside of our inner circle (construed broadly). Where I disagree with Singer is in the solution to this dilemma. Singer seems to believe that reason alone should be sufficient to overcome the evolutionarily derived limits to our concern for others. Interestingly, however, Singer gestures towards empathy—without naming it as such—as a motivating factor in how we allocate our care and concern for others. He says, “One way of arriving at . . . a[n] [ethical] decision . . . is to imagine myself living the lives of those affected by my decisions, and then ask what decision I prefer.”¹³ In Chapter Four I argue that this is a component of empathy. In Chapter Two I begin to argue that empathy is the driving force behind our moral decisions and that if we wish to achieve Singer’s goal of an “expanding circle” of care and concern for others we may need to rely on human enhancement technologies geared towards empathy enhancement. I continue that argument in Chapter Six.

In Chapters Three and Four I switch gears to a discussion of people with diminished empathy—namely, psychopaths. The discussion of psychopathy allows me to

¹³ SINGER, *supra* note 7, at 101.

develop a model of empathy that justifies a baseline standard for holding people responsible for their actions, regardless of how much empathy individuals *feel* towards others. Chapter Three is largely descriptive and provides an overview of the diagnosis of psychopathy as well as a description of how the legal system treats psychopathy.

In Chapter Four I turn my attention to the moral assessment of psychopathy. Specifically, Part I of Chapter Four considers arguments by two philosophers, Jeffrie Murphy and Stephen Morse, that the psychopath is not morally responsible for her actions because she lacks empathy and therefore cannot be motivated to act on the basis of empathy. In Part II, I develop a theory of empathy that allows for the psychopath—and people in general—to be held morally responsible for their actions. I begin Part II with a discussion of the psychological and philosophical literature on empathy. I pay particular attention to Stephanie Preston and Frans de Waal’s account of empathy. Preston and de Waal view empathy broadly, as encompassing all of the “empathic” emotions such as emotional contagion, sympathy, and affective empathy, as well as cognitive empathy.

Relying on Preston and de Waal’s description of empathy, I argue that the concept of empathy is best understood as involving two components. The first component is what I call *the emotive component of empathy* (ECE). ECE is the emotional, affective, response we have towards seeing others in pain. ECE can also be triggered through cognitive processes, per Preston and de Waal’s model. As I argue, psychopaths unquestionably lack ECE. The second component is what I call *the rational component of empathy* (RCE). I derive this component from cognitive empathy as described by Preston and de Waal—but it is an extension of that concept. Cognitive empathy, broadly, is trying to put yourself in the place of another with the hope that an affective empathic response will be triggered.

As an example, parents use this technique in the hopes of triggering affective empathy in their children. RCE does not (always) involve the triggering of an affective response, however. Rather, RCE involves the ability to put oneself in the place of another and realize that one would want to be helped (or not to be harmed) if one were in the other's shoes. I argue that RCE, but not ECE, is necessary for moral responsibility.

In the latter part of Chapter Four, in response to Murphy and Morse, I argue that psychopaths meet the moral responsibility conditions presented by John Fischer and Mark Ravizza. I argue that the psychopath has guidance control; she is moderately reasons responsive; and she feels herself to be a fair target of others' reactive attitudes. I go on to argue that the psychopath's lack of ECE is not relevant to her moral responsibility. The psychopath has an operative RCE and, thus, is capable of understanding that she would not want to be harmed if she were in the other's shoes. In this vein, I accept that there may be some people who lack RCE and that these people—for example, people with severe autism—may not be morally responsible for their actions.

In chapters five and six I return to a discussion of empathy enhancement. Chapter Five is largely descriptive and is meant to provide scientific grounding for the possibility of empathy enhancement. This supplements the moral ground for such enhancement that I argued for in chapters one and two. Thus, in Chapter Five I provide an overview of the current state of brain science and brain manipulation. I argue that technologies such as deep brain stimulation and synthetic chemical technology, in combination with forthcoming results of the United States' BRAIN Initiative and Europe's Human Brain

Project, herald a not-too-distant future where empathy enhancement may become a possibility.

In Chapter Six I imagine that empathy enhancement is available and at least some people have taken advantage of it. I raise two broad questions in Chapter Six. In Part I, I consider whether the enhanced empath has free will. I argue, in part, that the enhanced empath would have free will insofar as she does not feel that her will is being coerced by some outside force. Perhaps more controversially, I argue that coerced empathy enhancement for certain people (say, as a condition of parole for psychopaths) would also not rob the individual of her free will.

In Part II of Chapter Six, I consider what effect, if any, empathy enhancement would have on ascriptions of responsibility. In addressing this question, I engage with Michael Slote's work on empathy. Slote argues that the moral distinctions we make are correlated with distinctions of empathy. I offer the caveat that the moral distinctions we make are correlated with distinctions of ECE and not with distinctions of RCE. This point is important because, as I argue in Chapter Four, it allows us to say that someone who lacks all ECE is still responsible for her actions as long as she has an operative RCE. I use this distinction to argue that special obligations, in other words, obligations to certain individuals that go above and beyond the obligations we owe to people in general, are grounded in distinctions of ECE. Another way of looking at this, as I argue, is to say that when someone has greater ECE towards another, acts that would be supererogatory if directed towards people in general become obligatory with respect to the specific other. I go on to argue that the empathy enhancement procedure would target ECE and not RCE. Thus, I conclude that the enhanced empath might have an increased sphere of special

obligations since her ECE would attach to more people. However, following Michael Hardimon, I offer the caveat that any new obligations we might impute to the enhanced empath would only stand if they were determined to be *reflectively acceptable*.

In the Conclusion of my dissertation, I take a stab at addressing one of the potential counterarguments to my argument in Chapter Six, namely, that the enhanced empath might be subject to an expanded sphere of special obligations. The counterargument, as I see it, is that the enhanced empath might be seen as an *interloping* empath. The idea is that special obligations (or instances where actions that would be supererogatory become obligatory) are historical in nature. In other words, while empathy is the key element that gives special obligations their power (or so I argue in Chapter Six), the genesis of that ECE is significant. Thus, a parent has greater ECE towards her children *because* she is their parent—she birthed the child (or adopted it), and developed an attachment over time. As another example, a soldier in a squad has greater ECE towards her cohorts *because* of their shared struggle and training. The enhanced empath might suddenly feel the same level of ECE towards a child as does the child's mother, or towards a soldier as does a member of that soldier's squad. However, because that ECE is lacking in historical genesis the enhanced empath has no *right* to the greater moral obligation that normally attaches to greater degrees of ECE.

In the Conclusion I offer a tentative response to this concern based on analogies to other minority groups who have been denied equal standing based on the argument that they have no right to the privileges and obligations that attach to a way of life that they have no historical connection to. I view this reasoning as a sketch of what might constitute a response to the concern of the interloping empath—not as a fully developed

argument. My hope is that this concern, along with other issues raised by my dissertation, will form the basis for further philosophical discussion regarding the moral status of the enhanced empath.

Chapter One

HUMAN ENHANCEMENT

Introduction

In this chapter and the next, I will attempt to place the issue of empathy within the context of human enhancement technologies. I will begin with a brief overview of human enhancement technologies, including historical context. I will discuss various arguments against enhancement and present counterarguments. The purpose of this chapter is to provide context and grounding for the chapters that follows.

Human Enhancement Technologies

When we hear the term “human enhancement,” it is easy to call to mind images from dystopian science fiction: the genetically determined beings of Aldous Huxley’s *Brave New World*,¹ or the clones in Philip K. Dick’s *Do Androids Dream of Electric Sheep?*² Enhancement is often viewed in these radical forms—as something that human beings might strive towards in the future, with possible disastrous consequences, and wholly different from anything that has come before.³ In this vein, some researchers have treated the term “human enhancement” as nearly synonymous with direct and deliberate genetic engineering.⁴ Others have argued that human enhancement encompasses a host of

¹ ALDOUS HUXLEY, *BRAVE NEW WORLD* (1931).

² PHILIP K. DICK, *DO ANDROIDS DREAM OF ELECTRIC SHEEP?* (1968).

³ *See, e.g.,* MAXWELL J. MEHLMAN, *TRANSHUMANIST DREAMS AND DYSTOPIAN NIGHTMARES: THE PROMISE AND PERIL OF GENETIC ENGINEERING* pt. 2 (2012).

⁴ *See generally* NICHOLAS AGAR, *LIBERAL EUGENICS: IN DEFENCE OF HUMAN ENHANCEMENT* (2004).

emerging technologies—nanotechnology, biotechnology, robotics, information technology, and cognitive science (NBRIC)—that will allow for radical change to the human body and mind in the coming years.⁵

While there is nothing inherently wrong with thinking of human enhancement in more radical terms, such definitions obscure the fact that human beings have been “enhancing” ourselves, in some sense, for millennia. Consider the following definition: human enhancement technologies are any deliberate efforts to boost human capabilities above the species-typical level or statistically normal range of functioning.⁶ While there are some specificity problems with this definition,⁷ it at least gives us something to work with. From this definition, it becomes clear that all manner of human activity might be considered as “human enhancement.” If we imagine that the “species-typical” range of functioning for human beings in ancient times would have included an inability to read, then literacy becomes an enhancement. If we imagine the same with respect to certain

⁵ See *generally* CONVERGING TECHNOLOGIES FOR IMPROVING HUMAN PERFORMANCE: NANOTECHNOLOGY, BIOTECHNOLOGY, INFORMATION TECHNOLOGY, AND COGNITIVE SCIENCE (NBIC) (Mihail C. Roco & William Sims Bainbridge eds., 2003).

⁶ See Norman Daniels, *Normal Functioning and the Treatment-Enhancement Distinction*, 9 CAMBRIDGE Q. OF HEALTHCARE ETHICS 309, 314 (2000); Fritz Allhoff et al., *Ethics of Human Enhancement: 25 Questions & Answers*, 4 STUDIES IN ETHICS, LAW, AND TECHNOLOGY 1, 3 (2010).

⁷ For example, a “statistically normal range of functioning” implies a bell-curve where various human traits are distributed along the curve. If we consider one of those traits—I.Q.—there would be a large grouping of people (68% of the population) between 85 and 115. See EARL HUNT, HUMAN INTELLIGENCE 424 (2011). The specificity problem occurs when we try to pinpoint where some level of intelligence falls outside of the “normal” range of functioning. At some point we have to arbitrarily define what constitutes the normal range of functioning within the bell curve. However, as I mentioned, this specificity problem does not hamper using “statistically normal range of functioning” as a heuristic for exploring the concept of enhancement.

diseases in the nineteenth century, then vaccines become an enhancement. Indeed, Allhoff et al. define human enhancement in broad terms, as any activity we as humans use to improve our bodies, minds, or abilities.⁸ This includes activities such as learning how to read, exercising to build stronger bodies, and pursuing advanced academic degrees. These kinds of activities are not likely to shock anyone who lives in the modern world—and that is the point Allhoff et al. wish to make. As previously noted, the phrase “human enhancement technologies” often calls to mind futuristic scenarios where human beings have merged with machines and become cyborgs, or where the process of birth has become wholly subject to scientific determination. While these futures certainly involve technologies that fall under the broad umbrella of “human enhancement technologies,” Allhoff et al. want to make it clear that we already strive to enhance ourselves in myriad, uncontroversial, ways. This, it is hoped, will help to alleviate some of the anxiety that arises when people consider enhancement.

Allen Buchanan extends this effort at normalizing the concept of enhancement by considering the implementation of enhancements throughout human history. As human beings we have strived to better our station in the world since we took our first steps onto the African savannah. Once we left the relative safety of the jungle canopy and exposed ourselves to the dangers of the open plains, it became our brains—above all else—that allowed us to survive and thrive. One of the greatest functions of intelligence is in finding solutions to problems. The process of problem solving involves several steps.⁹ First, a

⁸ Allhoff et al., *supra* note 6, at 8.

⁹ This is my own, simplified, description of problem solving. For my purposes, *viz.*, describing the connection between intelligence and problem solving ability, it is sufficient. For an overview of various theories that consider problem solving in more

problem is identified. That problem might be something directly related to survival, for example, how to kill prey that are faster and stronger than us, or it might be related to comfort, for example, how to make sleeping in the cold with very little body hair more comfortable. Second, a means is discovered—either by invention, innovation, or dumb luck—that provides a solution to the problem. Incidentally, even if the solution comes about through dumb luck, it still requires intelligence to recognize that what has happened represents a solution to a problem as opposed to one of the countless random occurrences that occur in life but have no problem-solving efficacy. It was through use of our intelligence that human beings were able to develop agriculture, commerce, medicine, education, and every other historical stepping-stone that has brought us to our current state.

The preceding description of intelligence as problem solving fits nicely within the definition of human enhancement as activities that are pursued for the purpose of improving our bodies, minds, or abilities. Without intelligence there is no means by which to pursue enhancement. The point I am making here is that there *is* a distinction between the kinds of changes that occur through purposeful enhancement and those that occur through the unthinking mechanism of natural selection. With respect to the former there is an intelligent actor making decisions that lead to improvements in human capabilities. On the other hand, in the latter case there are only the undirected workings of natural selection. Any enhancement in human capability that arises through the process of evolution is not a human enhancement *technology*, because technology involves the

detail, see Yingxu Wang & Vincent Chiew, *On the Cognitive Process of Human Problem Solving*, 11 COGNITIVE SYS. RES. 81 (2010).

application of knowledge for a purpose—and evolution is devoid of purposeful ends or goals.¹⁰

Human Enhancement Technologies as Interfering with Nature

This clear line between purposeful human enhancement and the blind workings of evolution does not necessarily imply that similar lines may be drawn between the kinds of enhancements that humans have pursued in the past and those that modern science and technology may allow us to pursue today. History is replete with examples of technology luddites who warned of the disastrous impacts that would follow the implementation of a new human enhancement—and those warnings were often a result of an attempt to distinguish the technology at issue from that which had come before. For example, there was vigorous resistance to the proliferation of the printing press in part due to fears that it would allow the masses to gain access to knowledge hitherto reserved for the elites.¹¹ Similar reactions have accompanied the introduction of technologies from the industrial revolution to the computer revolution.¹² This is not to suggest that all concerns about technology advances are, or have been, reactionary in nature. For example, there was

¹⁰ See, e.g., Jean-Louis Monestes, *Human Evolution and Contextual Behavioral Science*, in THE WILEY HANDBOOK OF CONTEXTUAL BEHAVIORAL SCIENCE 100, 110 (Robert D. Zettle et al., eds., 2016) (“[T]he consensus among evolution scientists is that evolution has no goal, and is nondirectional.”). See generally RICHARD DAWKINS, THE BLIND WATCHMAKER (1986); STEPHEN JAY GOULD, WONDERFUL LIFE (1989); DANIEL DENNETT, DARWIN’S DANGEROUS IDEA (1995); STEPHEN JAY GOULD, FULL HOUSE (1996). I will address this point in more detail later in this chapter.

¹¹ See Vaughan Bell, *Don’t Touch That Dial! A History of Media Technology Scares, from the Printing Press to Facebook*, SLATE (Feb. 15, 2010), http://www.slate.com/articles/health_and_science/science/2010/02/dont_touch_that_dial_single.html.

¹² *Id.*

significant opposition to the development and use of nuclear weapons in the middle of the twentieth century. Even Einstein, who developed the theory that led to the possibility of nuclear weapons, opposed the development of nuclear technologies, fearing that such awesome power in the hands of governments would lead to disaster.¹³ Of course, there were others, such as Robert Oppenheimer, who made the contrary argument: that development of nuclear technologies would lead to a *reduction* in warfare.¹⁴ The point is not that Einstein was right and Oppenheimer was wrong. The point is that, whatever one's ultimate views on the matter, it would be difficult to call Einstein's concerns "reactionary." His concerns represented genuine dangers of the new technology of nuclear power, and these concerns were backed up by an understanding of the science behind the technology, and by rational conclusions about how some governments might use that technology. These kinds of concerns are distinct from technological "ludditism," whereby any new technology or enhancement is seen as a threat, whether or not there are rational reasons to be concerned about the effects of the particularly technology.

A Case Study: The Anti-Vaccination Movement

Given the strong reactions that have accompanied the advent of so many technologies, the reaction to the threat of biomedical human enhancements should not come as a surprise. When Edward Jenner developed the first vaccine in 1796, his discovery had two conflicting outcomes. On the one hand, there was widespread celebration of the revolutionary approach to smallpox prevention. Over 100,000 people in Europe and the

¹³ See WALTER ISSACSON, *EINSTEIN: HIS LIFE AND UNIVERSE* 490 (2007).

¹⁴ See THÉRÈSE DELPECH, *NUCLEAR DETERRENCE IN THE 21ST CENTURY* 25 (2012) (quoting Oppenheimer: "[the atomic bomb] has made the prospect of war unendurable.").

United States received the vaccine by the year 1800.¹⁵ On the other hand, much of the public was skeptical of Jenner's discovery, especially because the procedure required taking pus from a cow and inserting it into the human body.¹⁶ Presaging the current climate of the vaccine debate, opposition to vaccinations intensified in the 1830s as the rate of smallpox infection declined.¹⁷ There were reasons for concern: sterilization was still a long way off, so other diseases could spread through inoculation.¹⁸ This reasonable concern was rarely at the forefront of the anti-vaccination argument, however. Instead, there were wild concerns about government conspiracies and cover-ups, accompanied by generalized claims that vaccines were inherently poisonous.¹⁹ One of the anti-vaccinationist's most vociferous arguments was that vaccines were "an attempt to swindle nature."²⁰ Proponents of this view held that the body's natural immune system was being hijacked and, thus, made ineffective by the introduction of "unnatural" vaccines.²¹

¹⁵ Alexandra Minna Stern & Howard Markel, *The History of Vaccines and Immunization Familiar Patterns, New Challenges*, 24 HEALTH AFFAIRS 611, 614 (2005).

¹⁶ *Id.* at 613.

¹⁷ *Id.* at 617.

¹⁸ *Id.* at 616.

¹⁹ See J Leask & S Chapman, 'An Attempt to Swindle Nature': Press Anti-Immunisation Reportage 1993-1997, 22 AUSTL. & N.Z. J. PUB. HEALTH 17 (1998).

²⁰ *Id.*

²¹ See *id.*

Resistance to vaccinations intensified after England passed the first mandatory vaccination law in 1853.²² In London, the Anti-Vaccination League was founded, and by the 1880s there was several anti-vaccination groups and journals throughout Europe.²³ Similar movements sprang up in the United States. Following decades of intense debate, the Supreme Court ruled, in *Jacobson v. Massachusetts*, that the government had the authority to enforce compulsory vaccination laws.²⁴ Recognizing the danger of allowing individuals to opt-out of vaccination, the Court reasoned that personal liberties must sometimes be forfeited for the overall welfare of the community.²⁵

Following the Court's decision, the anti-vaccination movement lost momentum in the early part of the twentieth century. Indeed, when the polio vaccine was introduced in 1953, the public flocked to receive the treatment for a disease that had devastated up to 58,000 Americans per year in the preceding decades.²⁶ Following a mass vaccination effort, in 1961 there were only 161 reported cases of polio in the United States.²⁷

²² Robert M Wolfe & Lisa K Sharp, *Anti-Vaccinationists Past and Present*, 325 BMJ 430, 430 (2002).

²³ *Id.*

²⁴ 197 U.S. 11 (1905).

²⁵ *Id.* However, the Court included an exemption for opposition based on religious beliefs.

²⁶ Atif Kukaswadia, *History of Epidemiology: Jonas Salk and the Eradication of Polio*, PUBLIC HEALTH PERSPECTIVES (June 4, 2013), <http://blogs.plos.org/publichealth/2013/06/04/history-of-epidemiology-jonas-salk-and-the-eradication-of-polio/>.

²⁷ Ginny A. Roth, *Celebrating Salk*, NIH (March 26, 2014), <http://circulatingnow.nlm.nih.gov/2014/03/26/the-first-polio-pioneer/>.

Despite the short reprieve in anti-vaccination sentiment provided by the success of the polio vaccine, anti-vaccination rhetoric began to take hold once again in the 1970s after the introduction of the diphtheria, tetanus, and pertussis (DTP) vaccine.²⁸ However, it was not until 1998 that the modern anti-vaccination movement took full flight. In that year, a physician named Andrew Wakefield began proliferating concerns that the measles, mumps, and rubella (MMR) vaccine was associated with an increase in the incidence of autism.²⁹ The British medical journal, *The Lancet*, published Wakefield's paper, leading to an immediate public reaction against vaccinations.³⁰ The scare spread to the United States following Wakefield's appearance on CBS's *60 Minutes*.³¹ The medical community was immediately skeptical. Wakefield's study involved twelve children, all of whom were self-selected, and no control group.³² Furthermore, the gastrointestinal problems that Wakefield identified as being caused by the MMR vaccine and, in turn, causing autism, did not predate the symptoms of autism in many subjects.³³ Most alarming, it was revealed that Wakefield had been paid enormous sums of money by a

²⁸ See *History of Anti-Vaccination Movements*, HISTORYOFVACCINES.ORG (Dec. 18, 2014), <http://www.historyofvaccines.org/content/articles/history-anti-vaccination-movements#Source 8>.

²⁹ Jeffrey S. Gerber & Paul A. Offit, *Vaccines and Autism: A Tale of Shifting Hypotheses*, 48 CLINICAL INFECTIOUS DISEASES 456, 456 (2009).

³⁰ *Id.*

³¹ Brian Deer, *Exposed: Andrew Wakefield and the MMR-Autism Fraud*, BRIANDEER.COM, <http://briandeer.com/mmr/lancet-summary.htm> (last visited Oct. 21, 2015) (collecting Deer's entire investigative work on the issue).

³² Gerber & Offit, *supra* note 29, at 456.

³³ Indeed, even if Wakefield's results were as he stated, his conclusion would be based solely on the logical fallacy of *post hoc, ergo propter hoc*.

plaintiff's attorney to conduct research demonstrating a link between the MMR vaccine and autism.³⁴ In an attempt to ease the public's mind, the medical and epidemiological communities conducted dozens of large-scale studies, all of which provided conclusive proof that there was no link between the MMR vaccine and the incidence of autism.³⁵ And in 2004, *The Lancet* formally retracted the publication of the Wakefield paper.³⁶ Nonetheless, the damage had been done. To this day, many people continue to believe in a link between the MMR vaccine and autism. In certain parts of the country, often in areas of relative affluence and high education levels, the vaccination rate has dropped substantially—in some cases leading to outbreaks of diseases such as measles that had all but been eradicated in prior decades.³⁷

Interestingly, the same arguments that motivated anti-vaccination movements of prior centuries have resurfaced following the MMR / autism controversy.³⁸ Amongst these concerns is the old argument that vaccines interfere with nature.³⁹ Supporters of this view decry vaccines as a dangerous enhancement that goes against nature and will ultimately do more harm than good. These “anti-vaxxers” believe, inter alia, that

³⁴ Deer, *supra* note 31.

³⁵ Gerber & Offit, *supra* note 29, at 456.

³⁶ Deer, *supra* note 31.

³⁷ See Clyde Haberman, *A Discredited Vaccine Study's Continuing Impact on Public Health*, NY TIMES, Feb. 1, 2015, http://www.nytimes.com/2015/02/02/us/a-discredited-vaccine-studys-continuing-impact-on-public-health.html?_r=0.

³⁸ See *Box A: Anti-vaccination arguments, past and present*, THE BMJ, <http://www.bmj.com/content/suppl/2002/08/22/325.7361.430.DC1>, for a comparison of past and present concerns about vaccines.

³⁹ *Id.*

interfering with the body's natural reaction to exposure to certain diseases such as chicken pox disrupts the body's natural immune response to such exposure and therefore weakens the body's overall immune response, ultimately leaving it more susceptible to more dangerous diseases.⁴⁰

The reaction of the “anti-vaxxers” is, of course, devoid of scientific backing. Their claim that certain human enhancement technologies interfere with nature, however, has been applied to a whole host of human enhancement technologies by intelligent researchers. Thus, it is important to take the claim seriously. Note that I am not suggesting that all concerns about human enhancement technologies are as absurd as the “anti-vaxxer” argument. Rather, I am simply pointing out that many, more legitimate, worries ultimately rest on the same foundation: that enhancement interferes with nature.⁴¹ For example, one might argue that creating a human-chimpanzee chimera would be wrong because, inter alia, it is “unnatural.” My argument in the next few paragraphs will

⁴⁰ See, e.g., *Vaccines Interfere with your Body's Ability to Naturally Kill Viruses, Bacteria, Parasites and Tumor Cells*, PRESS CORE (Jan. 15, 2011), <http://presscore.ca/vaccines-interfere-with-your-body%E2%80%99s-ability-to-naturally-kill-viruses-bacteria-parasites-and-tumor-cells>; Joseph Mercola, *Flu Vaccines: The Horrible “Immune System Mistake” Millions Will Make This Year*, MERCOLA.COM (Jan. 3, 2012), <http://articles.mercola.com/sites/articles/archive/2012/01/03/flu-shot-increase-flu.aspx>. Some parents participate in “pox parties” where they intentionally expose their children to the chicken pox virus, believing that allowing the body to fight off the virus “naturally” will lead to a stronger immune system than if their children are given the vaccine. See Alice G. Walton, *The Chicken Pox Party*, THE ATLANTIC (Nov. 25, 2011), <http://www.theatlantic.com/health/archive/2011/11/the-chicken-pox-party-parents-caught-infecting-kids-with-virus/248768/>.

⁴¹ There is a distinction between a worry that interfering with nature is, ipso facto, wrong, and a worry that interfering with nature may lead to unintended bad consequences. I address both in this chapter. At this point in the chapter, I am only addressing the former concern.

not suggest that all concerns about human-chimpanzee chimeras are untenable; rather, I will argue that that simple argument that something is “unnatural” is quite weak.⁴²

Interference with Evolution

There are two facets to the claim that certain human enhancement technologies interfere with nature. The first is what Buchanan calls *the analogy of the master engineer of evolution*. According to this view, “[t]he human body and mind, highly complex and delicately balanced as the result of eons of gradual and exacting evolution, are almost certainly at risk from any ill-considered attempt at ‘improvement.’”⁴³

This view assumes that (1) evolution acts as a designer, crafting organisms that are optimally adapted to their environment and (2) human intelligence cannot hope to compete with the “wisdom” of evolution—thus, any attempts to “improve” upon evolution will inevitably lead to disaster.⁴⁴ The view of evolution as a designer can be dismissed rather easily. Beginning with Darwin, evolutionary theorists have emphasized that evolution is not a designer, and there is no reason to think that any organisms—including human beings—are optimally adapted to their environment. Because evolution picks up and discards traits in a haphazard manner as environmental pressures (along

⁴² I put “unnatural” in quotes because it, and the accompanying term “natural,” are difficult to define. I will address the ambiguity of these terms in the following sections.

⁴³ ALLEN E. BUCHANAN, *BEYOND HUMANITY?: THE ETHICS OF BIOMEDICAL ENHANCEMENT* 156 (2011).

⁴⁴ Of course, there *are* situations where attempts to “improve” upon evolution would likely result in disaster. For example, if scientists were to haphazardly insert genes from other species—genes that produce beneficial results in those species—into human germ lines. The only point I am making here is that it is not inevitable that *all* attempts to improve upon evolution will result in disaster.

with genetic drift) ebb and flow, the resulting organisms are a hodge-podge of traits, some well-suited to the environment and other abysmally lacking in form or function. Obvious examples include the chronic back pain that is endemic to humans as a result of our knuckle-walking ancestors, as well as the multitude of debilitating diseases that strike us only after we have surpassed the age of procreation (and that therefore cannot be altered by evolutionary forces). On the other hand, the view that human intelligence may do more harm than good in trying to “improve” upon evolution is not so easily refuted. Before addressing that concern, let me begin by considering the second facet to the claim that certain human enhancement technologies interfere with nature.

The Natural / Artificial Distinction

The second facet of the worry that certain human enhancement technologies interfere with nature is that there is a distinction between the natural and the artificial (or the unnatural). This facet can be broken down into two separate questions. First, is there a viable distinction to be made between the natural and the artificial? Second, if such a distinction can be made, does the distinct concept of “the natural” have normative force? Allhoff et al. suggest that one cannot separate the natural from the artificial without engaging in arbitrary line drawing.⁴⁵ One might draw the natural / artificial distinction between those things that arise in nature without the intervention of human beings and those things that arise as a result of human intervention. In that case, everything that existed before the advent of *Homo sapiens* would be considered natural. But it would likewise entail that everything that human beings have had a hand in—from using fire to

⁴⁵ Allhoff et al., *supra* note 6, at 4-5.

cook meals to reading (or writing) the plays of Shakespeare—is “unnatural” or artificial. Of course, there is nothing wrong with this distinction per se. Its failure lies in the fact that insofar as someone was motivated to make a distinction between the natural and the artificial, she would likely have something juicier in mind than a distinction between nature sans humans and nature plus humans.

Indeed, the opposition to vaccines (specifically, the opposition that considers vaccines to be unnatural) appears to involve a different kind of natural / artificial distinction. Let me define the distinction this way (I consider this to be a very charitable definition): *something is natural if it has been engaged in by humans throughout most of our history, and something is unnatural if it involves modern technology in a way that interferes with the natural (as the natural has just been defined).*⁴⁶ This definition avoids the problem I raised above with defining the natural / artificial distinction in terms of nature sans humans vs. nature plus humans. With this new definition, certain human activities are considered natural while others are considered unnatural. Yet this new distinction brings with it a whole host of new problems. First and foremost is what I call the *time-lapse problem*. The *time-lapse problem* simply identifies that the word “modern” is specific to the contemporary time in history from the perspective of the people living in

⁴⁶ “Anti-vaxxers” oppose vaccinations for other reasons as well. For example, many anti-vaxxers continue to believe in a correlation between vaccination and autism. However, the fact that vaccines are “unnatural” is a hidden premise in most anti-vaccination argument. Thanks to Ami Palmer for pointing this out. For an example of that premise made explicit (in the context of religion), see Megan Heimer, *God Does Not Support Vaccines*, LIVING WHOLE (July 7, 2014), <http://www.livingwhole.org/god-does-not-support-vaccines/>. See also every comment thread accompanying a pro-vaccination article. My purpose here is to present a charitable definition of what “unnatural” might mean in this context, given that anti-vaxxers do not define the term beyond comparing it to what is “natural,” which they also leave undefined.

that time. As such, the invention of fire would have been modern by the standards of the humans alive at that time and thus would be considered unnatural based on the definition. Yet from our current perspective fire is natural since it is something humans have used throughout most of our history. The extension of the *time-lapse problem* is that things that the definition now labels as unnatural will become natural after a sufficiently long period of time. Thus, in approximately 201,000 years vaccines would be considered natural under the definition (given that *Homo sapiens* has been on earth for approximately 200,000 years and so in 401,000 years vaccines will have been in use for “most” of human history). Surely, opponents of vaccination would not be comfortable in saying that given enough time vaccines will suddenly become natural as opposed to unnatural. Thus, even this charitable definition of the natural / artificial distinction fails to describe a meaningful distinction.

The Essential Good of the Natural

There are other definitions of the natural / artificial distinction that could be tried, and for my purposes I will assume for the moment that at least one of them succeeds in describing a meaningful distinction. This will allow me to address the second question I raised above: does the concept of “the natural” have normative force? In other words, does that fact that some act is natural mean that we *should* engage in that act and, conversely, that we should avoid engaging in acts that are unnatural (or artificial)? Certainly, the argument that certain things are unnatural rests on just such a normative foundation. But an essentialist argument for the natural as a normative force is unlikely to persuade any serious thinker. Thus, the advocate for the natural must be able to unpack the normative element of the term. One way that advocates for the natural have attempted

to unpack the normative element of the natural is by arguing that nature has more “wisdom” than our human intellect and, thus, we are better off sticking with the natural. This view is akin to the conception of *the master engineer of evolution*, described above. Yet it goes a step further: rather than simply identifying the risks associated with interfering with evolution, it conceives of the natural as being uniformly good or helpful (at least in a larger sense) to human beings.⁴⁷ I will refer to this view as *the essential good of the natural*.

According to this view, the things that human beings receive through the natural are the best things for us. There are two ways to flesh out this view. The first requires consideration of the rich body of work on the natural / artificial distinction in environmental philosophy.

Lessons in the Natural / Artificial Distinction from Environmental Philosophy

The environmental ethicist Robert Elliot has argued that there is value that can be derived from an experience of the natural world that cannot be derived from an artificial reproduction of the natural world.⁴⁸ According to Elliot, even a perfectly accurate reproduction of nature, down to the smallest detail, would not hold the same value as

⁴⁷ A prominent proponent of this view is Dr. Jack Wolfson—a self-styled “natural” cardiologist based in Green Valley, AZ. He believes (and his large patient roster indicates that his ideas have traction amongst a certain population) that “we should be getting measles, mumps, rubella, chicken pox, these are the rights of our children to get it [sic].” Terrance McCoy, *Amid Measles Outbreak, Anti-Vaccine Doctor Revels in His Notoriety*, WASH. POST, Jan. 30, 2015, <https://www.washingtonpost.com/news/morning-mix/wp/2015/01/30/amid-measles-outbreak-anti-vaccine-doctor-revels-in-his-notoriety/>. Dr. Wolfson also advocates for a “paleo” diet that includes only “foods that our ancestors have been eating for millions of years.” *Id.*

⁴⁸ See Robert Elliot, *Faking Nature*, 25 INQUIRY 381 (1982).

nature itself. This difference in value can be traced to the historical element of nature. For most people who have visited the Grand Canyon, there is a sense of awe and wonder that accompanies the experience of seeing such an immense and beautiful spectacle. Yet, this feeling of awe is not simply a result of the canyon's aesthetic beauty. A significant portion of the wonder we feel in the presence of the canyon is staring out into a massive wound cut into the earth and realizing that it is the result of hundreds of millions of years of natural phenomena. When we catch a glimpse of the tiny Colorado River far below the canyon's edge, we feel overwhelmed by the power of time, which has allowed such a small stream of water to carve such an immense tract into the land. Now imagine that a mining company were to request access to a portion of the canyon with the stipulation that after they had extracted some portion of the minerals in the area they would use the most sophisticated methods available to restore the area to its previous appearance. Surely, there would be mass protest to such a proposal; and people would scoff at the idea that a reproduction of the canyon—even a perfect one—could hold the same value as the original. The absence of deep time would form a glaring hole in the experience of the newly restored canyon.

As Elliot points out, the inadequacy of human-formed reproductions of nature is akin to the relative lack of value we place in reproductions of original art. Even though most of us are satisfied to own reproductions of paintings by Van Gogh, Picasso, or Michelangelo, if the originals of these paintings, held in museums, were to burn up in a fire, we would surely feel that something of special value had been lost—despite the fact that thousands of perfect reproductions exist throughout the world. Similarly, if I owned what I believed to be an original Picasso, I would be devastated if I were to find out that

it was, in fact, a skillfully made forgery. And my disappointment would not stem entirely from the loss of monetary value. I would also lose the feeling of awe and connection to history and to the artist who rendered the work. A similar feeling of disappointment would surely accompany a realization that the portion of the Grand Canyon that one is looking at is, in fact, an artificial restoration project. There may be a different sense of awe that accompanies witnessing humankind's power to shape the landscape, but it would not be fungible with a sense of awe in the face of nature.

An important question about Elliot's natural / artificial distinction is whether its significance depends upon the knowledge of the person experiencing the phenomenon. Elliot argues that it does not. According to Elliot, "there can be loss of value without the loss being perceived."⁴⁹ In other words, even if we truly believe that the landscape we perceive is the true Grand Canyon with all of its historical significance, if we are mistaken about the facts and we actually perceive an artificial canyon, there is a loss of value. Here, we can analogize to the "experience machine": a device that immerses you in a virtual reality that appears identical to the real world. The thought experiment asks: would you enter the world of the experience machine if you could have a happier, more fulfilling life in the artificial world? The rub of the question is that you would not know, once inside the world of the machine, that the experience was artificial. Elliot would say no—that there would be something of value lost in living in an artificial world even if one were not aware of its artificiality.

This idea is illustrated in many fictional works. In the film *The Matrix*, the human race has been enslaved by machines that use the biological human body as a power

⁴⁹ *Id.* at 385.

source.⁵⁰ Each human being's biological body is kept in a pod in the real world. The machines had discovered that human bodies could not be sustained without mental experience so they created an artificial, virtual-reality, world that perfectly mimicked the state of the world at the end of the twentieth century. Each human being is connected to this "matrix" and has a shared experience with her fellow human beings. These human beings live out their lives in the Matrix, totally unaware that their experience is artificial. However, a few people in the Matrix have a sense that something is off about the world in which they live. Some of this group discovers a way to escape from the Matrix and get back into the real world. There, they form a resistance movement to the machines with the goal of destroying the Matrix and returning humanity to reality. The resistance group believes that there is something fundamentally wrong with subjecting human beings to a false reality, even if those human beings believe their experience to be real and would likely choose to stay in the Matrix if given the choice.

A related story is Plato's allegory of the cave. In this story, a group of human beings is imprisoned in a cave, where their only experience of reality is in the shadows of things projected onto the cave wall by a fire that burns behind them.⁵¹ Plato has us imagine that one of the prisoners is accidentally freed and able to turn towards the fire and escape out of the cave. While he would be terrified at first, he would soon realize that the false reality he had been subjected to in the cave was pale in comparison to the real world. As in *The Matrix*, Plato imagines that if the man returned to the cave to tell the other prisoners about his revelation, most of them would not want to hear it. Yet, the man

⁵⁰ THE MATRIX (Warner Bros. 1999).

⁵¹ PLATO, REPUBLIC 509d-513e.

would feel a compulsion, even a duty, to “free” the other prisoners from their false reality and show them what they have been missing. The man would believe that the other people are missing out on something of greater value—even though they are content with their false version of reality.⁵²

There is some force to the argument that an experience of the natural world, rather than a reproduction, has inherently greater value and, thus, is something we should strive to preserve. However, there are two difficulties that come to mind. One with respect to the concept of nature as understood by Elliot (and others), and one with respect to how this specific natural / artificial distinction might apply to human enhancement technologies. The first difficulty is that there may not be anything that truly fits within the category of “nature,” if we define nature as spaces that have been wholly untouched by human activity. The environmentalist William Cronon has made such a critique of the natural / artificial distinction.⁵³ According to Cronon, the idea of nature as something wholly separate from human activity is mistaken. First, it is factually wrong because there is no place on earth that has not been touched by human activity, either through direct contact with the land, depletion of surrounding resources, or through the influence of fossil fuels and other greenhouse gases. Second, and more importantly, the idea of nature as something we should preserve from human influence devalues our place in

⁵² Of course, in Plato’s allegory of the cave, the “real” is the world of forms or ideas, which is a more abstract concept than presented in *The Matrix*. However, the analogy holds since the purpose of the allegory is to demonstrate that there is something more valuable about the “real,” even if most people are more content with a false representation of reality.

⁵³ See William Cronon, *The Trouble with Wilderness; or, Getting Back to the Wrong Nature*, in UNCOMMON GROUND: RETHINKING THE HUMAN PLACE IN NATURE 69 (ed. William Cronon, 1995).

nature—and especially devalues the experience of communities who make their lives on and through the land. As Cronon points out, much of the “uninhabited” wilderness we now view with almost sacred reverence is only that way due to the forced expulsion of native peoples.⁵⁴

Cronon also points out that our reverence for “nature” is often lopsided and demonstrates, more than any respect for an objective nature, “a reflection of our own unexamined longings and desires.”⁵⁵ As an illustration of this point, Cronon points to the fact that almost all of our attention and resources go into preserving mountainous regions such as Mount Rainier and the Grand Canyon, and there is little attention paid to the less dramatic natural landscape of the Midwestern grasslands.⁵⁶ Cronon’s point is not that we should stop trying to preserve natural wonders and replace them with human-made copies. Rather, he asks us to see ourselves for what we are—as part of nature: “Our challenge is to stop thinking of such things according to a set of bipolar moral scales in which the human and the nonhuman, the unnatural and the natural, the fallen and the unfallen, serve as our conceptual map for valuing and understanding the world.”⁵⁷

A second, and related, problem with the natural / artificial distinction from Elliot’s take on environmental philosophy is that the arguments in favor of an experience of nature over a replica are not easy to translate into discussions of human enhancement technologies. Consider one of Elliot’s arguments: that the sense of reverence and history

⁵⁴ *Id.* at 79.

⁵⁵ *Id.* at 69.

⁵⁶ *Id.* at 72.

⁵⁷ *Id.* at 88.

we feel in the presence of nature has value. Part of the appreciation of nature, as Elliot and others (including myself) see it, comes from a feeling of being small and temporary in the face of something grand and ancient. As Cronon puts it, nature has a tendency to bring out our own “unexamined longings and desires.”⁵⁸ Being in the presence of something so utterly inhuman—something with no will, no desire, no ambition, and no fear—condenses and amplifies the otherness of our human lives, which are filled with hopes, dreams, anxieties, and fears.

Now consider our relationship to our own bodies. Certainly, we can feel a kind of reverence in our etiology. We can recognize the historical contingencies that brought us, among so many potential yet unrealized beings, into existence. And we can feel humbled in the face of eons of development that, through unwilled natural selection, have led to our enormously complex bodies and minds. Yet, quite unlike our experience of nature, we can just as easily find flaws in our natural vessels. While our biological substrates, like the rest of nature, may not have a purpose or a goal, *we*, as thinking beings, do live through teleology. We have aspirations for our lives and as often as not our bodies and minds stand as impediments to the realizations of our goals.

When we face obstacles in nature, at least in theory, with enough resources we can sidestep those obstacles by moving to another location. Thus, we can achieve our goals while leaving nature intact (again, in theory). Not so with our own bodies and minds. If a person wishes to have a realistic chance of raising a child to adulthood in a disease-ravaged part of the world, she must vaccinate that child against preventable disease. If a person born with a genetic defect that leaves them without legs wishes to

⁵⁸ *Id.* at 69.

walk, she must obtain artificial legs. And it may very well be that if human beings as a species wish to survive or to obtain lasting periods of peace and wellbeing, we must enhance our bodies and minds in a way that conflicts with letting nature “take its own course.”

In proceeding paragraphs, I will argue against letting natural evolution control our destinies. For now, my goal has been to show that whatever significance the natural / artificial distinction may have in the environmental realm does not provide insight into how we should think about our relationship to our own bodies.

The Essential Good of the Natural According to Michael Sandel

There is another interpretation of *the essential good of the natural* that deserves attention. This interpretation is distinct from the analogy to the natural / artificial distinction from environmental philosophy, and it provides a counterpoint to my argument that while there may be value to leaving nature be in the environmental context, the direct harm that can be done to our bodies and minds by letting nature control our lives outweighs any justification to not interfere in the natural order.

According to this interpretation, everything that happens as a result of the natural is not necessarily for the best, in and of itself. Rather, some things that happen as a result of the natural are, in themselves, bad, but they provide us with the opportunity to appreciate the good things in life. A version of this view has been most widely defended by Michael Sandel.⁵⁹ Sandel argues that one of the most important human goods is an

⁵⁹ Michael J. Sandel, *The Case Against Perfection*, THE ATLANTIC (April 2004), <http://www.theatlantic.com/magazine/archive/2004/04/the-case-against-perfection/302927/>.

appreciation for a sense of “giftedness” that derives from a recognition that whatever good things happen to us are only partially a result of the efforts we may have engaged in to bring them about. This view implies a distinction that is similar to a natural / artificial distinction. But in terms of Sandel’s argument, it might be fairer to consider the distinction as between those things that come about as a result of human will and those that come about without the interference of human will. As such, Sandel argues that we should view our talents and powers as—in significant measure—a “gift” of nature because significant aspects of those talents and powers are not a result of human will.

According to Sandel, an appreciation of the giftedness of nature requires us to recognize the limits of human will to control—or interfere with—nature, or the natural. This is because an appreciation for the giftedness of life requires that we have the opportunity to experience the unfortunate negative events that life often throws our way. In other words, without the bad there can be no appreciation of the good. Sandel’s concern, then, is that we might use our human will to interfere with the natural to the extent that the domain of contingency in which we live will be significantly diminished, thus leaving us with fewer opportunities to appreciate the giftedness of life. For Sandel, a sense of giftedness is a good in and of itself. Thus, one potential consequence of his view is that even if we were able to develop a world—through use of our human will—that dramatically increased our ability to better survive, either through increased health, longevity, or comfort, that world would not *in fact* be better for us because it would lack (or provide for significant reduction in our ability to appreciate) a sense of giftedness.

There is a caveat to this consequence of Sandel’s view, however. For Sandel, the motivation behind human will is more important than the consequences that stem from

that will. Thus, if our motive is to obtain mastery over nature, then no good result will outweigh the damage done to our sense of giftedness. Indeed, Sandel sees the kind of high-pressure tactics that parents engage in in an attempt to mold their children to some ideal as objectionable for the same reason as direct and deliberate human enhancement technologies: they both seek mastery over nature.⁶⁰ If, on the other hand, our motive is to make life just a little bit easier or more comfortable while at the same time eschewing any notion that we might one day be able to dominate nature entirely, then there is no risk of losing our sense of giftedness; therefore, similar ends as might result from a motive based on mastery would be acceptable under the latter motivation.

Recall that those who adhere to the *essential good of the natural* believe that what is natural is always good for us, even if the immediate consequences might appear to be bad. Sandel's view can be used to bolster the argument of those who believe in this view of *the essential good of the natural* because it provides a more sophisticated exposition of the good that is achieved by accepting the natural in all its forms—good and bad. However, the increase in sophistication and subtlety provided by Sandel's view likely comes with the cost of having to abandon *the essential good of the natural* as a justification for disapproving of more benign human enhancement technologies such as vaccines. If Sandel is most concerned with technologies that purport to give us “mastery” over nature—or with the motives that lie behind the invention and application of those technologies—then it seems that many of the enhancement technologies we currently take for granted, for example, vaccines, fall outside the purview of that concern.

⁶⁰ *Id.* (“The hyperparenting familiar in our time represents an anxious excess of mastery and dominion that misses the sense of life as a gift. This draws it disturbingly close to eugenics.”).

Sandel appears to concede this point, as he must if he wants to save his argument from the objection that its consequence would be that vaccines and medical therapies should be abandoned. However, given the way that Sandel frames his argument—that we should appreciate the “unbidden” gifts and detriments that nature bestows upon us—it isn’t clear that he can simply sweep technologies such as vaccines out of the way. Sandel states that “[m]edical intervention to . . . prevent illness . . . does not override a child’s natural capacities but permits them to flourish.”⁶¹ But why not? It seems that, at least with respect to diseases that are the result of genetics, disease is as much a “gift or detriment” of nature as, say, a genetic predisposition to high or low IQ. It’s not at all clear why, based on Sandel’s argument, there should be a distinction between genetic predispositions that make life more difficult through disease and genetic predispositions that make life more difficult through, say, low IQ. If Sandel opposes intervention in the latter case, it seems that he should oppose intervention in the former case as well. The converse is that if Sandel is willing to accept interventions such as vaccines and other medical interventions then he should be willing to accept other kinds of enhancements that would, in a sense, allow a person’s “natural” capacities to flourish.

It is surely possible to see a genetic predisposition to disease and a genetic predisposition to low IQ as falling in the same camp. In both cases, a technology that disrupted the natural predisposition might allow the child to flourish with respect to other traits. An intelligent child might be hampered in her pursuits by disease, in which case a vaccine would allow her trait of intelligence to be fully realized. By the same token, imagine another child who is a naturally gifted musician, but whose low intelligence

⁶¹ *Id.*

hampers her ability to fully realize that gift through learning to read music and understand musical theory. In this case, just like the case of the child with a disease, an enhancement—this time with respect to IQ—would allow the child to fully realize her musical gift.

Sandel states that he is primarily concerned with humankind's attempts to have "total mastery" over nature. On this interpretation, Sandel's view seems quite reasonable. It does not say that all interferences with the natural order are bad. But the consequence of this interpretation is that *the essential good of the natural* is undermined. Those who believe that things like vaccines interfere with the natural order will not be satisfied by this interpretation of Sandel's view. On the other hand, if we focus on Sandel's concern with "openness to the unbidden," it seems that Sandel must either abandon his view that vaccines and other medical interventions that target (at least) genetic predispositions to disease are acceptable, or he must allow that other kinds of enhancements, say to IQ, are acceptable as well. If he does the former, then his view becomes untenable. If he does the latter, then his view no longer supports, in any sense, *the essential good of the natural*. Thus, on either interpretation of the Sandel's view—mastery or openness to the unbidden—it seems that he is unable to salvage *the essential good of the natural* from conceptual confusion.

Buchanan's Argument From Biological and Social Harmony

Regardless of any conceptual confusion I might impute to *the essential good of the natural*, it is impossible to deny the intuitive appeal of this view (or, at least, some formulation of it) to some serious thinkers. As such, I think it is important to let the view stand for a moment, if only to let its adherents know that they are not being dismissed out

of hand. Instead of abandoning the view as it stands, it will be more productive to consider some of its consequences, including views that rely on it as a foundation. If those consequences, or views, turn out to be untenable or undesirable, the underlying view—*the essential good of the natural*—will (or, at least, should) lose its appeal.

So what are some of the views that depend upon—or follow from—the view of *the essential good of the natural*? Buchanan identifies a view that he calls the Argument from Biological and Social Harmony.⁶² This view is related to the *analogy of the master engineer of evolution* but it extends the dependencies that are of such concern in the former view from pure biological dependencies to the realm of social institutions. Simply put, this argument suggests that human biology and human social institutions exist in harmony with one another such that the former supports the latter.⁶³ Even more than providing support, our current social institutions *depend* upon our biology—as it stands—for their survival. There is an undeniable grain of truth to this view. Our social institutions—marriage, family, child rearing, government, etc.—are all a result, in one way or another, of biological evolution. Even if these institutions are not a *direct* result of natural selection, in other words, they were not necessarily “selected” for their benefits to reproductive fitness, they are at the very least a result of our attempts to manage our place in the natural world. If our biology were to be radically altered, we would likely need to discover new ways to manage our place in the world, which would likely result in at least some significant changes in many of our social institutions. But notice that I specified *radical* alteration to our currently existing biology. It is not easy to see how human

⁶² BUCHANAN, *supra* note 43, at 161.

⁶³ *Id.*

enhancements that mitigate problems associated with disease, disability, or aging, would disrupt our social institutions to any worrisome degree.⁶⁴ For example, there are no obvious institution-disrupting effects that would come about if scientists were able to change our biology such that we were able to biosynthesize vitamin C—something no humans are currently able to do.⁶⁵ Indeed, there are a significant number of human idiosyncrasies that appear to have no role in our social institutions (other than providing ample grist for the complaint mill).

The Precautionary Principle

A defender of the Argument from Biological and Social Harmony might respond that while there are no obvious disrupting effects of changing our “natural” biology, all intelligent people know that the biological world—including human biology—is interdependent to such a great extent that a disruption in one area can have unforeseen effects in other areas. This defense smacks of the famous (or infamous) “precautionary principle,” which holds—in one formulation of this notoriously slippery concept—that when we do not fully understand the potential consequences of some action the best thing to do is nothing.⁶⁶ The precautionary principle is usually associated with the environmental movement, but it is not a stretch to suggest that it might be applied to the

⁶⁴ As noted, more radical enhancements, e.g., diminishing empathy in soldiers, could disrupt social institutions. Similarly, if human longevity were suddenly doubled there would be disruptions to practical institutions such as social security. However, at this point I am simply arguing that there is nothing wrong with enhancement per say, although I concede that there may be disrupting effects to specific, radical, enhancements.

⁶⁵ BUCHANAN, *supra* note 43, at 161.

⁶⁶ *See id.* at 200, for some formulations of the precautionary principle.

context of human enhancement technologies. Indeed, several authors have done so.⁶⁷ There are many formulations of the precautionary principle. In general, the various formulations of the precautionary principle revolve around two issues: 1) how to allocate the burden of proof; 2) how to determine the parameters of what constitutes an acceptable degree of risk.

Strong Precaution

One formulation of the precautionary principle holds that it is up to the proponent of a new—and potentially risky—technology to prove that the technology will not be harmful. Several groups have attempted to engrain a strong precautionary principle in public policy. The 1982 United Nations World Charter for Nature states that “where potential adverse effects are not fully understood, the activities should not proceed.”⁶⁸ In 1998, a conference of environmentalists published their own strong wording of the precautionary principle: “precautionary measures should be taken even if some cause and effect relationships are not established scientifically.”⁶⁹ This is a *strong* version of the principle because it requires a precautionary approach if there is even a small risk of harm. Strong versions of the principle are controversial because they can easily be taken as a near total ban on technological innovation. In the realm of human enhancements a strong precautionary principle would almost certainly require abstinence from research and

⁶⁷ See, e.g., NICHOLAS AGAR, *HUMANITY’S END* 157 (2010).

⁶⁸ 1982 World Charter for Nature, G.A. Res. 37/7, U.N. Doc. A/RES/37/7, at 11(b) (Oct. 28, 1982).

⁶⁹ WINGSPREAD CONFERENCE ON THE PRECAUTIONARY PRINCIPLE, WINGSPREAD STATEMENT ON THE PRECAUTIONARY PRINCIPLE (1998).

experimentation because every innovation brings with it the risk of negative consequences. One can always imagine a “worst case scenario,” and under a strongly precautionary principle this would require the proponent of the innovation to disprove the threat of every catastrophic outcome raised by critics. This is an insurmountable obstacle even for the most benign seeming innovations. A medical innovation such as vaccination would never pass muster under such a strong precautionary regime. Indeed, the whole range of historical enhancements—from agriculture to the printing press—would be unable to meet the burden of proof. The fact that even the most taken-for-granted human enhancements—things that no reasonable person would argue should have been forgone in the past due to some cautioned threat of societal disruption and harm—are unable to withstand the force of a strong precautionary principle is arguably reason enough to dismiss the principle outright.

The biggest problem with strong precaution is that it creates paralysis. Indeed, as Cass Sunstein points out, the problem is worse than paralysis because doing nothing also creates risks.⁷⁰ This creates a cascade of never-ending risks and counter-risks. As such, strong precaution offers no practical guidance for how we ought to act. However, there are other formulations of the precautionary principle that are not so easily dismissed.⁷¹

⁷⁰ Cass R. Sunstein, *The Paralyzing Principle*, 25 REGULATION MAGAZINE 32, 34 (2003).

⁷¹ There are many variations of the *strong* precautionary principle and not every variation is as susceptible to the criticisms I have given here. However, for my purposes it is not necessary to address each variation in turn. Rather, I make the working assumption that most strong versions of the principle are difficult to defend and that weak formulations provide more realistic means of taking precautions without stagnating innovation.

Weak Precaution

So-called *weak* precautionary principles contrast quite sharply with their strong counterparts. Weak formulations of the precautionary principle, like their strong counterparts, come in many varieties. But it is fair to say that there are three commonalities that define a precautionary principle as weak. All three features are in contrast to the commonalities underlying strong formulations of the precautionary principle. These commonalities are as follows: 1) some threshold of potential harm that must be met; 2) the likelihood of the potential harm coming about must be based in (at least) somewhat substantial evidence; 3) the burden of proof is placed on the party advocating for precautionary action.⁷² Under a generally weak precautionary principle, an opponent of some human enhancement technology would have to prove the threat of harm, and that harm would have to be more than speculative, in other words, there would have to be some basis for thinking the threat of harm is real and there would have to be reason to believe that the potential harm was, at least, somewhat severe.

Consider how this weak principle would apply to an opponent of vaccinations. The opponent would first have to specify some potential harm that is worth taking seriously if it is a legitimate threat. The statement that “vaccines are unnatural” is, as I

⁷² See LINDA CAMERON, ENVIRONMENTAL RISK MANAGEMENT IN NEW ZEALAND – IS THERE SCOPE TO APPLY A MORE GENERIC FRAMEWORK? 12 (2006), <http://www.treasury.govt.nz/publications/research-policy/ppp/2006/06-06/tpp06-06.pdf> (citing the Rio Declaration and the United Nations Framework Convention of Climate Change as examples of weak precautionary principles). For more in-depth discussions of the weak precautionary principle, see Noah M. Sachs, *Rescuing the Strong Precautionary Principle From Its Critics*, 4 U. ILL. L. REV. 1285, 1292 (2011). The European Union uses a version of the weak precautionary principle as derived from the 1992 Rio Declaration. See Consolidated Version of the Treaty of the Functioning of the European Union, art. 191, Sept. 5, 2008 O.J. (C 115) 47, 132.

have previously shown, not a legitimate threat on its own. On the other hand, the oft-touted (and fully debunked) warning that vaccines cause autism is surely a threat worth taking seriously. But under a weak precautionary regime, simply identifying a threat is not enough. It would then be up to the opponent of vaccinations to demonstrate that the threat of autism resulting from vaccinations is real. This would require identifying some evidence that vaccinations have been shown to cause autism in humans or, if we imagine that the worry was raised before the implementation of vaccines in humans, evidence that there is good reason to believe that vaccines will cause autism in humans.

When viewed in the context of enhancements that have already proven to be beneficial (with minimal accompanying harms) it is easy to think of a weak precautionary principle as superior to a strong version. When we have been the beneficiaries of the benefits provided from some enhancement for nearly a century we are unlikely to think that the enhancement should not have been implemented, even if we know that there were significant risks of harm at the time of implementation.⁷³ In evaluating a weak formulation of the precautionary principle it is much more enlightening to consider how such a principle would apply in the case of a new enhancement that has not yet been implemented—thus avoiding hindsight bias. When viewed in this context the limitations

⁷³ While this feeling of having made the right choice has some pull, it is not rational. Saying that we are glad that we made some choice because the consequences turned out to be beneficial is different than saying that the *ex ante* risks *at the time* would not have cautioned against implementing some technology. As a crude analogy, imagine a person is playing roulette and decides to bet his entire fortune on black. Even if the ball lands on black and the person wins, it still would not be the case that at the time of the bet the person made a reasonable choice since the risk of losing far outweighed the chance of winning.

of a weak precautionary principle that places the burden of proof on the opponent of a new innovation become apparent.

Surely, few of us would want to live in a society where scientists had *carte blanche* to engage in testing of human enhancement technologies. Even if these scientists were devoid of bad motives there is always the risk that the striving for success—even if that striving is towards the betterment of humanity—might make a person blind to the risks involved in pursuing some enhancement. Thus, we generally expect scientists to be subject to some kind of regulatory regime that requires them to consider the risks involved in their research. Almost all scientific research in the United States is subject to some form of regulatory regime. In fact, most researchers are subject to several forms of regulation. For example, a drug researcher working at a university who hopes to develop a drug that increases cognitive capacity in humans would be subject to at least three forms of regulation. She would first be subject to the ethical guidelines that govern her field—a form of self-regulation. She would also be subject to the research guidelines provided by her university, which would likely require her to demonstrate at least a *prima facie* case for the safety of her research. Finally, if she wished to have her drug introduced into the population she would be subject to the governmental regulatory regime administered by the Food and Drug Administration.

One reason regulation in fields involving technological innovation is so ubiquitous is likely the fact that we believe researchers who are working on innovations that involve risks have a positive duty to take these risks into account when making decisions. In this respect, many of us are unlikely to feel comfortable with a precautionary regime that places the burden of proof entirely on the opponents of an

innovation. Thus, it seems that weak formulations of the precautionary principle—like their strong counterparts—suffer from many shortcomings that makes it difficult for us to square them with our intuitive feelings about how individuals and institutions allocate (and should allocate) risk-prevention strategies. However, before entirely abandoning precautionary principles as realistic means for guiding innovation I would like to discuss one other formulation of the precautionary principle that has received much attention.

Maximin Precaution

Thus far I have discussed versions of the precautionary principle that deal with two principal issues: 1) allocating the burden of proof; 2) sketching the parameters of what constitutes an acceptable degree of risk. I have already provided a brief description of the difficulties associated with the first issue. The second issue provides its own set of difficulties. Recall that under strong formulations of the precautionary principle even speculative harms must be taken seriously and disproven by the proponent of the innovation in question. I have discussed the difficulties with this approach. Yet, while weak formulations of the precautionary principle provide more intuitively plausible standards for what counts as a risk worth taking seriously, they fail to provide any action-guiding principles. Under a weak precautionary regime we are still left with the following questions: what exactly counts as a risk worth taking seriously? And relatedly, once we have identified a risk that is worth taking seriously, how should we go about determining whether to move forward with the innovation in question?

Stephen Gardiner has attempted to provide a formulation of the precautionary principle that helps answer these questions by providing action-guiding principles.⁷⁴ Gardiner bases his formulation on John Rawls' description of the maximin principle, which means "maximize the minimum." In other words, the maximin principle has us consider the range of options before us and then choose the course of action that has the best worst outcome. The maximin principle is best understood by way of example, which will make its intuitive appeal quite clear.

Imagine you are a football player who is offered two contracts. Under the first contract you are offered \$10 million over the course of four years, but the contract is severed if you get injured during that time, in which case you do not receive the remainder of the money. Under the second contract you are offered \$15 million over a ten-year period with a guarantee that you will receive the full payment over that ten-year period even if you are injured. The maximin principle would dictate that you choose the second contract.

If you choose the first contract and you do not get injured you will potentially end up earning far more money than you would under the second contract. However, if you are injured within the four-year period the maximum you could receive would be \$10 million (assuming you are injured near the end of the fourth year). Furthermore, an injury would likely make you a far less attractive candidate for other teams following the completion of your first contract. In the worst-case scenario you will end up with \$0 (assuming you are injured very early on in pre-season training), and you will be a poor

⁷⁴ Stephen Gardiner, *A Core Precautionary Principle*, 14 THE JOURNAL OF POLITICAL PHILOSOPHY 33 (2006).

candidate for recruitment by another team due to your injury. On the other hand, under the second contract the worst thing that could happen would be that you get severely injured and become ineligible for play with any team after the end of your contract, yet you would still be left with \$15 million at the end of the ten-year period.

This example gives some intuitive plausibility to the maximin principle as action guiding. But I deliberately crafted the example to make the right choice seem fairly obvious from the football player's perspective. Now imagine the situation is changed such that under the first contract the football player would receive \$20 million over four years (with the same constraints as the prior example) and under the second contract he would receive only \$500,000 over ten years (\$50,000 per year). The choice seems far less clear now. In the first example the difference in the two contracts was the difference between ending up potentially extremely rich and ending up very rich. In the second example the difference is between ending up potentially enormously rich and ending up solidly middle class. However, under the maximin principle *the football player should still choose the second contract*. Even though the potential gains have shifted drastically in the second example it is still the case that choosing the second contract leaves the football player with the best worst-case scenario.⁷⁵

One can easily see the beginnings of an argument ad absurdum here. Gardiner provides another tweak that makes the irrationality of adopting a maximin strategy as a

⁷⁵ Initially, this might seem reasonable. Especially when extended to choices involving societal policy as a whole, all things being equal, the best worst-case scenario is what we would hope politicians would strive for. The problem for the maximin principle, as I have described it thus far, is that it does not operate as all things being equal. As I will describe in the next few paragraphs, this initial formulation of the maximin principle requires choosing the best worst-case scenario regardless of the size of the risk involved.

general principle even more clear.⁷⁶ Imagine that in the first two-contract scenario described above it is now the case that both contracts include a provision for injury, in other words, under both the four-year \$10 million contract and the ten-year \$1 million contract the player will receive the full amount even if he is injured and unable to play. But now suppose that accepting the first contract requires the player to fly to Minneapolis, while the second contract allows the player to stay in his home city of Milwaukee. In this case the maximin principle would dictate that the player choose the second contract—despite the clearly superior financial incentives of the first contract—because the fact that he would have to fly to Minneapolis means that he might get in a plane crash, thus turning the first contract into the option with the *worst* worst-case scenario.

Moving back into the realm where precautionary principles typically apply—namely, decisions that have the possibility of negatively affecting the environment—it is not difficult to think of examples where maximin thinking would be fatal to experimentation and innovation. A prescient example of this kind of thinking involves the Large Hadron Collider (LHC). Prior to the collider being turned on critics picked up on subtle nuances in some of the calculations respecting particle collisions at extremely high energy that suggested that there was an extremely small possibility that such collisions would create mini black holes, and an even smaller possibility that those mini black holes would cause a chain reaction of black hole creation that would subsume the entire earth. Wild stuff. But under the maximin principle, the small risk would be sufficient to derail the entire collider operation. In order to avoid the paralyzing effects of

⁷⁶ Adapted from Gardiner, *supra* note 74, at 46.

a generalized maximin principle, Gardiner provides three criteria that he suggests should be employed when adapting maximin thinking to a precautionary context.⁷⁷

First, maximin thinking should only be employed in contexts where our understanding of the likely probabilities of some outcome are either entirely lacking or lacking to such a degree that we can do little better than employ what I call linguistic probability heuristics (e.g., “what are the chances?!”). In the latter case of the football player deciding on a contract where signing the first contract would require him to fly to Minneapolis, this first criteria would exclude maximin thinking as a viable decision-making tool given that we have a fairly good understanding of the likelihood of a plane crashing. With respect to the LHC it is even more clear that this criterion allows us to discount the general maximin principle because in that case we know precisely the likelihood of the black hole nightmare scenario—the equations tell us the probabilities to a tee. On the other hand, this criterion does not necessarily place human enhancement technologies outside the reach of general maximin thinking because with many of these technologies there is simply no way to determine the likelihood of many of the potential negative outcomes that opponents worry about.

Consider again the *analogy of the master engineer of evolution*, which states that “the human body and mind, highly complex and delicately balanced as the result of eons of gradual and exacting evolution, are almost certainly at risk from any ill-considered attempt at ‘improvement.’” Ignoring for a moment the criticisms I have already leveled against this view, it seems that the word “risk” here cannot refer to some quantifiable probability of a negative outcome. Rather, the use of risk here appears to be a stand-in for

⁷⁷ Gardiner adopts these criteria from JOHN RAWLS, *A THEORY OF JUSTICE* 134 (1999).

uncertainty, namely, the uncertainty that comes with attempting to make changes to a system whose mechanisms and interconnections we do not fully understand. I would not go so far as to classify this as a linguistic probability heuristic, however. Instead, it is a worry that our lack of knowledge justifies precautionary thinking because that very lack of knowledge entails that disaster might follow.

Sandel's concern with a waning appreciation for the giftedness of life is another example of a negative outcome for which it makes little sense to speak in terms of probabilities. If anything at all can be said in terms of probabilities from Sandel's point of view, it is that there is a 100% likelihood that a quest for mastery over life will lead to a culture that lacks a sense of giftedness and is, thus, lacking in an important moral quality. But this, of course, is making fast and loose use of the term "probabilities," since Sandel's worry is based not on science but rather on rhetoric and philosophical argument.⁷⁸ Thus, under Gardiner's first limiting criterion for the use of the maximin principle as a guide to decision making, Sandel's concern makes the cut.

The second limiting criterion (related to the first) proposed by Gardiner is that the maximin principle should only apply to situations where the predicted negative outcome is "unacceptable." By this Gardiner means to distinguish between outcomes that would be negative but that would not be considered utterly disastrous and outcomes that no reasonable person would be willing to live with.⁷⁹ On the face of it, it might appear that this condition does little to help us out of the absurdity of applying maximin thinking to

⁷⁸ See BUCHANAN, *supra* note 43, at 2-6, for a critique of Sandel's (and other enhancement critics') reliance on "murky rhetoric."

⁷⁹ Gardiner, *supra* note 74, at 47.

situations where there is *some* likelihood of a disastrous outcome even if we are unable to quantify that likelihood. Consider that even though we excluded the case of the football player deciding between two contracts—one of which requires him to fly to Minneapolis—under Gardiner’s first limiting criterion, it is easy enough to tweak the hypothetical such as to make the possibility of a negative outcome fall under the category of uncertainty rather than risk. For example, we might say that instead of the risk of the plane crashing there is a possibility that another passenger on the plane will have a mental breakdown and stab the football player in the heart. This is an outcome whose likelihood cannot be quantified⁸⁰ so it makes it past Gardiner’s first criterion and also past the second criterion because clearly an outcome where the football player is stabbed to death on a plane is unacceptable.

⁸⁰ There are several problems with trying to quantify such as outcome as a risk rather than an uncertainty. First, no one has ever (as far as my research indicates) had a mental breakdown on a plane and stabbed another passenger. But, of course, the fact that something has never happened before does not mean that it will not happen in the future. However, when we have no past occurrences of specific human behavior to work with it becomes a guessing game to determine the likelihood that some behavior will occur in the future. With plane crashes, probabilities are determined by dividing the total number of plane crashes (of the number of crashes within a specified period) by the total number of flights. This gives us a precise number. With the scenario of someone having a mental breakdown and stabbing another passenger, the best we can do is use a linguistic probability heuristic such as, “it seems pretty unlikely.” One might argue that there are statistics about murder rates in general and about rates of mental breakdowns in general, and thus we can extrapolate those general numbers to the scenario of someone having a mental breakdown on a plane and stabbing another passenger. The problem is that murder rates (and rates of mental breakdowns) are highly context dependent. *See, e.g.,* ERICH GOODE, *DEVIANT BEHAVIOR* 113 (2015). The problem is further compounded when we consider specific means of murder, e.g., the chances of being killed by a gun as opposed to a knife. We can come up with something like a specific murder rate for planes (the 9/11 hijackings, for example, would count as murders) but we cannot do so specifically for murders committed by someone having a mental breakdown and stabbing another passenger. *See generally* Cass Sunstein, *The Limits of Quantification*, 102 CAL. L. REV. 1369, II.B (2014).

Gardiner attempts to dispel the difficulties provided by the second criterion by suggesting that the worry about unacceptable outcomes should only be applied to those outcomes that are realistic.⁸¹ He acknowledges that we are unlikely to be able to construct necessary and sufficient conditions for what constitutes a realistic outcome, as opposed to a fanciful one, but he argues that this does not stop us from recognizing cases where the threshold has clearly been crossed. As an example, Gardiner points to the threat of global climate change.⁸² In recent decades we did not fully understand what impact continuing industrialization might have on the climate, yet there was plenty of science showing how an increase in greenhouse gases in the atmosphere might lead to an overall increase in the global temperature. Even though the threat of global climate change could not be quantified, there was enough good science for us to think that the threat was real, and thus the maximin principle would dictate a precautionary approach.

This example represents a situation where the potentially negative outcome is entirely realistic. But let me return again to the example of the football player who is deciding between contracts. Recall the situation where the player has the option of a better contract, but he must fly across the country to sign it—thus taking on the risk of getting into a plane crash. In that scenario, it certainly seems realistic that the football player might die in a plane crash if he flies to Minneapolis. Thus, the “realistic outcome” criterion does not appear to save the maximin principle on its own. Many outcomes are realistic, in the sense that they have happened before and could happen again, yet not sufficiently probable such that they should guide our decision-making processes.

⁸¹ Gardiner, *supra* note 74, at 51.

⁸² *Id.* n.62.

Even more troubling, when we attempt to apply the “realistic outcome” criterion to the concerns related to human enhancement technologies that I described previously, the result is far less clear-cut. Is the potential for catastrophic damage to our biology (and, thus, to our ability to survive as a species) as a result of tampering with the slow pace of evolutionary change realistic? Surely, a concern about such catastrophe is not entirely absurd. But absurdity presents a rather low bar. Are we to discount only those outcomes that depend on utterly far-fetched scenarios such as aliens descending to earth in order to save us from our Promethean hubris? Or should we let the error fall in the other direction and only consider those outcomes that any reasonable person would consider to be realistic? And what about more abstract concerns such as that put forward by Sandel? Is the threat of losing our sense of “giftedness” a sufficiently realistic outcome of our attempts to enhance ourselves?⁸³ It is not clear that any of these questions can be answered by Gardiner’s criterion.

If I am being charitable to Gardiner, the best response from the perspective of the “realistic outcome” criterion would be my suggestion that only those outcomes that any reasonable person would consider realistic should count towards applying the maximin principle to the choice at hand. As such, it would seem that the large majority of objections to human enhancement technologies would be thrown out (insofar as they would be susceptible to maximin thinking). However, for the moment I will assume that objections to human enhancement technologies withstand the “reasonable outcome”

⁸³ Losing our sense of “giftedness” would not likely not qualify as an unacceptable outcome to Rawls, who is concerned with more practical concerns such as worldwide environmental catastrophe or the widespread enslavement of human beings, but Sandel surely views losing this sense as unacceptable.

criterion. This will allow me to address Gardiner's third criterion, which illustrates what I consider to be the fundamental issue that surrounds objections to (certain forms of) human enhancement technologies.

Gardiner's third criterion for a decision being subject to maximin thinking is that the potential upsides of not taking a precautionary approach are relatively small. As Gardiner puts it, under this third criterion "the decision-makers care relatively little for potential gains that might be made above the minimum that can be guaranteed by the maximin approach."⁸⁴ Buchanan frames the maximin principle under Gardiner's third criterion as follows: "to follow the maximin rule is to act as if one were extremely harm averse—as if all that mattered was avoiding harm."⁸⁵ Perhaps the most charitable way to formulate this criterion is to say that it means that maximin thinking should only apply in situations where the best worst outcome is good enough that it is not worth the risk of trying for something better. Putting it this way allows the maximin principle to apply to many of the environmental contexts that the precautionary principle was designed for in the first place. For example, the maximin principle—under this third criterion—would almost certainly require that greenhouse gas producing companies take measures to reduce the release of pollutants into the air and to mitigate the effects of the pollutants that the company cannot stop from being released into the atmosphere.

The best worst case scenario in that situation is that it turns out that we were wrong about the devastating impact of greenhouse gases and, thus, the companies' unnecessarily spent money on pollution reduction and mitigation—costs that were surely

⁸⁴ *Id.* at 47.

⁸⁵ BUCHANAN, *supra* note 43, at 202.

passed on to customers. This scenario does not seem so bad, especially when we compare it to the best best case scenario—where consumers save money by virtue of companies’ not having engaged in pollution reduction and mitigation—and then to the worst worst case scenario—where the environmental science was right and companies’ failure to engage in pollution reduction and mitigation leads to wide-spread environmental damage. This is the kind of case that maximin thinking was built for. Indeed, environmentalists’ common rejoinder to people who claim that global climate change is not man made and we can do nothing to prevent it is this: what’s the harm in trying?

At “worst,” requiring companies to reduce and mitigate pollution will lead to a cleaner and more enjoyable planet, even if it turns out that we can do nothing to stop the ultimate catastrophic effects of global climate change. And the risk of foregoing the use of polluting technologies seems quite minimal—it can be measured in a few dollars paid by consumers. Thus, applying a maximin approach in this situation seems entirely reasonable, even commonsensical. At the same time this situation seems wholly different from one in which the technology in question is being developed or implemented, not for its minor economic value over other means, but for the purpose of creating a significant and tangible benefit.

For example, consider a technology that weaves together metal and computer circuits with the human body in such a way as to allow a person to lift more weight or to walk or run for very long periods of time. This sort of cybernetic enhancement would allow not only for workers, soliders, and otherwise fit individuals to work more efficiently—a benefit more parallel to the environmental example described above—but would also allow for older individuals to maintain their healthy functioning for much

longer and would provide relief to many people, even healthy people, who suffer from the aches and pains that are endemic to human existence. These latter benefits are significant and far-reaching in their implications for the quality and duration of human life and are thus of a different kind from the benefits of minimal added productivity and cost-savings benefits that were discussed in relation to polluting technologies. In this case, Gardiner's third criterion fails: when it comes to technologies that have such transformative potential, we are very concerned—and rightly so—with the potential upsides of the technology. As Buchanan states: “enhancements may be needed not just to improve our situation relative to the status quo, but also to sustain the good we now enjoy[.]”⁸⁶ Thus, when it comes to human enhancement technologies, “we do care about gains, not just losses.”⁸⁷

Conclusion

In this chapter I have attempted to assuage some of the concerns raised by opponents of human enhancement generally. In the next chapter I will delve specifically into empathy enhancement, and I will argue that empathy enhancement at least represents a means of

⁸⁶ *Id.*

⁸⁷ *Id.* There are several ways to interpret this statement. From the perspective of the individual, it might not be worth a 1% risk of immediate death to obtain a 99% chance of living a life with better quality and longer duration. However, from the perspective of society as a whole, having one in one hundred people die suddenly might be an acceptable trade-off for a far-reaching benefit. In any case, the point is that neither the individual nor society as a whole would take only the risks into consideration in making the decision. The potential benefits would play a large role in the decision-making process. See Cass Sunstein, *Cost-Benefit Analysis and the Environment*, 115 ETHICS 351, 382 (2005), for further discussion of why Gardiner's third criterion should generally be abandoned.

achieving some of our moral goals and at most represents a moral imperative once the means to achieve it become available.

Chapter Two

EMPATHY AND THE EXPANDING CIRCLE

Introduction

In this chapter, I will consider the role of empathy in our moral obligations to others. I will begin by discussing the role of evolution in morality, including its limitations. I will present Allen Buchanan's description of a "human enhancement enterprise," which includes the view that certain enhancements are required—either morally or prudentially—if we are to survive as a species or, less prosaically, if we wish to achieve the goals of justice, peace, and other higher-order goods. Within that context, I will consider Peter Singer's notion of "the expanding circle" of concern for others. I will present an argument that artificial enhancement of human empathic response may provide a means by which to expand our sphere of concern—and thus our willingness to help others who may fall outside our "naturally" imbued empathic response.

Ultimately, in this chapter I hope to persuade the reader that a human enhancement project that seeks to enhance human beings' ability to experience empathy is a worthwhile project to pursue. Indeed, I will argue that we may be morally obligated to pursue such a project if the technology to do so becomes available. As with the preceding chapter, my purpose here is to provide a motivating groundwork for the rest of this dissertation.

Evolution and the Moral Compass

Evolution, as discussed in the preceding chapter, cannot be relied upon to produce morally desirable outcomes. Any tendency towards morality in human beings is a result

of, to begin with, random chance and thereafter, the further randomly selected-for traits that allow human beings to engage in moral thought.

Evolution itself sets no goals, let alone goals that might align with what we have determined—through use of our reason and moral thought¹—to be the ideals towards which humanity should strive. As Richard Dawkins put it: achieving a moral world requires that we “rebel against the tyranny of the selfish replicators.”² Dawkins is not being literal when he refers to genes (and, in turn, evolution) as “selfish.” Rather, the word is used to emphasize the point that evolution does not care about our needs and wants. Worse than being selfish, however, evolution is wholly indifferent to human suffering, human goals, and the human sense of morality. Even if it is true that the “arc of the moral universe . . . bends towards justice,”³ it does so entirely as a result of our human efforts to overcome the nasty and brutish instincts laid down by the indifferent forces of evolution.

Of course, it is not the case that evolution has created us as entirely amoral beings who are wholly indifferent to the suffering of others. Rather, a simple kind of morality is built into our genes, regardless of our ability to reason explicitly about right and wrong. Indeed, it seems quite impossible to imagine that our species could have survived the early days of our existence if there were not some built-in moral compass that provided at least a weak guide for human behavior. Indeed, there is plenty of evidence indicating that

¹ I take no stance here about the interconnection, or lack thereof, of reason and moral motivation. Suffice to say that reason and moral motivation, whether related or not, allow us to think and act outside the bounds of evolutionary dictates.

² RICHARD DAWKINS, *THE SELFISH GENE* 205 (1976).

³ Speech by Martin Luther King, Jr., Montgomery, AL, March. 25, 1965.

human beings are not alone in possessing a biological moral compass. Cooperation and reciprocity have evolved in many species. Monkeys and apes have been shown to respond to situations of unfairness with similar indignation as that of human beings.⁴ Other mammals, too, appear to have a biologically driven moral imperative that might make us nod our head in approval of their seemingly moral decisions in the face of injustice.⁵ We cannot know what manner of conscious experience attends the indignant experience of mammals other than ourselves, but I do not think it is too great a leap to suggest that the seemingly moral actions of other mammals are likely driven more by unconscious biological processes than by conscious thought.

At first, it might seem odd to consider that behavior that strikes us as demonstrably moral in nature could be the result of a dumb biological moral compass that reacts to situations of injustice without truly recognizing them as such. But it ought not to. Though we tend to pat ourselves on the back when we engage in behavior that we consider to be particularly altruistic, the truth is likely that most of the moral behavior we engage in from day-to-day is unconscious. This should not come as a surprise given that so much of the behavior that we associate with our “higher” reasoning abilities can occur quite well without the direction of the conscious mind. Who hasn’t had the experience of arriving home from work only to realize that one has little to no memory of the trip? While this kind of automaton-like behavior can be alarming and a bit disturbing at times,

⁴ See Sarah F. Brosnan & Frans B.M. de Waal, *Monkeys Reject Unequal Pay*, 425 NATURE 297 (2003).

⁵ See generally Sarah F. Brosnan & Frans B.M. de Waal, *Evolution of Responses to (Un)Fairness*, 346 SCIENCE 1251776 (2014).

it is certainly not uncommon. In reality, the majority of our being in the world appears to be devoid of conscious experience or intervention.⁶

None of this is to suggest that apes and other mammals are pure automatons who lack conscious experience of the world and their place in it. Rather, I am simply making the argument that moral behavior—or, more precisely, what we would view as moral behavior from a third-person perspective—can likely exist as a purely biological phenomenon, driven by the forces of evolution that would lead to a moral decision being selected for its overall reproductive fitness.⁷

The Limits of Evolutionary Ethics

There are dangers in this line of thinking, however. Peter Singer has said, “[e]thics is part of the natural human condition.”⁸ Yet, Singer is acutely aware of the limitations of this “natural ethics.” For Singer, if our ethical principles depend for their justification on a foundation of biology then we are lost. If evolution provides the terms of what constitutes moral behavior by rewarding behavior that leads to greater fitness, and we recognize that much behavior that our rational minds would consider to be immoral also might produce greater fitness, we leave ourselves open to the conclusion that morality should be defined

⁶ See John A. Bargh & Ezequiel Morsella, *The Unconscious Mind*, 3 PERSP. PSYCHOL. SCI. 73, 76 (2008) (“[S]ocial cognition research over the past 25 years has produced a stream of surprising findings regarding complex judgmental and behavioral phenomena that operate outside of awareness.”).

⁷ Whether this fitness applies to groups, individuals members of groups, or individual genes is not germane to my interest here.

⁸ PETER SINGER, *THE EXPANDING CIRCLE: ETHICS, EVOLUTION, AND MORAL PROGRESS* 23 (2011).

just as behavior that leads to reproductive fitness.⁹ Examples of this kind of thinking are replete throughout the history following Darwin’s great discovery. “Social Darwinism,” the idea—in part—that we are justified in promoting policies that serve to bolster the “stronger” members of society while culling the “weaker” members because evolution would naturally do the same work, has been used by countless despots to justify barbaric programs of genocide and imperialism. And many academics have taken the idea—albeit, in a more palatable form—seriously and proposed arguments in favor of policies that work in tandem with the natural flow of evolution. For example, Edward O. Wilson, in his book *Sociobiology* and other works, has argued that “the time has come for ethics to be removed temporarily from the hands of philosophers and biologicized.”¹⁰

Wilson and other defenders of a “biologicized” basis for ethics have been thoroughly lambasted for their views, which depend in large part on the naturalistic fallacy—that an ought can be derived from an is.¹¹ Singer has taken a sober look at the arguments presented by Wilson and other defenders of the view that ethics is, at its root, derived from biological imperatives. In *The Expanding Circle*,¹² Singer argues that if Wilson’s view of ethics holds, then “ethics ultimately rests on subjective judgments

⁹ *Id.* at 84.

¹⁰ EDWARD O. WILSON, *SOCIOBIOLOGY* 287 (1980).

¹¹ See, e.g., Elizabeth Allen et al., *Against “Sociobiology,”* NY REVIEW OF BOOKS (Aug. 7, 1975), <http://www.nybooks.com/articles/archives/1975/nov/13/against-sociobiology/> (“[Wilson] present[s] [us] with yet another defense of the status quo as an inevitable consequence of ‘human nature.’”).

¹² SINGER, *supra* note 8.

which are immune from criticism.”¹³ This conclusion follows from Wilson’s argument, which holds that only an ethics that is tied directly to a biological imperative has objective force.¹⁴ As such, Wilson argues that whatever else we might call “ethics” is nothing more than an outpouring of emotions for which we create post hoc justifications.¹⁵

If Wilson is right about ethics, then these post hoc justifications are nothing more than an attempt on our part to convince ourselves that our actions and decisions are more than simply the result of crude, subjective, emotional responses. Under this view of the world, once we recognize the truth of our predicament we would have to accept that we are entirely unjustified in judging the actions and decisions of others based on any objective standard—other than the “objectivity” of a biological imperative. In that case we would be left with the conclusion that all decisions based on “ethics” are equally justifiable in that all of them are entirely *unjustifiable*. As Singer puts it: “If ethical judgments were nothing but the outflow of our emotional control centers, it would be as inappropriate to criticize ethical judgments as it is to criticize gastronomic preferences.”¹⁶ Both would be equally lacking a deeper foundation than the contingencies of nature and nurture.

¹³ *Id.* at 85.

¹⁴ *See* WILSON, *supra* note 10, at 3.

¹⁵ *Id.*

¹⁶ SINGER, *supra* note 8, at 85.

For Singer—and presumably for most people—this outcome is as unacceptable as it is unsupported. There is a long tradition of philosophical thinkers who have argued that ethical judgments are, as least some of the time, a result of our reasoning about right and wrong. In other words, there is a distinctly rational component to ethical judgments that distinguishes those judgments from mere emotional responses. The dominant view among philosophers is that our ability to reason allows us to choose the foundations of our ethical judgments—our “ethical premises,”¹⁷—in a nonarbitrary way.¹⁸ Thus, we are able to see ethical judgments as a form of argument. Where the premises do not support our ethical conclusions we are rightly subject to criticism. This is not to say that people always, or even most of the time, conform their behavior to the dictates of the ethical conclusion that is most supported by rational argument. The fact that we are able to think rationally about something does not imply that we will make use of that ability. And even where we do purport to be engaging in rational thought about our ethical judgments, we are still faced with a panoply of competing factors—ego, prejudice, habit, and plain stubbornness—that often interfere with our reasoning about ethical issues.

It is telling that when faced with strong arguments that challenge our ethical judgments we generally experience some manner of cognitive dissonance. Our reaction to these situations is often something like, “I know you’re wrong, but I don’t know why,” or, “I know that there are more persuasive reasons that justify my ethical judgments but I just can’t come up with them at the moment.” Interestingly, our strain to justify (at the

¹⁷ *Id.* at 86.

¹⁸ *Id.*

very least, to ourselves) the soundness of our ethical judgments even in the face of irrefutable argument weighs in favor of the view that ethical judgments are rational in nature. If ethical judgments were just emotional outpouring, we would feel no more need to justify them than we do to justify our gastronomic preferences. Unlike our tastes in food, which we generally recognize to be contingent and subjective preferences, we feel a stronger attachment to our ethical beliefs—an attachment that compels us to attempt to justify our beliefs.¹⁹

The Expanding Circle

Singer is most concerned with how reason dictates our obligations to one another. According to Singer, rational thinking about our ethical duties leads to the conclusion that many of our commonsense attitudes are misguided. For example, he argues that there is no defensible reason to discount the suffering of non-human animals.²⁰ As the argument goes, there are many human beings who have a level of cognition that is equal to or more diminished than many non-human animals. We would rightly find it abhorrent to say that due to this diminished capacity it is permissible to test experimental drugs on these people or to use them as a source of food. At the same time, many of us are perfectly willing to allow non-human animals—with the same level of cognition as young children or some disabled adults—to be used as means to our own ends without taking into account their suffering.

¹⁹ There is, of course, another way to justify our ethical judgments—religion. Religious teachings form the objective ethical premises for many people. I will not be addressing this third way.

²⁰ See generally, PETER SINGER, *ANIMAL LIBERATION: A NEW ETHICS FOR OUR TREATMENT OF ANIMALS* (1975).

For Singer, our revulsion at the former and our casual acquiescence in the latter is based entirely on irrational emotional reactions. We are evolutionarily predisposed to feel an attachment to members of our own species, above and beyond that which we feel towards members of other species. While this may be a reason for our preference, it is not a rational reason. To see it as such would be to fall right back into the trap of the naturalistic fallacy and the argument made by Wilson, that the ultimate ethical premise just is the biological imperative handed down by evolution.

Less controversially, Singer argues that rational consideration of our ethical obligations towards one another leads inevitably to the conclusion that we are unjustified in extending our altruism to the people we know to a different degree than we extend that same altruism to other, less well-known people.²¹ He is not oblivious to the fact that our kin preference is a result of evolutionary pressures which dictated that those who preferred close kin to outsiders were more likely to survive. During the period that human beings arguably were subject to the greatest evolutionary pressures our species was composed of small tribes of hunters and gatherers. These tribes competed for relatively scarce resources and each individual depended on the contribution of other members of the group. Individuals within one group had nothing to gain from individuals within another group—indeed, they had much to lose. A human living during this time who was especially trusting of strangers or willing to share the bounty of her labor with members of all groups would likely have perished before reaching reproductive age. Thus, one can easily see how the blind forces of evolution could create a sense of ethical obligation

²¹ See generally, SINGER, *supra* note 8.

between human beings while simultaneously limiting the tendency of humans to share the concern created by those obligations with people outside one's in-group.

As Singer notes, the ambit of our ethical concern has naturally expanded as human beings went from living in small, isolated, groups to densely connected societies. This should not be particularly surprising. The things that justified early humans' loyalty to members of their own group—resource sharing, mutual dependency, etc.—are present in most modern societies. In the United States, for example, we are dependent on our families and friends to some degree. But we are just as dependent on our individual states and the nation as a whole for our wellbeing. While early hunter gatherers could get by with a small circle of ethical concern, since all of their ability to survive was vested in a very few individuals, modern humans cannot afford to exist within such a limited scope of ethical concern.

Reciprocal Altruism

Singer ties much of the development of ethical concern to the concept of “reciprocal altruism.” Robert Trivers, the originator of the term, defines altruism as “behavior that benefits another organism, not closely related, while being apparently detrimental to the organism performing the behavior, benefit and detriment being defined in terms of contribution to inclusive fitness.”²² The problem for evolutionary biologists is that it is counterintuitive to imagine that altruistic behavior would ever develop through evolutionary processes because it appears, *prima facie*, to be a trait that would lead to *decreased* reproductive fitness. Yet, as Trivers points out, “given the universal and nearly

²² Robert L. Trivers, *The Evolution of Reciprocal Altruism*, 46 QUART. REV. OF BIOL. 35, 35 (1971).

daily practice of . . . altruism among humans today, it is reasonable to assume that it has been an important factor in recent human evolution and that the underlying emotional dispositions affecting altruistic behavior have important genetic components.”²³ Trivers presents a convincing theory, “reciprocal altruism,” which explains how altruistic behavior could lead to evolutionary advantages. According to the theory, altruistic behavior that involves an actual benefit to the recipient that is greater than the potential detriment to the actor will provide an evolutionary advantage to the actor insofar as the actor can rely on the recipient to act in a reciprocally altruistic manner at some time in the future when the actor is in need.²⁴

This solution is as simple as it is ingenious. A modified version of a scenario described by Singer exemplifies the wisdom in Trivers’ theory.²⁵ Imagine you walk by a lake and see a child drowning in shallow water. If you do not save the child she has a 95% chance of dying. You have a 5% of dying if you save the child (perhaps you trip and drown in the shallow water). From an evolutionary perspective, on first glance it might seem that the decision that would accrue the greatest fitness benefit would be to let the child drown. After all, 5% is greater than 0%, which is the chance you would die if you did not help the child. *Prima facie*, in the long term evolution should select for the individuals who do not rescue the child. After all, even a 5% death rate can add up over time, leading to a larger number of non-drowning-child-savers in the population. But

²³ *Id.* at 48.

²⁴ *See id.*

²⁵ *See* SINGER, *supra* note 8, at 16. The example is simplified with the hopes that it will be intuitively obvious (without the need for mathematical equations) that reciprocal altruism could be selected for by evolution.

there is more to the story. There is a chance that the would-be drowning-child-saver will one day find herself in a similar predicament, where she is dependent on the altruism of another to save her from drowning. Clearly, it is to her benefit to be saved, and being saved will lead to an increased chance of her genes being passed on to future generations. From the perspective of her genes, the best scenario is one in which she does not risk her life to save another but others risk their lives to save her. But this scenario is not stable, and the individuals who risked their lives without reciprocal acts of altruism being performed in their direction would eventually see their genes die out. The stable scenario is where each person acts altruistically towards others when the risk of death to the actor is less than the risk of death to the recipient of the altruistic act. In that case, altruistic acts are selected for overall and the trait would not be weeded out.

The Limitations of Reciprocal Altruism

Although the scenario presented above is convincing, further analysis shows that it does not call for moral celebration. First, consider the limited reach of a genetically selected-for altruism. This kind of altruism would likely only extend over the individual's in-group, as described above. Indeed, it is not fanciful to suppose that the selection pressures between groups, contrary to those operating within groups, would lead to the extinction of between-group altruistic tendencies. After all, an altruistic act from a member of one group to a member of another group is unlikely to be repaid.²⁶ Thus, a between-group altruist would be like a within-group altruist where no member of the group

²⁶ Recall that this discussion is limited to the time period when human beings lived in discrete groups who rarely interacted, competed for limited resources, and existed during a time of strong overall evolutionary pressures.

reciprocated—a scenario that would inevitably lead to the extinction, or at least the attenuation, of that trait. Singer acknowledges this problem when he says that “group altruism . . . work[s] best when coupled with a degree of hostility to outsiders.”²⁷

Second, there is the problem of cheating.²⁸ As described above, a group where only a few members act altruistically, in other words, where there is not widespread reciprocal altruism, is not a stable configuration from an evolutionary perspective. At the same time, reciprocal altruists within a group are at an evolutionary advantage compared to their non-reciprocally altruistic brethren (again, assuming the risk of engaging in altruistic behavior is outweighed by the benefit that one would receive from a similar altruistic act). If a large number of individuals within a group were cheaters, in other words, they were the recipients of altruistic acts but did not reciprocate, the dynamic would quickly shift towards a non-stable configuration as the few honest altruists died out without passing on their generous genes. The key word here, however, is “widespread.”²⁹

There does not appear to be any evolutionary reason that limited amounts of cheating behavior would not proliferate. It is not difficult to imagine an individual who

²⁷ SINGER, *supra* note 8, at 20.

²⁸ Trivers is careful to define “cheaters” in a purely biological sense, without reference to morality. TRIVERS, *supra* note 22, at 48. For my purposes, this isn’t necessary.

²⁹ Widespread cheating as a losing strategy has been demonstrated using the Prisoner’s Dilemma. Robert Axelrod and William D. Hamilton used two hundred iterations of a Prisoner’s Dilemma game to determine the best strategy. In the end, “tit for tat” won out. In other words, the best strategy was to be fair: don’t be the first to defect; retaliate against defection but immediately forgive when the other player cooperates. While this strategy might not lead to a higher score than one’s partner in a single round, over many iterations it consistently wins out. See Robert Axelrod & William D. Hamilton, *The Evolution of Cooperation*, 211 SCIENCE 1390 (1981).

did not reciprocate but managed to stay under the radar. Indeed, this individual might go unnoticed with no effort on her part. This seems less likely in a smaller group where acts of altruism are routinely shared between individuals, however. But add in the element of cunning, and as groups grow in size the ability to get away with not reciprocating altruism seems less far-fetched. An individual might pretend to act altruistically without ever really risking her life, or she might always be the last one to arrive in times of need. In either case, it is entirely possible that the rest of the group would not notice the cheater. If that were the case, the skillful cheater would be at an advantage. Recall the drowning child example. If the skillful cheater is the child, she would gain the benefit of being saved without the 5% risk of dying involved in helping another. In the long term, and again assuming that the number of cheaters remains relatively small, this would lead to a small advantage in terms of survival and the skillful cheater gene would gain a niche hold.

These two unfortunate outcomes of the evolution of reciprocal altruism in humans—hostility towards outsiders and a small, but not insignificant, tendency to “cheat”—have led to a somewhat paradoxical state of affairs for human beings in the modern world. On the one hand, there is a marked tendency on the part of most people to want to ease the suffering of others when doing so is not too great a sacrifice. It would not be absurd to call this tendency an instinct, given its likely evolutionary origins. At the same time, there is an undeniable instinct in us to want to cheat if we know we can get away with it. Of course, this trait is stronger in some and weaker in others. From a purely evolutionary perspective it is not surprising that there would be a small number of people who almost never even consider cheating, while there are others who are faced with the

constant desire to cheat. Evolutionarily endowed traits are generally not evenly distributed throughout the population. Rather, the distribution is more likely to approximate a bell curve.³⁰

While our acts of altruism are admirable, so our acts of cruelty and indifference—so often directed towards people who are not members of our particular in-group—are deplorable. But again, from a purely evolutionary perspective the fact that human beings seem to be uniquely skilled at showing great compassion and concern for members of our families, friends, and neighbors while simultaneously displaying a sometimes almost psychopathic indifference to the suffering of people in other countries, or people who are of a different race or religious doctrine, should not come as a surprise. This dual nature within us is the essence of tragedy. The nature of that tragedy becomes especially clear in circumstances where we are faced with the hypocrisy of our protectionist distribution of empathy and altruism towards our fellow human beings. The tragedy comes from the fact that we realize irrational forces that we in no way condone when examined in the light of reason have driven us.

From Natural Ethics to Justification-Based Systems

For Singer, and many others, the solution to the problems of cruelty and unequal distribution of resources that plague humankind is the application of reasoned judgment. Our instinct towards altruistic action may have derived from biological evolution, but the consequences of that instinct have led us to a remarkable place. Unlike other species of animals, human beings have taken our evolutionarily granted traits and used them to craft

³⁰ DAVID A. BAUM ET AL., THE PRINCETON GUIDE TO EVOLUTION 223 (2013).

a system of rules that govern our conduct towards one another.³¹ The biological instinct to reciprocate altruistic acts yielded positive cognitive and emotional responses to those who share in our reciprocal altruism, and negative responses to those who do not. These cognitive and emotional responses became judgments, and judgments require justification. Because these justifications applied to types of behavior in general, and were not the idiosyncratic application of subjective preferences, they had to be formed in general, objective, terms. In other words, the justification proffered would only suffice if it applied to all members of the group.

Consider the alternative. Imagine that there is a group of humans living in a small community that resides in a town. The community has rules and generally agreed upon codes of conduct that come from shared notions of right and wrong. It is a community that relies on large group hunts in order to survive. During the hunts nearly every member of the community is gone from the town. Only two or three people stay behind during the hunts, and they are in charge of the town's affairs while the others are away. One day as the members of the town are preparing for a hunt, a member of the community, Dee, decides that she will elect to stay behind so that she can enact her master plan. Dee has realized that the rules of the town are written such that the group that stays behind during a hunt has total control over the town's affairs. Dee realizes that this includes the ability to pass new laws. She has decided to take advantage of the opportunity to pass a law that will provide her with a particularly generous share of the meat from the town's hunts from that point on. Dee convinces two other particularly slow and dull-witted members to

³¹ SINGER, *supra* note 31, at 92. Of course, animals may adhere to systems of rules about how to interact with one another, but these rules are not intentionally created.

stay behind with her because she knows that they will go along with whatever she suggests. The moment she sees the last member of the hunting party cross over the hill heading out of town, Dee calls a meeting of the three members that have stayed behind. She proposes the new law and then calls for a vote. Not understanding the meaning of the law and wanting to stay in Dee's good graces, the two other members of the group that stayed behind vote to pass the law.

As the hunting party returns and begins to divide up the shares of meat among the members of the group, Dee approaches and demands a greater share based on the new law. The other community members laugh and tell her, "sure, we'll give you your extra share in the morning." Satisfied, Dee returns home to sleep and await her greater share of meat that will arrive in the morning. However, when Dee wakes up she finds that no extra meat has been delivered to her home. She rushes to the town square and demands her extra share. Laughing again, the townspeople ask Dee why she should get an extra share. Dee explains that she passed a law that requires the community to provide her with additional shares of meat. When asked to explain the basis for the law, Dee says that she gets hungry in the evenings and wants more food to satiate her hunger. The townspeople roar with laughter. "We are hungry in the evenings too, Dee. Why should only you get an extra share of meat?" Dee flushes with indignation and says that she cannot believe that the townspeople would question her desire to have more than her equal share of meat. As the days go by, Dee continues to live in judgment of her fellow townspeople. She feels that they have done something wrong by denying her an extra share of meat. Amused by Dee's behavior, the townspeople continue to live their lives by the rules and standards upon which they have all agreed.

The Requirement of Objectivity

The preceding example demonstrates the absurdity of trying to justify one's judgments from a purely subjective point of view. Of course, there is nothing absurd about Dee wanting more than her fair share. Most of us are guilty of similar desires. The absurdity lies in Dee's belief that her own personal desires should create a binding normative standard for the entire group. Dee's motive is entirely selfish. This is not to say that there is no justification that Dee could proffer and which would be acceptable to the whole group. If Dee can convince the group that she is uniquely suited to perform some task that will benefit the group as a whole, and that she can only accomplish that task if she is particularly well-nourished, she may be able to convince the group that providing her with an unequal share of rations is in the best interests of all.

The key difference between the two justifications that Dee gives for her behavior is that the former justification only makes sense *from Dee's perspective*. On the other hand, the latter justification makes sense *from the perspective of all*. This, according to Singer, is the essence of moral justification: "In a dispute between members of a cohesive group of reasoning beings, the demand for a reason is a demand for a justification that can be accepted by the group as a whole."³² Singer is not alone. Many other philosophers have made the same point and argued convincingly for it, in their own unique ways. And again, while we might dispute the details of each philosopher's view, most people would not disagree that moral views should be justified by objective standards.

The problem arises when we attempt to put this generally agreed upon tenet into practice. As I stated earlier in this chapter, the world is replete with examples of people

³² *Id.* at 93.

who are experts at applying objective moral justifications to a select group of people while completely ignoring the perspective of those outside the select group. We allow—and even passively support—conditions in other countries that we would be appalled to see happen in our own backyards. For example, we buy clothing manufactured by children forced to work in brutal conditions in third-world countries. At the same time, we rail against the relatively minor inconveniences suffered by our own children in the first world with a force of true moral outrage. We justify collateral damage in wars fought in the Middle East while we condemn, in the strongest moral terms, those who allow for collateral damage in their attacks on the United States. We give money to charities in our hometowns that only benefit a few, when that same money could be used to benefit thousands more in poor countries. And, as Singer argues, we torture and kill other animals—who are just as susceptible to pain as ourselves—to satisfy our taste buds even when we could sustain ourselves perfectly well with a vegetarian diet.³³

When push comes to shove, we are unable to justify these hypocrisies, blind spots, and selective attentions with rational argument. Rather, we are left to rely on irrational justifications based on our feelings. We argue that it is only natural to care more for our friends and families than we do for strangers. And we extend that reasoning to include members of our city, state, and country, and, more perniciously, to our gender, sex, race, ethnicity, religion, and species. This is the common trope of the rich who oppose taxation for government welfare programs. They argue that it should be up to individuals to decide how to allocate their charity. Indeed, they claim that allowing individuals to donate money to programs of their choice, rather than distributing that

³³ See SINGER, *supra* note 20.

money throughout the population via government programs, will lead to better services for those in need.

While there may be some merit in the argument that private organizations are better, and more efficient, at providing services than is the government, the argument suffers from a fatal flaw. Private actors almost always allocate their charity in a biased manner. People who care about veterans will donate money to veteran's organizations. People who care about hurricane victims will donate money to groups that help hurricane victims. Inevitably, some groups of needy people will have far more supporters than other groups of, equally, needy people. A particularly maligned minority may receive almost no support at all. Towns with many rich churchgoers will be more able to help their poor population than towns with lower median incomes. The distribution of charity would in many ways mirror the distribution of human empathy: altruism towards the near and the familiar; indifference towards the stranger.

Impartial Ethics

Let me return to my earlier claim that, for Singer, the solution to the problems of cruelty and unequal distribution of resources that plague humankind is the application of reasoned judgment. It would be unfair to claim that this is all Singer's argument amounts to. He claims that were we more rational beings "we would use our resources to save as many lives as possible . . . and we would be no readier to kill children from great heights than face to face."³⁴ But Singer acknowledges that because we are not wholly rational beings the best ethical outcomes may come through "an ethic that accepts our inclinations

³⁴ SINGER, *supra* note 8, at 157.

and harnesses them so that, taken as a whole, the system works to everyone's advantage."³⁵ At the same time, Singer is not satisfied with this position. Rather, he argues that human beings should rebel against the dictates of our genetic heritage and strive for a morality that is impartial. Indeed, Singer's hope is that we will adhere to a system of impartial ethics that takes genetics into account in the sense that one of its goals is to produce a future of more rational human beings.³⁶

To some, Singer's proposition will reek of eugenics. At the same time, it would be difficult to argue that a world where people acted for the good of all rather than for the good of some would not be a better world, all else being equal. While I do not disagree with Singer's goal, I do disagree that the best means of reaching that goal is through convincing people to act more rationally. In addition to the argument I have presented thus far—that evolution has created a dual nature in us that makes us perfectly capable of applying a rational moral code to some while ignoring others—philosophers such as David Hume have argued that human beings are motivated at least as much by emotion as we are by reason. I may accept the conclusion of a rational argument yet feel no pull to act on that conclusion.³⁷ Indeed, I am unable to argue against the conclusion that supporting a food system that tortures animals is immoral and unjust. At the same time, I feel little to no compulsion to change my behavior. Perhaps this is an example where “the spirit is willing but the flesh is weak.” In any case, rational argument alone is sometimes a

³⁵ *Id.*

³⁶ *Id.* at 172.

³⁷ See DAVID HUME, A TREATISE OF HUMAN NATURE 252 (Everyman 2003) (1738) (“[R]eason . . . can have an influence on our conduct only . . . when it excites a passion[.]”).

poor motivator.³⁸ The things that make me really question my decision to continue eating meat are the images shown by groups such as PETA, of hobbled cows, and chickens that are unable to stand in their battery cages; or the story of Francis the Pig, who escaped from a Canadian farm and wandered free in the wilderness for weeks before being recaptured.

The Role of Empathy in Impartial Ethics

The pull of these stories lies in the fact that when I see a tortured animal, or an animal that longs for freedom, I imagine myself in a similar position and can thus relate to the pain and the longing that those animals feel. This is the very basis upon which Singer rests his ethical case. As Singer states: “One way of arriving at . . . a decision . . . is to imagine myself living the lives of all affected by my decisions, and then ask what decision I prefer.”³⁹ Less sophisticated versions of this thought include, “walk a mile in my shoes,” and, “try to see the world through the eyes of another.” In each telling, the purpose remains the same: by taking on the perspective of others we become better able to act in a way that takes the interests of others into account. One interpretation of this scenario is that taking on the perspective of another is a way of transferring our own, natural, self-interest to another being.⁴⁰ This may appear as a cold way of looking at empathy. Nevertheless, the fact remains that so long as we are able to feel that we are

³⁸ However, as I will argue in more detail in Chapter Four, one is still morally responsible for one’s actions if one is *capable* of being motivated by some good reason.

³⁹ SINGER, *supra* note 8, at 101. In Chapter Four I argue that perspective taking involves a rational, or logical, step.

⁴⁰ I discuss this concept in more detail in Chapter Four.

separate from others we will be more willing and able to act in ways that harm other beings.

The rational argument that we should act in a way that benefits as many people as possible, in other words, that we should abstract away from our own perspective when making ethical decisions, is easier to accept when it is accompanied by *feelings* of empathy.⁴¹ I italicize “feelings” because one could potentially see the logic behind Singer’s argument, in other words, one could say to oneself, “yes, it is irrational for me to place my interests above others when I could just as easily be those others as I am myself,” while still missing out on the emotional *oomph* that accompanies what we typically mean when we talk about the feeling of empathy. What follows from this, as I argue in more detail in Chapter Four, is that it can be said that there are two components to empathy: an emotive component and a rational component. While Singer plays to our emotions at times in books like *Animal Liberation*, he is primarily focused on the latter kind of empathy—rational empathy.

Artificial Empathy Enhancement

In my view—and it is a view shared by Buchanan and several other philosophers—if Singer, and others like him, wishes to see the dream of an ever-expanding circle of empathy realized, he should be open to the possibility of artificially enhancing the human

⁴¹ See MICHAEL SLOTE, *THE ETHICS OF CARE AND EMPATHY* 15 (2007) (“[D]ifferences in strength or force of empathy make a difference to how much we care about the fate or others[.]”). For Slote, the natural feelings of empathy that accompany a direct experience of someone in distress, as opposed to an indirect experience, highlight a relevant moral distinction between our obligations to those nearby and our obligation to those in far-off places. Thus, contra Singer, Slote argues that we do *not* have the same more obligations to help starving children in African as we would to help a drowning child directly in front of us. See *id.* at 23-27. I discuss Slote’s view in more detail in Chapter Four.

empathic response.⁴² Singer does not address the possibility of human enhancement technologies as a means of achieving the goal of a more just and fair world, but the concept of artificially enhanced empathy is in line with Singer's view that the way to achieve the best possible outcome for all may be to "accept[] our inclinations and harness[] them" for the good of all.⁴³ After all, as I have argued, it is not as though most people reject Singer's logical argument.⁴⁴ Rather, people find it difficult to connect to the means and ends of that argument on an emotional level, in no small part due to the evolutionarily created boundaries of our capacity for empathy. Certainly, the human ability to reason gives us one tool with which to overcome the conditions of our existence laid down by biology—to, in Singer's words, "neutralize or reverse what might otherwise be genetically advantageous consequences of selfish behavior."⁴⁵ But it may be that reason itself is not enough. Human beings are not robots; we are driven by emotion as well as by reason. But this does not imply that all emotional impulses are selfish or bad. As I have argued, the emotional pull of empathy—in conjunction with logical

⁴² For my purposes, it's not necessary to assume that an increase in empathy would lead automatically lead to more utilitarian attitudes. An increase in empathic response would be beneficial under other ethical systems as well. For example, the feminist philosopher, Nel Noddings, argues for a distinction between what she calls "natural caring" and "ethical caring." NEL NODDINGS, *CARING: A FEMINIST APPROACH TO ETHICS AND MORAL EDUCATION* 81-83 (1984). Natural caring involves the feeling of *wanting* to help someone, while ethical caring involves the feeling that it is morally right to help someone. In a sense, this distinction mirrors my distinction between the emotive and rational components of empathy (which I develop in Chapter Four). Enhancing empathy would lead to an increase in natural caring that would compliment feelings of ethical caring. This point will become more clear in the following chapters.

⁴³ SINGER, *supra* note 8, at 157.

⁴⁴ *But see* SLOTE, *supra* note 41.

⁴⁵ SINGER, *supra* note 8, at 171.

argument—creates the impulse in us to act in ways that take others’ wellbeing into account.⁴⁶

A cursory survey of human history makes it fairly clear that we have generally expanded our sphere of ethical concern as time has progressed. Our biological tendencies towards altruism may have evolved in the limited context of small group dynamics. Yet we have learned to live together in massive societies where the evolutionary forces that originally drove the adoption of altruistic behaviors are no longer operative. There are several explanations for this phenomenon. The first is Singer’s argument that ethics, though brought into existence through biology, is justified by principles that are not founded on crude biology à la Wilson. Thus, we have been able to adapt our evolutionarily derived altruistic impulses to situations where pure biological survival is not the driving force. This apotheosis has an almost mystical quality. In Singer’s words, “[t]he capacity to reason is a special sort of capacity because it can lead us to places we did not expect to go.”⁴⁷

Second, we have used our capacity to reason about ethical issues to create systems of rules. In early societies these rules were generally enforced through informal social systems.⁴⁸ As societies expanded it became more and more necessary to codify these systems of rules in formal written documents. Formalized systems of rules have many interesting characteristics when it comes to the evolution of ethical tendencies in

⁴⁶ However, as I argue in Chapter Four, the rational component of empathy, but not the emotive component, is necessary for moral responsibility.

⁴⁷ SINGER, *supra* note 8, at 88.

⁴⁸ See MARY C. BRINTON & VICTOR NEE, *THE NEW INSTITUTIONALISM IN SOCIOLOGY* 92 (2001).

societies. In many ways, formalized rules take over the role of evolution in ensuring that ethical behaviors are maintained in a population. Where blind natural selection served to normalize altruistic tendencies in early human beings, the role of ethical curator is now filled by intelligently designed systems of rules. These rules allow for the expansion of ethical behaviors beyond small in-groups because they provide incentives and disincentives for behavior beyond those provided for by our individual, idiosyncratic, psychologies.

Despite all of this, we are constantly reminded of the evolutionary origin of our altruistic behaviors when we see the cruelty and indifference that human beings are able to display towards the Other. This brings us full circle to my earlier discussion of the indifference of evolution. Recall from Chapter One that one of the arguments put forward by opponents of human enhancement technologies is that those technologies would interfere with the “wisdom” of nature. I hope that I have argued convincingly in this chapter that nature is anything but wise. Rather, it blindly brings traits into existence with no concern for whether those traits conform with what we, as human beings, value. Reason may get us a long distance from the place on the map where evolution dropped us off. But even if reason will continue to lead us down the road towards greater care and concern for our fellow beings, it moves with a snails pace and often stops to rest. If the technology to increase our “natural” sense of empathy were available—and the technology were safe and effective—it would, at the very least, be something that we should consider implementing if we truly care about lessening the suffering of as many being as possible. At the most, if such technology were available we would have a moral duty to implement it as soon as possible.

Conclusion

In this chapter it has not been my goal to present a watertight argument for why it would be morally desirable, or morally required, for human beings to adopt human enhancement technologies that increased our sense of empathy if those technologies become available. Rather, I have only attempted to sketch a *prima facie* case for why the issue is worth talking about. Here, I concluded that artificial human enhancement of our natural empathic response is an important way in which the goals of our values may be realized. My purpose has been to provide groundwork for my discussion of the implications of artificial empathy enhancement in chapters five and six. In those two chapters, I will consider how people with enhanced empathy might be burdened by greater moral responsibility for their actions. Before getting to that discussion, however, it is necessary for me to consider the concept of empathy in more detail. In the next two chapters, I will use people on the opposite end of the empathic spectrum, namely, psychopaths, to evaluate empathy as a two-part concept involving a rational component and an emotive component.

Chapter Three

DIMINISHED EMPATHY

AN OVERVIEW OF PSYCHOPATHY

Introduction

When exploring areas of philosophical interest it is often helpful to think in terms of opposites. Right vs. wrong; conscious vs. unconscious; real vs. imagined; knowledge vs. belief. In some cases we have a more developed understanding of one side of the coin than the other. For example, it is generally easier to understand what it means for something to be unconscious than it is to understand just what constitutes consciousness. Other times, each side of the coin may be equally difficult to read yet without both sides we would have no context for the value of either. For example, we tend to think of morality in terms of right *and* wrong, rather than one or the other. And when we gain an insight into one it may in turn shed light on our understanding of the other. Our understanding that a lack of respect for the value of human life ought to be discouraged or punished informs our belief that an appreciation for the value of life ought to be encouraged and celebrated.

In chapters five and six of this dissertation I will be concerned with the ethical implications of artificially enhancing human empathic response beyond the boundaries provided by nature. While there is some precedent for understanding what it would mean for our empathic response to extend beyond the baseline—consider people like Martin Luther King Jr. or the Buddha—, by definition there are no examples of people whose empathic response reaches past the bounds of our evolutionarily created brains. In other words, there are no super empaths.

However, there is a whole population of people who live on the other side of the coin. Psychopaths are people who psychologists regard as having a profoundly decreased sense of empathy. Like a child born with no sense of smell, psychopaths recognize that other people appear to have some experience of the world that the psychopaths are missing out on.

In this chapter and the next, I will focus on how we relate to psychopaths from legal and ethical perspectives. Here, I will provide a brief history of psychopathy as well as a survey of how the legal system treats psychopaths—since the legal system provides at least a rough approximation of the majority’s ethical views. In Chapter Four, I will consider how certain philosophers have evaluated psychopaths, and I will present my own argument regarding the psychopath’s moral and legal responsibility.

A Brief History of Psychopathy

When most people think of a psychopath the image that comes to mind is of a deranged and cold-blooded person who acts purely for motives of self-interest with no regard for the feelings, or the lives, of others. If I asked a group of people to name a psychopath the most frequent response would undoubtedly be “Ted Bundy.” While it is generally acknowledged that psychopathy is associated with criminality—sometimes of the most brutal sort—it is not necessarily the case that every psychopath is a sadistic murderer. Several studies have suggested that psychopaths are prominent—at least compared to the general population—in business, as CEOs and entrepreneurs.¹ According to a study by P.

¹ P. Babiak, C.S. Neuman, & RD Hare, *Corporate Psychopathy: Talking the Walk*, 28 BEHAV. SCI. LAW 174 (2010); see Julian Lorkin, *Why a Psychopathic Tendency May be Good for Entrepreneurs*, UNIV. N.S.W. AUSTL. BUS. SCHOOL (Aug. 28, 2013),

Babiak, CS Neuman, and Robert Hare, psychopaths account for 4% of corporate CEOs—four times the prevalence of the general population.² This number may not be high enough to justify the fear propaganda spread by books like *The Sociopath Next Door*,³ but it is significant when we consider the characteristics attributed to psychopaths.

Psychopathy is a controversial diagnosis—if it can even be called that. Neither the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) nor the *International Classification of Diseases* (ICD) recognizes a diagnosis of psychopathy.⁴ While the term “psychopath” is of relatively recent origin, one can look to the distant past to find examples of historians, writers, psychiatrists, and psychologists, who describe people who embody the traits we now associate with psychopathy. One of Aristotle’s students, Theophrastus, wrote a book called *The Unscrupulous Man*, where he described behavior that resembles traits now associated with psychopathy. Theophrastus described the

<https://www.business.unsw.edu.au/news-events/news/why-psychopathic-tendency-good-for-entrepreneurs>.

² *Id.*

³ MARTHA STOUT, *THE SOCIOPATH NEXT DOOR* (2006). The first thing to note here is that psychopathy and sociopathy are, for the most part, synonymous terms. The difference lies in the theoretical etiology of the behaviors manifested. See ROBERT D. HARE, *WITHOUT CONSCIENCE: THE DISTURBING WORLD OF THE PSYCHOPATHS AMONG US* 23-25 (1999). In any case, there is no diagnostic test for sociopathy as distinct from psychopathy. The second note is that books like Stout’s are little more than “fodder for our worst—not to mention most implausible—suspicions.” Pamela Paul, ‘*The Sociopath Next Door*’: *Ruthless People*, *NY TIMES: SUNDAY BOOK REV.* (Mar. 6, 2005), http://www.nytimes.com/2005/03/06/books/review/the-sociopath-next-door-ruthless-people.html?_r=0. Unfortunately, there is an entire cottage industry of pop psychology that plays on the “exciting” and “alien” nature of the psychopath while exaggerating and misstating the science behind the disorder.

⁴ As will be discussed below, both diagnostic manuals recognize the related diagnosis of anti-social personality disorder.

unscrupulous man as one who, “[w]hen marketing . . . reminds the butcher of some service he has rendered him and, standing near the scales, throws in some meat, if he can, and a soup-bone. If he succeeds, so much the better; if not, he will snatch a piece of tripe and go off laughing.”⁵ This kind of “moral madness” was also described by the French psychiatrist Philippe Pinel in 1801. In his work trying to bring about reform in the abysmal asylums that were scattered throughout Europe in the nineteenth century, Pinel was struck by a particular kind of inmate—one who appeared to be mentally well from a traditional psychiatric perspective, yet was prone to impulsive, immoral, and self-defeating behavior. He called these behaviors *manie sans delire*—mania without delusion.⁶ In 1835, building on Pinet’s observations, James Cowles Prichard—an English psychiatrist—invented a new mental disorder called “moral insanity.”⁷

Although the term “psychopath” had been around since at least the early 1800s, it was not until 1888 that it was used in a context familiar to us today. In formulating his own version of moral insanity, Julius Ludwig August Koch described behavior that he referred to as exhibiting a “psychopathic disposition.”⁸ Koch’s description of the

⁵ Quoted in T. Millon, E. Simonsen & M. Birket-Smith, *Historical Conceptions of Psychopathy in the United States and Europe*, in *PSYCHOPATHY: ANTISOCIAL, CRIMINAL, AND VIOLENT BEHAVIOR* 3 (Theodore Millon et al., eds., 2002). A question that arises here is whether the psychopath’s behavior should be described as “evil.” While modern day explanations of behavior have moved away from pure moral ascriptions like good and evil, there is at least an intuitive sense that a person who commits wrongdoing with no compunction and no remorse is especially worthy of moral condemnation.

⁶ *Id.* at 5.

⁷ See JAMES COWLES PRICHARD, *A TREATISE ON INSANITY AND OTHER DISORDERS AFFECTING THE MIND* (1837).

⁸ See RICHARD F. WETZELL, *INVENTION OF THE CRIMINAL: A HISTORY OF GERMAN CRIMINOLOGY, 1880-1945* 48 (2000).

psychopath is eerily poetic. He describes the psychopath as exhibiting “anomalies in excitability, a lack of harmony, an eccentric, contradictory self, peculiarities, primordial instinctive impulses and outbursts and something periodic in behavior.”⁹ The modern-day character of Patrick Bateman, from Bret Easton Ellis’ remarkably disturbing 1991 novel, *American Psycho*,¹⁰ is the near-perfect personification of Koch’s words.

In 1918, the *Statistical Manual for the Use of Institutions for the Insane* was published.¹¹ The manual, which was a precursor to the modern-day *Diagnostic and Statistical Manual of Mental Disorders* (DSM), listed “constitutional psychopathic inferiority” as a diagnosis that included behaviors such as “criminal traits, moral deficiency, tramp life, sexual perversions and various temperamental peculiarities.”¹² In 1930, the American psychologist George E. Partridge published a paper, *Current Conceptions of Psychopathic Personality*.¹³ In the paper, Partridge argued for a narrower definition of psychopathy that included antisocial elements at its core.¹⁴

⁹ *Quoted in id.*

¹⁰ BRET EASTON ELLIS, *AMERICAN PSYCHO* (1991).

¹¹ AMERICAN PSYCHIATRIC ASSOCIATION, *STATISTICAL MANUAL FOR THE USE OF INSTITUTIONS FOR THE INSANE* (U. Mich. 2008) (1918).

¹² *Id.* at 28.

¹³ George E. Partridge, *Current Conceptions of Psychopathic Personality*, 87 AM. J. PSYCHOL. 53 (1930).

¹⁴ *See id.* Partridge also introduced the term “sociopath,” which he considered a better term for a disorder characterized by antisocial tendencies. While there is some disagreement today over whether the two terms—psychopathy and sociopathy—refer to the same group of people, the general consensus is that they are interchangeable. For my purposes, I will treat them as such.

Psychopathy in the Mainstream

The renaissance of psychopathy research—and the beginning of the popularization of the term in the wider culture—began with the publication of American psychiatrist Hervey M. Cleckley’s 1941 book, *The Mask of Sanity*.¹⁵ Cleckley surveyed examples of psychopathic personalities that he had encountered in his practice and concluded that psychopaths share many traits in common, including superficial charm, lack of remorse or shame, antisocial behavior, incapacity for love, and an unstructured life plan.¹⁶ Cleckley famously described the psychopath as a “subtly constructed reflex machine.”¹⁷ He vividly described the alien nature of the psychopath as follows:

[The psychopath] can mimic the human personality perfectly. This smoothly operating psychic apparatus reproduces consistently not only specimens of good human reasoning but also appropriate simulations of normal human emotion in response to nearly all the varied stimuli of life. So perfect is this reproduction of a whole and normal man that no one who examines him in a clinical setting can point out in scientific or objective terms why, or how, he is not real. And yet we eventually come to know or feel we know that reality, in the sense of full, healthy experiencing of life, is not here.¹⁸

In many ways, Cleckley’s description of the psychopath harkens to David Chalmers’ zombie thought experiment.¹⁹ Like Chalmers’ zombies, Cleckley’s psychopaths look and

¹⁵ HERVEY M. CLECKLEY, *THE MASK OF SANITY* (5th ed. 1982).

¹⁶ *Id.* at 369.

¹⁷ *Id.* at 338-39.

¹⁸ *Id.* at 369-70.

¹⁹ See generally DAVID J. CHALMERS, *THE CONSCIOUS MIND: IN SEARCH OF A FUNDAMENTAL THEORY* 94-99 (1996).

act like normal human beings, but they lack conscious experience—and least in terms of empathy—of what it *feels* like to be a caring, empathic human being.²⁰

Cleckley's work was influential in defining psychopathy outside the context of criminal behavior. He presented several case studies of psychopaths who had achieved material success and had risen to the highest positions in society. According to Cleckley, there were psychopaths in business, science, medicine, and even in psychiatry.²¹ In his studies, Cleckley found that psychopaths were often highly intelligent and, if they were able to maintain a consistent outward appearance of normality, able to achieve positions of authority. Indeed, Cleckley observed that many psychopaths were able to use their pathology to gain an advantage over other people who were weighed down by empathy and ethics.²²

In 1952, eleven years after Cleckley published his influential book, the first edition of the DSM came out and listed “sociopathic personality disturbance” as a diagnosis.²³ The four subtypes described a full range of what we now associate with the psychopathic personality: an inability to change in response to punishment, disregard for

²⁰ The analogy is not perfect, of course. Chalmers' zombies lack *all* experience, not just experience of certain aspects of what it means to be human. But the concept of the psychopath as a detached entity—acting without thought, conscience, and even sometimes without consciousness—is pervasive.

²¹ See CLECKLEY, *supra* note 15, at 21-26.

²² See *id.*

²³ AMERICAN PSYCHIATRIC ASSOCIATION, DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 38 (1952).

societal rules, sexual deviation, and propensity for addiction to drugs and alcohol.²⁴ However, from the DSM-II on, the terms “sociopath” and “psychopath” were dropped and replaced with “antisocial personality disorder” (APD).²⁵ The DSM V—the most recent edition of the manual—provides the following description of APD:

A pervasive pattern of disregard for and violation of the rights of others, occurring since age 15 years, as indicated by three (or more) of the following:

1. Failure to conform to social norms with respect to lawful behaviors, as indicated by repeatedly performing acts that are grounds for arrest.
2. Deceitfulness, as indicated by repeated lying, use of aliases, or conning others for personal profit or pleasure.
3. Impulsivity or failure to plan ahead.
4. Irritability and aggressiveness, as indicated by repeated physical fights or assaults.
5. Reckless disregard for safety of self or others.
6. Consistent irresponsibility, as indicated by repeated failure to sustain consistent work behavior or honor financial obligations.
7. Lack of remorse, as indicated by being indifferent to or rationalizing having hurt, mistreated, or stolen from another.²⁶

Despite the change in terminology, the impact of Cleckley’s description of the psychopath is clearly evident in the modern definition of APD. Indeed, in the section titled “Associated Features Supporting Diagnosis,” the DSM-V repeats many of the criteria coined by Cleckley. It talks about individuals with APD as projecting “a glib,

²⁴ Frederick L. Coolidge & Daniel L. Segal, *Evolution of Personality Diagnosis in the Diagnostic and Statistical Manual of Mental Disorders*, 18 CLINICAL PSYCHOL. REV. 585, 589 (1998).

²⁵ See, e.g., AMERICAN PSYCHIATRIC ASSOCIATION, DSM-II: DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 301 (1968).

²⁶ AMERICAN PSYCHIATRIC ASSOCIATION, DSM V: DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS 659 (2013).

superficial charm,” and displaying a total “lack [of] empathy . . . [for the] suffering of others.”²⁷

I spend so much time on the history and diagnosis of psychopathy (and its cousin, APD) because I wish to make it clear that these people do, in fact, exist; that the idea of a being, human in biology, appearance, and general demeanor, but utterly lacking in the qualities that we most associate with humanity, is not a myth dreamed up by hack authors of fiction or lazy psychologists who could find no other way to classify this group and so labeled them “psychopaths” as a catch-all for those human creatures that cause things to go bump in the night. Unfortunately, psychopaths are all too real. It is surely the case that certain opportunistic psychologists and journalists have played up the percentage of psychopaths in the general population and the threat of these individuals to the rest of us, blessedly normal, human beings. Nevertheless, the concept of psychopathy is well defined and has been shown to be a unique psychological construct through extensive psychometric testing.

Robert Hare and Lack of Empathy as a Hallmark Trait of Psychopathy

While Cleckley may be responsible for creating the modern idea of the psychopath, Robert D. Hare operated the bellows that breathed oxygen into the psychopathic personality as the modern day monster. In his 1993 book, *Without Conscience*, Hare described the existence of the psychopath as “a dark mystery with staggering implications for society.”²⁸ Among the traits Hare ascribes to psychopaths is a generalized lack of

²⁷ *Id.* at 660.

²⁸ HARE, *supra* note 3, at 2.

empathy. Unlike doctors or soldiers, who learn to repress their natural feelings of empathy in order to get the job done, Hare describes psychopaths as being completely unable to experience feelings of empathy.

Interestingly, another trait that Hare notes as being common in psychopaths is an inability to understand complex syntax or to appreciate the emotional content that is so frequently embedded in language.²⁹ Again, this should not come as a surprise. The emotional impact of certain phrases and expressions is often a result of the listener's ability to put herself in the shoes of another person. A world where human beings did not empathize with one another would surely be a world with an entirely different language than the emotionally laden languages that make up our world. Thus, a psychopath is unable to understand the “feel” meaning of language for the same reasons she is unable to display more than a shallow emotional affect: her world simply does not include most of the emotional experiences described by our language.

The PCL-R

In 1996, Hare wrote an article that emphasized the psychopath's inability to feel empathy—and the dire societal consequences of failing to recognize psychopaths as a unique brand of criminal that requires special attention from the criminal justice system.³⁰

In the article, Hare presented research demonstrating that psychopaths make up 25% of the prison population, despite being only 1% of the population as a whole.³¹ Hare's goal

²⁹ *Id.* at 130.

³⁰ See Robert D. Hare, *Psychopathy: A Clinical Construct Whose Time Has Come*, 23 CRIM. JUST. BEHAV. 25 (1996).

³¹ *Id.* at 26.

was to dispel the idea, popular among some psychologists, that psychopathy is a mythological disorder used to explain psychological and behavioral phenomenon for which we have no other good explanation. Thus, Hare presented his brainchild, the *Psychopathy Checklist* (later revised and now referred to as the PCL-R).

Hare had created the PCL-R in the late 80s and early 90s and put it through rigorous testing in the years before the publication of his 1996 article.³² The test consists of nineteen questions related to traits associated with coldness and lack of empathy. When Hare created the test, he modeled the categories after Cleckley's prototypical psychopath.³³ A clinician interviews a subject and rates him on a scale between 0-50, with a score of 30 being the cutoff for a diagnosis of psychopathy.³⁴ Scores on the test fall along a spectrum, so it is not that a person who scores a 29 is not a psychopath while a person who scores a 30 is a psychopath. However, the higher one's score, the more psychopathic traits one can be said to have. According to Hare, the test shows remarkable consistency and reliability³⁵—which demonstrates not only that psychopathy is a stable psychological construct but also that Cleckley, whose model of psychopathy was used to create the PCL-R, was an insightful clinician. Interestingly, Hare found that psychopathy

³² See Robert D. Hare et al., *The Revised Psychopathy Checklist: Reliability and Factor Structure*, 2 PSYCHOL. ASSESSMENT 338 (1990).

³³ See generally Robert D. Hare, *A Research Scale for the Assessment of Psychopathy in Criminal Populations*, 1 PERS. & INDIVIDUAL DIFFERENCES 111 (1980).

³⁴ ROBERT D. HARE, *THE HARE PSYCHOPATHY CHECKLIST-REVISED* (1991).

³⁵ Hare, *supra* note 32, at 340.

only loosely correlates with the DSM diagnosis of APD. According to Hare's research, there are less than half as many psychopaths as there are people with APD.³⁶

Hare's primary goal in the 1996 article was to convince more jurisdictions to adopt the PCL-R as a tool for determining the suitability of inmates for parole. In that vein, Hare described several studies that he and other psychologists had performed demonstrating that the higher an inmate scores on the PCL-R, the more likely that inmate is to reoffend.³⁷ Hare was alarmed with what he had witnessed during some of his studies. For example, he saw innumerable cases where psychopaths had been assessed by prison psychologists and been described as having made "remarkable progress" in the treatment of their disorder. Judges and parole boards would hear this evidence, as well as the heartfelt testimony of the psychopath, conclude that the inmate was reformed, and allow her back into the world. Yet Hare and others found that psychopaths derive almost no benefit from therapy. Indeed, Hare reported one study—since repeated³⁸—showing that therapy can actually *increase* the recidivism rate for psychopaths. Hare explained the reason for this counterintuitive result: "group therapy and insight-oriented programs help

³⁶ Hare, *supra* note 30, at 31. While both men and women can be psychopaths, about twice as many men score over 30 on the PCL-R. Data from large sets of male and females inmates show that 15.7% of men and 7.4% of women score over 30. *See* Tonia L. Nicholls et al., *Psychopathy in Women: A Review of its Clinical Usefulness for Assessing Risk for Aggression and Criminality*, 23 BEHAVIORAL SCIENCES AND THE LAW 779, 785 (2005).

³⁷ Hare, *supra* note 30, at 39-41.

³⁸ *See* Michael C. Seto & Howard E. Barbaree, *Psychopathy, Treatment Behavior, and Sex Offender Recidivism*, 14 J. INTERPERSONAL VIOLENCE 1235 (1999).

psychopaths to develop better ways of manipulating, deceiving, and using people but do little to help them understand themselves.”³⁹

Hare’s test is not without its critics. First, there are the critics who argue that scores on the PCL-R are given too much weight by parole boards considering the dearth of evidence demonstrating any correlation between higher scores and recidivism.⁴⁰ Second, there are critics who argue that the PCL-R does not actually measure psychopathy as the term is understood by the wider psychological community. Rather, these critics argue, the PCL-R puts too much emphasis on criminal behavior as a telltale sign of psychopathic tendencies.⁴¹ Indeed, a person *must* display antisocial behavior to be qualified as a psychopath under the PCL-R.⁴² While a criminal history can sometimes be viewed as a proxy for anti-social behavior, in a country such as the United States where the vast majority of inmates are incarcerated for victimless crimes,⁴³ there is something perverse about equating behavior that is statutorily defined as criminal with the psychological diagnostic criteria of displaying anti-social tendencies.

³⁹ Hare, *supra* note 30, at 42.

⁴⁰ See, e.g., Daniel C. Murrie et al., *Field Validity of the Psychopathy Checklist-Revised in Sex Offender Risk Assessment*, 24 PSYCHOL. ASSESSMENT 524 (2012) (finding that scores on the PCL-R are not predictive of sexually violent recidivism).

⁴¹ See Glenn D. Walters, *The Trouble with Psychopathy as a General Theory of Crime*, 48 INT. J. OFFENDER THERAPY COMP. CRIM. 133 (2004); Jennifer L. Skeem & David J. Cooke, *Is Criminal Behavior a Central Component of Psychopathy? Conceptual Directions for Resolving the Debate*, 22 PSYCHOL. ASSESSMENT 433 (2010).

⁴² See Skeem & Cooke, *supra* note 41, at 433.

⁴³ E. ANN CARSON & WILLIAM J. SABOL, BUREAU JUST. STAT., PRISONERS IN 2011 1 (2012), <http://www.bjs.gov/content/pub/pdf/p11.pdf> (83% of federal prisoners in 2011 were incarcerated for victimless crimes: 48% for drug offenses and 35% for public-order crimes).

Nonetheless, Hare predicted that the PCL-R would increasingly be used in determining an inmate's eligibility for parole, and would also be used to civilly commit psychopathic individuals.⁴⁴ He was right. In addition to the PCL-R's extensive use by parole boards, the test has been used as a diagnostic tool to justify the indefinite civil commitment of sexual psychopaths.⁴⁵ For Hare, this surely comes as good news. According to his research, psychopaths are far more likely to reoffend than their non-psychopathic brethren.⁴⁶ And almost by definition, psychopaths are incorrigible. If they are programmed through genetics to be incapable of feeling empathy, there seems to be little hope for rehabilitation other than some kind of draconian program that subjects the psychopath to an electric shock when her behavior starts to become "antisocial."⁴⁷

These outcomes, predicted by Hare, indicate what might appear to be an obvious observation at first glance: psychopathy is considered an aggravating factor by courts and by the public. After all, the traits that define psychopathy are hardly flattering: callous lack of empathy, parasitic lifestyle, pathological lying, grandiose sense of self worth, etc.—these words and phrases do not ring a sympathetic tone. Nevertheless, there is

⁴⁴ Hare developed a separate test for use in the civil context: the PCL-SV ("short version."). See Hare, *supra* note 30, at 32.

⁴⁵ Records from appellate courts throughout the country indicate that from 2005 to 2011, the PCL-R was used to civilly commit 214 sexually violent predators (SVP). David DeMatteo et al., *The Role and Reliability of the Psychopathy Checklist—Revised in U.S. Sexually Violent Predator Evaluations: A Case Law Survey*, 38 LAW & HUM. BEHAV. 248, 250 (2013). The actual number will be much higher since only a small percentage of cases get appealed.

⁴⁶ See Hare, *supra* note 30, at 39-43.

⁴⁷ Something akin to Alex's plight in *A Clockwork Orange*. See generally ANTHONY BURGESS, *A CLOCKWORK ORANGE* (1962).

research indicating that our attitudes towards psychopaths are partly shaped by what we consider to be the origin of the psychopath's pathology.

Psychopathy as a Mitigating or Aggravating Factor?

In 2012, the University of Utah conducted a study that asked 200 judges in nineteen states whether they would consider evidence of psychopathy in a criminal defendant as an aggravating or a mitigating factor.⁴⁸ When presented on its own—through descriptions of a defendant's behavior and state of mind—the judges tended to view psychopathy as an aggravating factor. Presumably, the judges were ascribing a certain degree of agency to the psychopath's behaviors and states of mind. In other words, the judges likely viewed the psychopaths as responsible for their own predicaments. It is easy to see why the judges would come to this conclusion. In many ways, psychopathy does not seem like a mental disorder, at least in the way that we usually think of that term. Psychopaths do not hear voices directing them to commit criminal acts. They are not plagued by delusions that make it impossible for them to differentiate fantasy from reality. They are not foaming at the mouth or rambling to themselves incoherently on the street. Perhaps most importantly, other mental disorders generally have descriptions that are morally benign. Indeed, we tend to describe mental disorders using sympathetic language. A schizophrenic *suffers* from paranoia and hallucinations. A person with bi-polar disorder is *subject* to uncontrollable fluctuations in mood. Not so with psychopathy.

⁴⁸ Lisa G. Aspinwall, Teneille R. Brown, & James Tabery, *The Double-Edged Sword: Does Biomechanism Increase or Decrease Judges' Sentencing of Psychopaths?*, 337 SCIENCE 846 (2012).

The very definition of psychopathy is imbued with morally significant language. Psychopaths are *callous*. Psychopaths are *devoid of empathy*. Psychopaths are *selfish*. All of these words imply agency. One can act callously. One can fail to display empathy. One can be selfish. Other mental disorders tend to lack this implied agency in their descriptions. Again, consider schizophrenia and bi-polar disorder. It requires much more work to put the symptoms of those two disorders in terms of agency. Saying that the schizophrenic hallucinates does not imply agency. Similarly, no agency is implied in saying that the person with bi-polar disorder has rapidly fluctuating moods. It seems that the only way to turn those descriptions into agency-laden terms is to add the word “choose” before describing the trait. But that results in phrases that seem quite strange to our ears. For example, saying that a schizophrenic *chooses* to hallucinate sounds plainly wrong. To our ears, it does not seem fair to ascribe agency to the traits of schizophrenia because this implies a sort of victim-blaming—that schizophrenics are responsible for their own suffering. On the other hand, we are comfortable ascribing agency to the psychopath because we feel that traits like callousness, lack of empathy, and selfishness *are* subject to one’s will. Thus, we feel little hesitation in thinking that a selfish, callous, un-empathic person has a bad character and is morally condemnable.

Interestingly, this initial reaction was disrupted in the judges from the Utah study when they were provided with a biological basis for a defendant’s psychopathy. There is plenty of evidence suggesting that psychopathy is the result of defects in the brain. Several studies have highlighted the amygdala as an area of the brain that is defective in

psychopaths.⁴⁹ The amygdala is at least partially responsible for processing emotional experiences. Specifically, the amygdala is the part of the brain that allows us to recognize the emotional content of things we see and hear in the world.⁵⁰ It should come as no surprise, then, that a defect in the amygdala would lead a person to lack empathy towards others. At least part of our experience of empathy involves a purely emotional connection with other people. When I see a person in pain, I have an immediate sensation of pain—usually in the form of sadness—that is entirely reactive and emotional in nature. In other words, I do not have to stop and think to myself: “this person is in pain and it could just as easily be me that is in pain and I know how terrible it is to be in pain, therefore I empathize with this person and feel her pain.” This kind of logical thought is surely a component of empathy (as I describe later in the following chapter) but it is not the only part—and it is certainly not the most immediate part.

If there is a biological basis for psychopathy—as the researchers in the Utah study suggested to one group of judges—our initial reactions to the psychopath come into question, as is presumably what happened with the judges who were provided with such a biological basis. The words that describe a psychopath—the words so imbued with a sense of agency—suddenly take on a sympathetic quality. Callousness might become “an inability to feel”; lack of empathy might become “alienation from human experience.” This reflects our intuitive feeling that a person should not be held responsible—or, at least, should be held to a lesser degree of responsibility—when her acts cannot be

⁴⁹ See, e.g., R.J.R. Blair, *The Amygdala and Ventromedial Prefrontal Cortex: Functional Contributions and Dysfunction in Psychopathy*, 363 PHIL. TRANSACTIONS ROYAL SOC’Y. LONDON B 2557 (2008).

⁵⁰ *Id.* at 2558.

considered as the result of her will but rather as a result of some biological brain abnormality. Thus, in the same way that we are less inclined to ascribe blameworthiness to a schizophrenic who is suffering from delusions when she commits a crime, we are less likely to morally condemn a psychopath who lacks control over her pathology.

Case Study: Psychopathy and Volition

One case that reflects this intuition has been much discussed in the psychological literature. Alex was a normal, well-adjusted, forty-year-old man who suddenly developed an intense interest in child pornography.⁵¹ He had never before displayed any signs of aberrant behavior and was, in fact, appalled at his sudden compulsions. In this sense, Alex was not like a psychopath because he did feel guilt and remorse for his actions. Nonetheless, his case is instructive in that it demonstrates how our attitudes towards what we normally consider to be volitional behavior can change when we discover that the behavior is the result of a brain abnormality. As time went on, Alex accumulated more and more child pornography. Eventually, his prurient interests turned towards his prepubescent stepdaughter. Alex's wife noticed that he was "making subtle sexual advances" towards their daughter and she was, understandably, horrified. She searched the house and discovered Alex's stash of child pornography and promptly turned Alex into the police. Alex was removed from his home, charged with child molestation, and diagnosed with pedophilia before being placed in a 12-step recovery program. However, Alex was unable to control his sexual appetites and repeatedly made advances to the

⁵¹ See Jeffrey M. Burns & Russell H. Swerdlow, *Right Orbitofrontal Tumor with Pedophilia Symptom and Constructional Apraxia Sign*, 60 ARCHIVES NEUROLOGY 437, 437 (2003).

staff. The judge in his case determined that Alex was incorrigible and was determined to throw him in prison.

The night before sentencing was to be carried out, Alex was checked into the University of Virginia Hospital after complaining of a severe headache. An MRI revealed a tumor impinging on Alex's amygdala and orbitofrontal cortex. Both brain areas are associated with moral knowledge acquisition, and damage to either area can result in psychopathic behavior.⁵² Again, it is not clear that Alex could rightly be described as a psychopath, at least in the full Cleckley / Hare sense of the word. Yet, there is surely something psychopathic about his behavior. Alex knew that what he was doing was wrong, yet he did not stop himself from doing it. Indeed, even after being removed from his home and sent to a mental institution, Alex did not stop from engaging in sexually inappropriate behavior. Recall that one of the defining traits of psychopathy is a failure to respond to punishment.

Yet, in a sense, Alex was one of the lucky ones. After doctors removed the tumor from his brain Alex's pedophilic desires disappeared as if by magic. Alex, who the day before had met every criteria in the DSM-IV for a pedophile, was once again his normal, well-functioning self. It is not clear exactly how Alex managed to avoid a jail sentence—this part of the story is missing from the article published by the doctors who worked on the tumor—but within seven months Alex was determined to be entirely “cured” and he returned home to his wife and daughter. All was well until he once again began to experience intense headaches and was once again possessed by the urge to start viewing and collecting child pornography. This time, Alex went straight to the hospital. His

⁵² *Id.* at 438.

doctors saw that the tumor in his brain had regrown and once again excised it. Again, like magic, Alex was his old self.

Alex's case is interesting because the behaviors that he engaged in were all volitional. He was not forced to look at child pornography—he certainly was not forced to hoard a large collection. Alex had control over his limbs. He could have chosen not to click on the links. He could have chosen to dispose of his child pornography collection at any time. Unlike a schizophrenic, Alex was not being directed by voices to engage in immoral acts. Neither was he under a delusion that what he was doing was morally acceptable. Rather, he was aware of the immorality of his behavior—he even believed that it was immoral, in other words, he felt guilt and remorse—yet he continued to engage in the behavior. All of these things cry out for judgment. Just as our instinctive reaction to the psychopath is to hold her responsible for her immoral acts, our initial reaction to Alex is to judge him as a bad person. Yet, this moral condemnation is overcome the moment I hear that Alex's compulsions were likely caused by the tumor in his brain. Suddenly, Alex's movements seem lacking in volition.⁵³ Even the complex behaviors of searching out, obtaining, and hoarding large amounts of child pornography, start to feel like an automated script that is being carried out *through* Alex rather than *by* him.

I think that the same analysis can apply to how the judges in the Utah study changed their attitudes—and, in turn, their sentencing decisions—upon being told that

⁵³ Of course, while Alex may not have been able to resist his impulses, he was able to put himself in a position where it would be impossible for him to act on those impulses. Indeed, when the impulses recurred after his first brain surgery Alex went right back to the doctor because, presumably, he knew that was the only way to stop himself.

psychopathy has a biological basis. This information turned the diagnosis of psychopathy from one of aggravation to one of mitigation. The judges suddenly became sympathetic to the psychopath's attitudes and behaviors, or at least felt that those attitudes and behaviors were not rightly ascribable to a *choice* made by the individual (now) suffering from psychopathy. Just as our attitudes towards Alex soften when we learn that a tumor is the cause of his aberrant behavior, the judges in the Utah study softened towards psychopathy when it was given grounding in biology.

Again, this result should not be particularly surprising, because it reflects our intuitive view that a person should not be held responsible—or should at least be held to a lesser degree of responsibility—when her acts cannot be considered a result of her will but rather as a result of some biological brain abnormality. Thus, if someone is suffering from a schizophrenic delusion when she commits a crime, we are less likely to hold her blameworthy. On the other hand, if she tries to offer the excuse of something like, “well, I did it because I don’t care about other people’s suffering,” we are likely to feel that she is blameworthy. Yet, if experts identify some specific brain defect that made it so that she was *incapable* of caring about other people’s suffering, we might be willing to say—as the judges in the Utah study did—that such evidence should weigh in favor of mitigating her sentence.

Indeed, there are real-world examples of cases in the United States where juries found evidence of psychopathy to be a mitigating factor, and thus reduced the sentences

of criminal defendants.⁵⁴ Specifically, these juries were convinced that the psychopaths' inability to feel empathy at least partially absolved them of their crimes. The defense attorneys in these cases argued that their clients' inability (or decreased ability) to feel empathy meant that they should not be judged for failing to take moral reasons into account when they acted to cause the harm for which they stood accused. In some sense, the attorneys were attempting to convince the jury that their clients' behaviors—fueled by their biologically-imbued psychopathic natures—were more akin to reflex actions than volitional decisions to act. The case of Alex is again illustrative here. If we assume that the tumor impinging on Alex's amygdala and pre-frontal cortex was, in fact, the cause of his sexual outbursts—in the sense that it reduced his impulse control to a point where he could not resist his sexually-deviant urges—it seems as though Alex bears less moral or legal responsibility for his acts.⁵⁵ As the argument goes, the tumor acted upon Alex in the same way that a condition such as epilepsy might act upon a person, causing that person's arms to flail out and hit someone in the face as a result of a seizure. Certainly, no juror would find the epileptic guilty of a crime or in violation of some moral dictate.

The Model Penal Code: Mental States and Volitional Action

Insofar as the analogy to a reflex action holds, one could make the argument that Alex—or the psychopath acting under the controlling influence of a biological abnormality—should not be held legally responsible for her acts to *any* degree. This follows from the

⁵⁴ See Kate Kelland, *Biology Gives American Psychopaths a Legal Break*, REUTERS (Aug. 16, 2012), <http://www.reuters.com/article/2012/08/16/us-science-psychopaths-biology-idUSBRE87F12M20120816>.

⁵⁵ Although the legal system might be justified in civilly confining him to prevent him from harming others.

notions of criminal responsibility that developed in the United States through the common law and were eventually codified in the *Model Penal Code* (MPC) and subsequently through state and federal statutes that adopted large portions of the MPC.⁵⁶ Under the MPC, the requirements for criminal responsibility can be broken down into two broad categories: *actus reus* and *mens rea*.⁵⁷ The *actus reus* is the act—or, in certain circumstance, the failure to act—that leads to the harm for which one may, at least preliminarily, be held criminally responsible.⁵⁸ In the case of murder, for example, the *actus reus* might consist of pulling the trigger on a gun.

Importantly, in order for a person to be held criminally responsible for her conduct, the *actus reus* must be volitional. In other words, one cannot be held accountable for acts that are outside of one's control. Specifically, Section 2.01 of the MPC exempts the following actions from criminal responsibility:⁵⁹ (a) a reflex or convulsion; (b) a bodily movement during unconsciousness or sleep; (c) conduct during hypnosis or resulting from hypnotic suggestion; (d) a bodily movement that otherwise is not a product of the effort or determination of the actor, either conscious or habitual.

⁵⁶ See JOSHUA DRESSLER & STEPHEN P. GARVEY, *CASES AND MATERIALS ON CRIMINAL LAW* 5 (6th ed. 2012).

⁵⁷ See *id.* at 127. *Actus reus* is defined in MODEL PENAL CODE § 2.01 (1962) and *mens rea* is defined in *id.* § 2.02.

⁵⁸ Albin Eser, *The Principle of “Harm” in the Concept of Crime: A Comparative Analysis of the Criminally Protected Legal Interests*, 4 DUQ. L. REV. 345, 386 (1965), defines *actus reus* more precisely as follows: “[A]*ctus reus* is to be interpreted as the comprehensive notion of act, harm, and its connecting link, causation, with *actus* expressing the voluntary physical movement in the sense of conduct and *reus* expressing the fact that this conduct results in a certain proscribed harm[.]”

⁵⁹ MODEL PENAL CODE, *supra* note 57, § 2.01(2).

Thus, if the acts of a psychopath are akin to reflexes, the psychopath gets off the hook. However, I am quite certain that no serious person would be willing to give the psychopath such a free pass. Yet, in the case of Alex it seems that the judge was willing to view Alex's behavior if not as a reflex, then at least as something more akin to "a bodily movement that otherwise is not a product of the effort or determination of the actor."

Another section of the MPC is relevant here. Section 4.01 provides the elements of an "insanity" defense. The section states that "[a] person is not responsible for criminal conduct if at the time of such conduct as a result of mental disease or defect he lacks substantial capacity to either appreciate the criminality [wrongfulness] of his conduct or to conform his conduct to the requirements of law."⁶⁰ As can be seen, there are two ways to get off the hook based on the MPC's insanity defense. The first is referred to as the "cognitive" prong because it deals with the inability of a person to appreciate the wrongfulness of her conduct. This prong is relatively uncontroversial and is used by most states and the federal government in determining whether someone is "insane" from a legal perspective.⁶¹ The bar for the cognitive prong is quite high, given that even a severely mentally ill person acting under the influence of her illness can be aware that her conduct is wrongful. For example, a paranoid schizophrenic who is commanded by voices in her head to kill her husband may be fully aware that her action is wrong, yet she may feel that the only way to quell her internal commander is to carry out the act. If the

⁶⁰ *Id.* § 4.01. The Model Penal Code's test for legal insanity incorporates the *M'Naghten* rule as well as the "irresistible impulse" test. *See id.* § 4.01 cmt. 2 at 165.

⁶¹ *Id.* § 4.01 cmt. 2, at 165 nn.1-2.

cognitive prong is the only avenue for her to escape a charge of murder, she is out of luck. This is precisely why the authors of the MPC decided to include a second avenue for those suffering from severe mental illness to avoid criminal liability.⁶²

The second prong is referred to as the “volitional” prong and it is far more inclusive than the cognitive prong. Under the volitional prong, one need only lack “substantial capacity” to conform one’s behavior to the law—with the caveat that the impairment is due to a mental illness.⁶³ Thus, a person who murders her husband will likely be able to avoid criminal liability if she is driven by incessant paranoid delusions. From a commonsense perspective, this aligns with most people’s intuition that a person who kills due to a nearly irresistible command from some disembodied internal demon is less—if at all—morally responsible for her behavior than someone who kills after rational, cool-headed, planning. Indeed, in 1972, John Hinckley—the would-be assassin of President Ronald Reagan—successfully used the volitional prong of the insanity defense to avoid a prison sentence.⁶⁴ Based on the court’s determination that Hinckley’s mental illness made it impossible for him to control his behavior, Hinckley was sent to a mental hospital.⁶⁵ The public outcry in response to this verdict led the federal government, and nearly every state, to revise or remove the volitional prong from their

⁶² See *id.* § 4.01 cmt. 2 at 165-68 (discussing the shortcomings and potential injustice of relegating insanity defenses to the cognitive prong).

⁶³ The caveat to this caveat is that the “mental illness” cannot be “manifested only by repeated criminal or anti-social conduct.” *Id.* § 4.01.

⁶⁴ See PETER W. LOW, JOHN CALVIN JEFFRIES JR., & RICHARD J. BONNIE, *THE TRIAL OF JOHN W. HINCKLEY, JR.: A CASE STUDY IN THE INSANITY DEFENSE* 114 (1986).

⁶⁵ See *id.* at 122.

respective insanity defenses.⁶⁶ Thus, it is now significantly more difficult for people suffering from serious mental illnesses to receive the benefit of an insanity defense even if, from a psychological perspective, they are clearly insane.⁶⁷

Psychopathy and the Insanity Defense Under the Model Penal Code

It seems quite clear that a psychopath would not be able to successfully raise an insanity defense based on the current legal definitions of insanity. The psychopath cannot meet the cognitive prong because she knows the difference between right and wrong. Indeed, psychopaths might even be viewed as having a hypersensitive awareness of the boundaries of moral and legal conduct. After all, psychopaths—due to their lack of empathy and consequent inability to intuit social norms—spend their lives developing a carefully constructed persona of normality. This façade is, by necessity, heavily rule-based since the psychopath has to ask herself specific questions such as, “is this behavior morally acceptable in the eyes of others?” in order to conform her behavior to societal standards. This is in stark contrast to how ordinary people form their place in the social, moral, and legal world. Unlike the psychopath, the average person develops her sense of right and wrong primarily through intuitive, socially constructed, cues. For example, I primarily know that torture is wrong because when I imagine someone being tortured I

⁶⁶ See CHARLES PATRICK EWING & JOSEPH T. MCCANN, MINDS ON TRIAL: GREAT CASES IN LAW AND PSYCHOLOGY 99 (2006). Some states simply modified their respective insanity defenses to make a claim under the volitional prong more difficult; however, the federal government eliminated the volitional prong entirely, reverting back to the old *M’Naghten* rule. See Insanity Defense Reform Act, 18 U.S.C. § 17 (2006) (passed in 1984).

⁶⁷ See generally Melinda Carrido, Comment, *Revisiting the Insanity Defense: A Case for Resurrecting the Volitional Prong of the Insanity Defense in Light of Neuroscientific Advances*, 41 SW. U. L. REV. 309, Part II.C. (2012).

feel revulsion, as well as a desire to bring the perpetrators to justice. Of course, I am also aware that society looks down on torture and that is against the law, but these are secondary considerations that inform, but do not create, my sense of the moral and legal unacceptability of torture. A psychopath, on the other hand, is bereft of these natural reactions and must think carefully about whether some act is considered right or wrong. Thus, there is a sense in which the psychopath can be said to have a better understanding of the boundaries between right and wrong than the non-psychopath.

An interesting question is whether the psychopath would qualify for an insanity defense under the volitional prong (assuming that prong were still in use throughout the United States). This would partially depend on whether we accept a biological basis for psychopathy—in the specific sense that psychopathy can be considered a mental disease. The requirement of a biological basis for the disorder stems from the MPC’s stipulation that abnormal behavior “manifested only be repeated criminal or anti-social conduct” does not qualify one for the insanity defense.⁶⁸ This returns us to the reasoning of the judges in the Utah study, who only treated a diagnosis of psychopathy as a mitigating factor after a biological basis for the condition was provided.

Assuming that we do accept a biological basis for psychopathy, the next question is whether it can reasonably be said that the psychopath “lacks substantial capacity” to control her behavior. A reasonable reading of the Utah study would lead to the conclusion that people believe there is at least some sense in which the psychopath is not in control of her behavior. Otherwise, why would the judges have treated a diagnosis of psychopathy as a mitigating factor? After all, there is a biological basis for *all* of our

⁶⁸ MODEL PENAL CODE, *supra* note 57, § 4.01(2).

actions. The fact that our behavior can be tied to our biology is clearly not, by itself, reason to excuse (or mitigate the consequences) of that behavior. If that were the case, then everyone who committed a crime would be able to get off the hook by providing evidence that their behavior resulted, in some part, from their biology. To anyone but the committed incompatibilist, this result would be entirely unacceptable and unjust. The American Psychological Association (APA) used this line of reasoning in supporting abolition of the volitional prong of the insanity defense during the fervor following Hinckley's trial. According to the APA, "The line between an irresistible impulse and an impulse not resisted is probably no sharper than that between twilight and dusk."⁶⁹ In other words, we all succumb to our impulses from time to time—and surely our momentary weaknesses of will are attributable, to some degree, to our biology⁷⁰—but we clearly do not hold ourselves, or others, unaccountable for our actions simply because we did not resist our impulses.

The Boundaries of Judgment for "Impulses Not Resisted"

Of course, there are exceptions. We are often willing to cut others (and ourselves) some slack when they give in to an impulse after they have endured a period of extreme

⁶⁹ American Psychiatric Association, *American Psychiatric Association Statement on the Insanity Defense*, 140 AM. J. PSYCHOL. 681, 685 (1983).

⁷⁰ Indeed, the strength of our wills is subject to the most mundane interferences. For example, a recent journal article demonstrated that our wills are subject to exhaustion—but they can be replenished with a simple administration of glucose (e.g., donuts). See Matthew T. Gailliot & Roy F. Baumeister, *The Physiology of Willpower: Linking Blood Glucose to Self-Control*, 11 PERS. & SOC. PSYCHOL. REV. 303 (2007).

pressure where they had to apply their wills in some other area.⁷¹ For example, while we would generally not tolerate outbursts of yelling from our significant others, we *may* be willing to let such an outburst pass if that person has just finished a long and difficult project at work only to find out she has been assigned another and it is due the following morning. At the same time, there is generally a limit to our willingness to let “impulses not resisted” go. If our significant other, instead of yelling, lashed out with physical violence whenever she was under pressure, most of us would presumably maintain that we have a zero-tolerance policy towards domestic physical violence—even if committed under stress.

The boundaries of this limitation are not immediately clear, however. If the explanation for our willingness to let some impulsivity go is that we believe that circumstances (plus biology, of course) can lead a person to lose control of her behavior, why shouldn’t that tolerance apply to *all* instances of stress-induced impulsivity? In other words, why let the yell go but not the fist? I think this question is central to understanding if the psychopath should be held morally or criminally accountable for her behavior. There is a good deal of literature suggesting that psychopaths tend to act on their anti-social impulses more often under periods of extreme stress. For example, Ted Bundy—the quintessential psychopathic killer—would frequently go for periods when he did not act on his impulses to rape, kill, and dismember, but there was almost always an outburst

⁷¹ As an extreme example, most states reduce the charge for killings committed in the “heat of passion,” e.g., the extreme stress caused by finding one’s spouse in bed with another person, from murder to voluntary manslaughter. See JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 530 (7th ed. 2015). In these cases, the person’s will is assumed to be weakened by the sudden “heat of passion.” See *id.*

of such behavior after stressful events in his life.⁷² Why are we not willing to excuse Bundy's outbursts as momentary, and understandable, weaknesses of will?

One answer is that Bundy's "impulsiveness" indicates a severely disturbed personality that exists at all times—not just at the moment of acting out. This intuition is in line with empirical work on psychopaths. Like other psychopathic killers, Ted Bundy engaged in long periods of careful planning where he would find a target, spend a great deal of time fantasizing about what he wished to do to the target, and finally act on those fantasies.⁷³ The fact that the final act often proceeded a period of intense stress in some other domain of his life does not diminish the fact that his ultimate act was the result of a long process of planning and obsession. This is in stark contrast to the person who yells at her spouse after a particularly stressful day. In general, such outbursts do not represent some long period of quiet resentment and fantasy about lashing out. Of course, if the outburst *is* the result of pent-up frustration and anger, there are bigger problems to address and the spouse will be less likely to let the incident go. In other words, the genesis of the impulsive action is as important as the immediate trigger of the impulse.

⁷² E.g., there was a brief lull in Bundy's killings when he transferred from law school at the University of Puget Sound to the University of Utah in August 1974. While taking the required first-year law school courses, Bundy realized that he could not measure up to his fellow students. He had always prided himself on his intelligence, so his inability to stand out among his classmates was devastating. Feeling overwhelmed and rudderless, Bundy went on a spree where he murdered four women in the course of two months. See POLLY NELSON, *DEFENDING THE DEVIL: MY STORY AS TED BUNDY'S LAST LAWYER* 55 (1994). In general, serial killers tend to go through cycles of murder that occur during periods of intense stress in their lives. See Heather Mitchell & Michael G. Aamodt, *The Incidence of Child Abuse in Serial Killers*, 20 J. POLICE & CRIM. PSYCHOL. 40 (2005).

⁷³ See Christopher J. Ferguson et al., *Defining and Classifying Serial Murder in the Context of Perpetrator Motivation*, 31 J. CRIM. JUST. 287, 288 (2003).

For most of us, we can rarely identify a genesis for our impulses beyond the time shortly preceding the incident. Perhaps we can recognize that we have had a particularly hard week, or month, and the build-up of stress finally had to be released. But that is typically as far as we can trace the source. On the other hand, the psychopath can tell a long and detailed story about the genesis of her impulses. Indeed, the final act—in Bundy’s case, murder and mutilation—is more like the concluding chapter of a horror novel, an almost inevitable consequence of an entire course of voluntary acts and indulged fantasies, than a jarring and incongruous outburst from an otherwise normal character.

Three Reasons to Resist Impulses

Another way to think about this issue is to consider that we are willing to forgive the occasional impulsive outbursts of our friends and family because we recognize that the person may not have been thinking straight in the moment. From there, we can ask ourselves what the person would have done if she *had* been thinking straight. What kinds of things would have motivated her to refrain from lashing out? There are a host of reasons that we can come up with for the normal person. I don’t think it’s a stretch to posit that there are three general considerations that she—along with most of the rest of us—would take into account. First, she would consider our feelings and decide—knowing that we will feel bad if she acts on her impulse to yell—that she cares about us and doesn’t want to hurt our feelings. Second, she would likely also consider the combined effect that the outburst would have on both her and us, in other words, on the relationship. And she would likely decide, when thinking straight, that she doesn’t want to damage the relationship over a silly outburst. Finally, she would probably consider the

impact that the outburst would have on her, alone. Even if she were angry at us, and not particularly concerned about the fate of the relationship in the moment, she might still refrain from acting on the impulse to yell if only to avoid having to deal with the fallout.

Hopefully it's obvious that the third motivation is the most self-oriented of the three. It considers the feelings of the other person only insofar as the reaction created by those feelings will impact the one who lashes out on an impulse. Thankfully, this is likely to be a small factor in the decision of our family and friends not to act in impulsive, hurtful, ways towards us. The primary motivations in normal people will tend to be the first two motivations described above—motivations that take the other person's feelings into account for their own sake, rather than just for the sake of wanting to avoid potentially irritating conflict.

Impulses Not Resisted and Tainted Character

I contend that it is (at least in part) the motivations behind what causes a person to refrain from acting out on her impulses that determines whether we are willing to let it go when she does act on her harmful impulses from time to time. To put this in perspective, imagine how different our reaction to an outburst of shouting from our loved one would be if we found out that she actually wanted to yell at us all the time and only refrained from doing so to save herself the hassle of having to deal with our inevitable emotional response (the third reason described above). In that case, we would almost certainly be less willing to forgive the outburst. Rather, we would begin to feel that these outbursts were a reflection of her true character and we may even start to feel apprehensive around her when she isn't lashing out because we would feel constantly on edge, like she was judging us and holding back her real feelings.

This might explain why we feel less inclined to forgive someone who lashes out with greater degrees of violence. The more extreme the outburst, the more likely we are to believe that the outburst is a reflection of a tainted character. If our loved one strikes us in the face after a stressful day, we will be less willing to forgive the act as an aberration precisely because we view such an outburst as indicative of greater flaw—or as a greater warning. With the psychopath like Bundy, whose outbursts included the most abominable acts we can imagine, this feeling is justifiably magnified to the point where we would find it much more difficult to believe that his outbursts were *not* a reflection of his true character. Indeed, we might begin to think that Bundy’s only motivation for not acting on his impulses all of the time was a desire for self-preservation and if he knew that he would not be caught he would have acted out his fantasies all of the time.

Conclusion

This explanation leads us back to the question of whether it makes sense to place the psychopath’s behavior under the volitional prong of the insanity defense. Yet the question now appears to be more complicated than simply whether the psychopath is subject to “irresistible impulses” or whether her behavior reflects “impulses not resisted.” After all, there must be some impetus for resistance of an impulse. Resistance does not occur in a motivational vacuum. In the next chapter I will consider arguments from two philosophers, that the psychopath should not be held morally (or criminally) responsible for her actions because she lacks empathy and is, therefore, incapable of being motivated on the basis of empathy. I will then present my own analysis of the psychopath’s responsibility, relying on an argument that empathy is best considered as a two-component concept, of which the psychopath only lacks one component.

Chapter Four

DIMINISHED EMPATHY

DEFINING THE PLACE OF EMPATHY

Introduction

In the preceding chapter, I provided a brief history of the diagnosis of psychopathy, including the salient features of the modern diagnosis. I also discussed how the law treats psychopathy, specifically with respect to whether psychopaths might qualify for the insanity defense. I described the two prongs of the insanity defense: the cognitive prong, which covers the inability to recognize right from wrong, and the volitional prong, which covers the inability to conform one's behavior to what is right (whether legally or morally). In this chapter I move to a more explicit philosophical treatment of psychopathy. In Part I, I consider the argument that psychopaths should not be held morally responsible for their acts because they lack empathy—and thus cannot be motivated to act based on empathy. In Part II, I present a counterargument where I posit that the concept of empathy can be broken down into two components: a rational component and an emotive component. I conclude that psychopaths only lack the emotive component of empathy, and the existence of the rational component of empathy is sufficient for the psychopath to be held morally, and legally, responsible for her acts. While I hope that my argument with respect to the responsibility of psychopaths will stand on its own, it will inform the next two chapters where I delve into the responsibility of enhanced empathes.

Part I

Psychopaths as Objects of Moral Judgment

Recall that the volitional prong of the insanity defense exists where a person lacks “substantial capacity” to conform her behavior to the law. A total absence of motivation to act in response to others’ feelings may qualify as something that substantially derails the psychopath’s ability to conform her behavior to the dictates of law. Recalling my argument from the preceding chapter: behavior does not occur in a motivational vacuum. In this vein, Stephen Morse and Jeffrie Murphy defend the position that psychopaths are not proper objects of legal or moral judgment.

Morse’s Argument that Psychopaths Lack Rational Capacity

Morse begins by describing the two criteria that must be present for someone to be held legally responsible for her acts: rational capacity and lack of coercion.¹ It would be more accurate to say that *at least* these two criteria must be met, but Morse is probably right to say that these two criteria are the most important for our willingness to hold someone criminally responsible for her acts.² In any case, Morse argues that psychopaths fail to

¹ Stephen Morse, *Psychopathy and Criminal Responsibility*, 1 NEUROETHICS 205, 207 (2008).

² I say *at least* because the two criteria described by Morse do not perfectly map onto the two prongs of the insanity defense. Rational capacity maps onto the volitional prong quite well, but lack of coercion seems broader than the volitional prong. Lack of coercion applies to any case of legal responsibility. If someone holds a gun to my head and tells me to rob a bank or I will be killed, I am not legally responsible for robbing the bank. However, it would not be the insanity defense that saved me in that case but rather a general defense of having been coerced. The volitional prong of the insanity defense certainly includes the defense of coercion, but it is a special kind of coercion that relates to mental illness.

meet the first criterion because they lack the full measure of rationality. Morse first states the uncontroversial position that in order for someone to be considered rational, she must be able to be responsive to reasons.³

To better understand this statement, imagine what it would be like to be devoid of this responsiveness. A person who was not responsive to reason would be unable to accept the truth of simple propositions that the rest of us take for granted. When I was told as a child that $2+2=4$ it probably took a visual representation to convince me of that proposition's truth. My parents may have gathered four common spoons and placed them on a table. I likely would have been able to identify that there were four spoons on the table. When my parents separated the spoons into two groups of two, which I could also recognize, and then drew them together into a group of four again, a proverbial light bulb almost surely went off in my head. From that display I was able to understand the reason that justified the truth of the proposition $2+2=4$. And from that point on I accepted it as obviously true. Sadly, there are surely people in the world who are unable to understand the truth of $2+2=4$ even after seeing a visual proof. Aside from young children, people who are afflicted with certain mental illnesses, such as paranoid schizophrenia, will often be unconvinced by even the clearest proof of an obvious truth. For example, a paranoid schizophrenic may believe that there is a conspiracy against her. If someone presents a series of irrefutable facts that disprove the conspiracy theory, the paranoid schizophrenic will hear the logical truths of the arguments but will be incapable of being *responsive* to

³ Morse, *supra* note 1, at 208. The argument behind this view is explicated in detail by Fischer and Ravizza, whom I will discuss in detail later in this chapter.

those truths, and thus she will remain convinced of the conspiracy.⁴ A paranoid schizophrenic who is incapable of being responsive to the proof of certain facts cannot rightly be faulted for failing to act in accordance with those facts.

Where a paranoid schizophrenic may be incapable of being responsive to facts that disprove a conspiracy theory, Morse argues that psychopaths suffer from a related lack of ability to be responsive to *moral* facts. According to Morse, “the psychopath has particularized deficits of rationality when moral concern and respect for others is in question.”⁵ Thus, Morse concludes that the psychopath should not be held morally or legally responsible for her anti-social actions. As Morse argues, this deficit in responsiveness to moral reasons is in contrast to the psychopath’s rational capacity with respect to other areas. The psychopath understands concepts like $2+2=4$. The psychopath also has rational capacity with respect to cause and effect. She understands that driving over the speed limit can result in a speeding ticket; that failing to show up for work can result in losing her job; that pointing a gun at someone’s head and pulling the trigger will likely result in that person’s death. The psychopath takes all of these things into consideration when she acts. In other words, she is responsive to these reasons. What Morse sees as critically important, however, is that the psychopath is unable to take moral reasons into account when making decisions. In other words, when the psychopath is weighing her options—sorting out all of the potential factors for and against some course of action—the pain and suffering of other people will never enter her calculus. Moral

⁴ See Chandra Kiran & Suprakash Chaudhury, *Understanding Delusions*, 18 INDUS. PSYCHOL. 3, 5 (2009), for an overview of this phenomenon.

⁵ Morse, *supra* note 1, at 208.

reasons will not be weighed and discarded in favor of more selfish reasons. Rather, they will not be considered at all. Note that in Morse's view, the psychopath is *aware* of concepts such as harm to others, but she is incapable of responding to them, just as the paranoid schizophrenic is incapable of responding to logical arguments that deflate her conspiracy theory.

This is in contrast to how we normally think of immoral or criminal behavior. We generally assume that criminals have access to the same broad range of reasons for and against some course of action as the rest of us—and that they are capable of being responsive to these reasons. For example, when someone decides to burglarize a house, she considers the potential financial gain and the risk of getting caught, along with the fear and loss of security (as well as money) that will be felt by her victim. The burglar weighs these factors and comes to the conclusion that her own gain is more important than the losses to her victim. In many ways, this is the cornerstone of our criminal justice system. Even when it is not fair—for reasons of social and economic inequality—to think that we all have access to the same set of motivating reasons, we treat individuals as though they conform to a normalized standard of human thought and conduct. This is what is meant by the “reasonable person” standard in law.⁶ Even if the reasonable person is a legal fiction, it is not a concept entirely lacking in merit. After all, it would be impossible to determine the exact ratio of reasons available to each person who commits a crime. As long as we can be relatively certain that everyone has access, in some degree, to the broad list of factors that go into decision-making—including moral reasons—it is

⁶ See RESTATEMENT (SECOND) OF TORTS §283 cmt. c (1965), for a legal definition of the “reasonable person.” *See also* JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 134 (7th ed. 2015).

prudent to hold people to a semi-uniform moral and legal standard.⁷ Thus, we hold people morally and legally responsible for their actions in large part because we see them as having considered all of the relevant factors and having come to the wrong—in the moral and legal sense—decision.⁸ Again, even where a person did not, in fact, consider all of the relevant factors, we take account of the fact that the person at least had the capacity to do so. The fact that someone acts without thinking, or in the heat of the moment, does not mean that she was incapable of thinking; rather, her failure to do so suggests recklessness or negligence on her part.

On Morse's view, the psychopath can hardly be accused of recklessness or negligence in the same way. When we scold someone for acting recklessly or negligently, there is an unspoken sentiment that the person knows better. In the language of the MPC, to act recklessly is to “consciously disregard” a known consideration.⁹ To act negligently

⁷ I say “semi-uniform” because there are ways for the criminal justice system to take the unique circumstances of a particular defendant into account at the sentencing and parole phases, so not everyone who is convicted of burglary will receive the same sentence even if they commit identical crimes. The same is true of moral responsibility: while our initial reaction to moral wrongdoing may be to measure the person against a uniform standard, our assessment of that person might change when we learn of mitigating factors unique to that person.

⁸ Of course, in the legal domain there is also a practical motive for using the “reasonable person” standard, viz., it is not epistemically possible to reach into each person's brain and figure out the contours of their individual responsiveness to all kinds of reasons. With moral responsibility the epistemic barrier is less important given that morality does not require that one suffer a consequence, e.g., one can kill a person and get away with it, thus not being subject to legal responsibility, yet one is morally responsible for the act whether or not one is caught. My use of the “reasonable person” standard here is just to highlight that we generally consider people to be responsive to reasons—including moral reasons—when they make decisions.

⁹ MODEL PENAL CODE § 2.02(2)(c).

is to fail to be aware of a consideration that one “should be aware of.”¹⁰ In both cases, there is an implication that one has the capacity to take the consideration into account in the broad sense that one is capable of being responsive to the consideration. In other words, a person cannot act negligently if she is incapable of being responsive to the consideration in question.

Morse applies ought-implies-can reasoning to the psychopath when he describes her as “color blind to moral reasons.”¹¹ The analogy appears to be this: we would not hold a color blind person responsible for hitting and killing a person with her car if she were unable to differentiate the person she hit from the background due to her inability to notice subtle color gradations. Similarly, according to Morse, since the psychopath is unable to recognize moral incentives and disincentives, we ought not to hold her morally (or criminally) responsible for violating a moral norm if she runs someone down with her car, because she is simply unable to take moral reasons into consideration. Indeed, Morse’ argument is that applying the normative language of moral responsibility to the psychopath makes as much sense as saying that tidal waves should not destroy cities, or that cancer should cease its unjustified assault on humanity. As Morse says, “the psychopath is not a member of the moral community, is not a person with whom moral engagement is possible.”¹² Blaming and punishing the psychopath is as pointless as blaming a tidal wave or “punishing” cancer. Of course, this is not to say that psychopaths

¹⁰ *Id.* § 2.02(2)(d).

¹¹ Morse, *supra* note 1, at 209.

¹² *Id.*

should not be *contained* (just as we try to contain the effects of tidal waves and of cancer). Indeed, Morse argues that psychopaths should be subject to civil confinement.¹³

Murphy's Argument that Psychopaths have no Rights

Note that I have presented Morse's view in a relatively positive light. Indeed, I admire his reasoning. Yet I disagree with his conclusion that psychopaths should not be held responsible for their anti-social behavior. Before beginning my criticism of Morse's view, I wish to discuss a similar argument put forth by Jeffrie Murphy.

Murphy takes a much more stark attitude towards the psychopath than does Morris. Murphy argues that psychopaths have no rights as persons.¹⁴ This is because psychopaths lack the attributes that make human beings deserving of rights and respect. According to Murphy, the very notion of "rights" and "respect" between human beings depends on "reciprocity of fairness." Reciprocity of fairness is the idea that one may not claim a right unless one is willing to "recognize and respect" the same right in others facing similar circumstances.¹⁵ In other words, it would be inappropriate for me to steal from another and then turn around and complain that my rights have been violated when another steals from me. As Murphy puts it: "Justice . . . can apply only to those having a sense of justice."¹⁶

¹³ *Id.* at 211.

¹⁴ Jeffrie G. Murphy, *Moral Death: A Kantian Essay on Psychopathy*, 82 *ETHICS* 284, 295 (1972).

¹⁵ *Id.* at 291.

¹⁶ *Id.*

Channeling Kant, Murphy suggests that punishment is a right.¹⁷ In the most common formulation of the argument, the criminal—as a member of the moral community—is entitled to the reciprocity that is inherent to any system of morality. Reciprocity, after all, applies to all duties and, therefore, to all consequences. As Murphy notes, to be a member of a moral community is to deserve both moral approbation as well as moral condemnation.¹⁸ In this sense, membership in the moral community is akin to a social contract. This helps explain the paradoxical idea that a negative condition such as punishment can be construed as a right. Even if the experience is unpleasant, the right to punishment affirms that one is—or, more importantly, has the capacity to be—a member of the moral community. As Murphy explains Kant’s view, punishment allows the criminal to be restored to her proper place in the moral community because it balances the scales of justice that she disrupted when she violated the rights of another member of the community.¹⁹ Retribution—the act of inflicting punishment on someone—puts the criminal, at least in principle, in the same position as the one she wronged. In this way, the wrongdoer is restored to equilibrium with the moral community of which she is a part.

Murphy argues that the psychopath’s lack of empathy means that she is incapable of engaging in reciprocity of fairness and, therefore, cannot be a member of the moral community, and thus the psychopath should be excused from responsibility and

¹⁷ *Id.*

¹⁸ *See id.* at 291.

¹⁹ See JEFFRIE G. MURPHY, *KANT: THE PHILOSOPHY OF RIGHT* 121-22 (1994) for an overview of Kant’s view.

punishment. According to this line of argument, the psychopath is constitutionally unable to defer her self-interest to the rights of others.²⁰ This is not to say, of course, that the psychopath cannot *appear* to defer self-interest, or that she cannot give up opportunities to indulge in her selfish desires; it is simply to say that when the psychopath engages in these personal “sacrifices,” she is still acting for her own self-interest, if in a more complex way. On Murphy’s view, when the psychopath acts in a seemingly selfless manner, she is hiding some motive—for gain or self-preservation. Indeed, according to Murphy, one of the most important differences between the psychopath and the average person is that the latter will often feel herself motivated to act by consciously selfless motives—or at least motives that rely on adherence to a normative moral code for their impetus. The psychopath will not have such feelings. As Murphy puts it, this is the distinction between moral obligations and simply prudential maxims.²¹

According to Murphy, the psychopath’s inability to participate in the moral community means that the psychopath should be considered more animal than man. As with an animal, Murphy states, “[w]e can act wrongly with respect to them, but they cannot be wronged.”²² Moreover, this lack of moral commune means that we cannot hold the psychopath morally responsible for her actions. Like a wild dog that bites our hand, we can get angry—we can even put the dog down—but we cannot rightly hold any moral feelings such as resentment, indignation, or disappointment towards the animal. To do otherwise would be to commit a category mistake, for we would be treating the

²⁰ Murphy, *supra* note 14, at 291.

²¹ *Id.* at 294.

²² *Id.*

psychopath as someone who had transgressed a moral precept. Murphy's argument suggests that moral precepts can only be violated by beings who are not only aware of those precepts, but are—or at least have the capacity to be—motivated by those precepts.

As Murphy sees it, the psychopath is utterly deficient in this regard. Unlike the non-psychopathic wrongdoer—who may be able to put moral reasons aside for a time, but will eventually find herself face-to-face with the guilt that naturally accompanies the violation of another member of the moral community's rights—the psychopath has no moral feeling to push aside, and no moral guilt to overtake her on the descent.

Conclusion to Part I

Morse and Murphy make cogent points regarding the moral status of the psychopath. Nevertheless, I disagree with their conclusions. In the remainder of this chapter I will argue that diminished—or total absence of—emotive empathic response, as the term is commonly understood, does not absolve a person of moral responsibility for her actions. To reach this conclusion, I will begin by laying out an explicit concept of empathy. I will argue that empathy is best understood as a two-part concept, encompassing an emotive component and a rational component. I will conclude that the latter component, which the psychopath does *not* lack, but not the former component, is necessary for moral responsibility.

Part II

Empathy as more than Emotional Responsiveness

As we have seen, Murphy views the ability to reciprocate fairness as a necessary condition for one being a member of the moral community and, thus, capable of being held morally responsible for one's actions. I will not dispute Murphy on this point. Where Murphy goes wrong, in my view, is the point at which he declares the psychopath as a being constitutionally incapable of engaging in reciprocity of fairness *because of her inability to feel empathy*. I take it that Murphy has the ordinary conception of empathy in mind—the conception that empathy involves emotional responsiveness to others' suffering. But why should a failure of emotional response—even if it is a systematic failure rather than a one-time aberration—be relevant to moral responsibility? In other words, Murphy's conclusion is question begging. It is not at all clear that emotional responsiveness to others' suffering is necessary for a person to be motivated to reciprocate fairness. Since Murphy builds his case against the psychopath's moral responsibility on the edifice of reciprocal fairness, his argument unravels if it can be shown that reciprocal fairness can rest on ground that is solidified by something other than emotional responsiveness to others' suffering. As I see it, such ground exists. To see why, it is necessary to delve a bit deeper into the psychological literature on empathy.

Just What is Empathy?

So far I have been playing somewhat fast and loose with the term “empathy.” I have treated empathy, loosely, as “emotional responsiveness to others' suffering,” and I have relied on the reader's intuitive notion of empathy. Aside from the obvious problems of

over inclusion with such a loose definition (e.g., emotional responsiveness could include rage or a desire to inflict harm, which would presumably not constitute forms of empathy), it ignores the complicated nature of empathy and the disagreements within the psychological and philosophical literature about just what empathy is. Indeed, there is no clear consensus regarding the nature of empathy.²³ One difficulty—or, at least, what some researchers see as a difficulty—is that the psychological literature generally views the term empathy as encompassing a host of emotions, the “empathic emotions,” and responses.

The empathic emotions include, inter alia, emotional contagion, sympathy, and affective empathy.²⁴ Emotional contagion is a primitive form of empathy seen in humans and some animals.²⁵ It occurs when one individual “catches” an emotional state from another individual, resulting in personal distress but not any feeling of connection or a desire to help the other individual.²⁶ For example, if you’re in a room filled with unhappy looking people, you might start to feel unhappy yourself. But this wouldn’t require you to feel any distress on the part of the other people in the room.

²³ See Stephanie D. Preston & Frans B.M. de Waal, *Empathy: It’s Ultimate and Proximate Bases*, 25 BEHAV. & BRAIN SCIENCES 1, 1 (2002).

²⁴ See *id.* at 4 tbl.2. See also Heidi L. Maibom, *(Almost) Everything You Ever Wanted to Know About Empathy*, in EMPATHY AND MORALITY 3 (Heidi L. Maibom ed., 2014) (describing, in symbolic terms, four of the empathic emotions).

²⁵ See Maibom, *supra* note 24, at 4; Preston & de Waal, *supra* note 23, at 13.

²⁶ See Maibom, *supra* note 24, at 3 (“Emotional contagion is often thought to be the most basic emotional reaction to others that is still empathic in nature.”).

Sympathy—also called empathic concern—involves feeling concern for another because of her situation, but not necessarily matching the other’s emotional state.²⁷ The researcher Martin Hoffman sees sympathy as more sophisticated than empathy because it involves taking a step beyond emotional identification and considering what is best for the other.²⁸

Affective empathy is probably the closest category to what most people think of when they think of empathy. Generally, it involves state matching, in other words, having the same emotional state as another person, and concern for the other person due to this matching.²⁹ In animals and humans, affective empathy can arise from direct perception of another individual.³⁰ For example, if I run into a friend on the street and she appears upset, I might immediately feel a sense of affective empathy towards her, without the need for any cognitive processing.³¹ It is also possible for humans (and possibly other apes)³² to engage in cognitive empathy—which is sometimes referred to as *true* empathy.³³

²⁷ *Id.* at 6.

²⁸ See MARTIN L. HOFFMAN, *EMPATHY AND MORAL DEVELOPMENT* (2000).

²⁹ See Preston & de Waal, *supra* note 23, at 4, 4 tbl.2. See also Maibom, *supra* note 24, at 3-4.

³⁰ Preston & de Waal, *supra* note 23, at 4-5.

³¹ See *id.*

³² *Id.* at 18-19.

³³ *Id.* at 4.

Cognitive empathy can come about in two ways: one can believe that another is in distress and thus feel a sort of attenuated affective empathy, or one can “imaginatively project” oneself into the place of another and imagine how one would feel in the other’s place.³⁴ Indeed, this latter form of cognitive empathy is often how children are taught to empathize with others. If a child is cruel or a bully to other children, we might ask that child to imagine how she would feel if the bullying were happening to her.³⁵ At least with the latter form of cognitive empathy, it becomes apparent that empathy is as much about oneself as it is about another. In “Interpretation Psychologized,” Alvin Goldman argues that understanding others can only take place by placing oneself in another’s shoes and imagining how *oneself* would feel in a similar situation.³⁶ This process is complex and involves imaginatively placing oneself in a context that might be very different from one’s own. Yet, it involves projecting *oneself* into the place of another—and *not* totally losing oneself and, in a sense, *becoming* the other. In this vein, the model of empathy developed by Stephanie Preston and Frans de Waal holds that “there is no empathy that is

³⁴ *Id.* at 18; Maibom, *supra* note 24, at 2.

³⁵ See MICHAEL SLOTE, *THE ETHICS OF CARE AND EMPATHY* 29 (2007).

³⁶ See Alvin I. Goldman, *Interpretation Psychologized*, 4 *MIND & LANGUAGE* 161 (1989). Note that this can sometimes lead to situations where it might be considered inappropriate for one to claim to empathize with another’s situation. For example, a privileged adult white male who claimed to empathize with the victims of female child sex trafficking in the third world might be ridiculed. See Stephanie D. Preston, *A Perception-Action Model of Empathy*, in *EMPATHY IN MENTAL ILLNESS* 431 (Tom F.D. Farrow & Peter W.R. Woodruff eds., 2007).

not projection, since you always use your own representations to understand the state of another.”³⁷

Preston and de Waal’s model of empathy has been influential in the fourteen years since it was published. Part of its influence likely derives from its treatment of empathy as a range of processes, rather than an attempt to nitpick intractable distinctions between various theoretical positions. For example, when Preston and de Waal say that individuals always use their own representations to understand the state of another, they are not saying that empathy is always self-oriented. Rather, they are saying that empathy is always informed by the self, and by one’s experiences, and that this is a necessary component of empathy—though not the only component. Whether or not any helping behavior that results from empathy is ultimately for the other or for the self-projected-into-the-other is unimportant and unnecessarily complicates the concept of empathy, according to the two authors.³⁸ A study by Daniel Batson, Shannon Early, and Giovanni Salvarani provides empirical support for this position, finding that subjects experience an

³⁷ Preston & de Waal, *supra* note 23, at 17. The two authors use this insight to refine one of the previously popular definitions of empathy given by Martin Hoffman. Hoffman defines empathy as “an effective response that is more appropriate for another’s situation than one’s own.” HOFFMAN, *supra* note 28, at 4 (2000). The refinement offered by Preston and de Waal is that empathy is a state that is more applicable to another’s state than to *one’s own prior state*. See Preston & de Waal, *supra* note 23, at 4. This new definition (although it is more like a broad description given that the two authors spend the entire article trying to conceptualize empathy) allows that empathy, or at least cognitive empathy, involves projecting oneself into the place of another and imagining how one would feel in the other’s place. The inclusion of *one’s own prior state* in the description is significant because it allows that the new state generated by empathy *is* applicable to one’s own *current* state; but this state is generated from encountering or imagining another in distress and projecting oneself into that person’s shoes.

³⁸ See Preston & de Waal, *supra* note 23, at 2, 9.

intractable mix of self-oriented and other-oriented concern when presented with a story of another in distress.³⁹

While both affective empathy and cognitive empathy involve projection under Preston and de Waal's model, cognitive empathy involves it more explicitly. With affective empathy, there often is no conscious cognitive element, in other words, there is simply the perception of another in distress and an automatic emotional response to that distress.⁴⁰ However, affective empathy can be achieved through cognitive empathic processes. I already mentioned how parents will ask their children to use cognitive empathic processes, with the hope of bringing about an affective response in the children. If a parent sees her child fail to respond in a situation where affective empathy should be automatically triggered, the parent might ask the child to imagine how she would feel if she were the other. Sometimes this imaginative projection will trigger a latent affective empathic response. Of course, adults are not immune to failures of affective empathy, and cognitive processes can fill in the gaps when our automatic responses fail us. Indeed, this is implicit in Peter Singer's appeal that we ought not to distinguish between the suffering and need of people in distant parts of the world and those on our doorstep. To see why, let me dig a bit deeper into the mechanism of cognitive empathy.

³⁹ See C. Daniel Batson, Shannon Early, & Giovanni Salvarani, *Perspective Taking: Imagining How Another Feels Versus Imagining How You Would Feel*, 23 PERS. & SOC. PSYCHOL. BULL. 751 (1997).

⁴⁰ See Preston, *supra* note 36, at 433.

Empathy as a Two-Part Concept

As I mentioned above, cognitive empathy—imaginatively projecting oneself into the place of another—can *sometimes* lead to an automatic affective empathic response. In Preston and de Waal’s model, helping is the response and it is triggered automatically.⁴¹ But this needn’t always be the case. If cognitive empathy involves imagining what one would feel like if one were in the position of another, followed by a response, then cognitive empathy involves—at least some of the time—a *rational* step. The process must go something like this: first, one imaginatively projects oneself into the place of another, trying to understand how one would feel in the other’s situation. But, unless an automatic affective empathic response is triggered at this stage, this will not be enough to lead to an outward response. So, second, if there is to be the possibility of an outward response, in other words, helping (or refraining from inflicting harm), one must be able to take a logical step from realizing that one would want help (or to not be harmed) in a similar situation, to understanding that, therefore, one should help (or refrain from harming) the other. This understanding, even if it is devoid of affective empathy directed towards another, is still empathy. It still involves understanding another person, even if that understanding rests of a more coldly logical basis than affective empathy. One can enter into cognitive empathic processes without triggering an affective empathic

⁴¹ Preston and de Waal point out that the likelihood of helping behavior increases with familiarity, similarity, and salience. Preston & de Waal, *supra* note 23, at 4. Also, there is evidence that when the distress of the other is too great, helping behavior falls off. *Id.* at 8.

response.⁴² Yet, as Preston and de Waal put it: “one can be helpful . . . through purely cognitive processes.”⁴³

Due to the logical step involved in cognitive empathy (that does not result in an automatic affective empathic response), I refer to this as the *rational component of empathy* (RCE). This is distinct from the kind of empathy that does not necessarily include a logical step. I group affective empathy, sympathy, and emotional contagion under this latter category, which I call the *emotive component of empathy* (ECE). I refer to them as components because while they usually go hand-in-hand, each can be achieved without the other. ECE is generally an automatic response, but if we are asked to justify that empathic response we will rely on cognitive empathic processes to do so. For instance, imagine that an acquaintance co-worker is laid off from her job. You feel an automatic empathic response (again, this could consist in affective empathy, sympathy, or emotional contagion) and when she asks for your help in finding another job you do so freely. This process would likely be primarily emotive. However, if another co-worker asked why you were helping this person, you would probably turn to RCE to explain. You might say, “I know she would do the same for me,” or, “If I were laid off I would hope that someone would do the same for me.”

So, ECE can act on its own or in combination with RCE. And RCE can sometimes lead to an automatic ECE response.⁴⁴ I have already suggested that RCE can also act on its own, but let me explain further.

⁴² See Maibom, *supra* note 24, at 2.

⁴³ Preston & de Waal, *supra* note 23, at 17.

The Rational Component of Empathy

RCE is a cognitive process. As I described above, it involves going beyond the act of imaginatively projecting oneself into another's shoes, to understanding that one should help (or refrain from harming) the other. There is at least one way that this step can occur. I hinted at it above in the example of the person who justifies why she is helping her co-worker. Let's consider that example in a bit more detail. I'll call the two co-workers Sally and Jane. As I originally described the case, Sally feels ECE towards Jane after Jane is laid off. RCE only comes in when Sally is called on to justify her reaction to Jane's situation. Sally imagines what it would be like to be Jane and states that she would hope that someone would help her were she in Jane's place. Thus, Jane has now engaged both ECE and RCE, with the former leading to the latter.

The process can operate in reverse as well, as it does when a parent—with the hope of triggering ECE—asks her child to imagine how she would feel in another's place. As I mentioned above, adults engage in this rational-to-emotive empathic process as well. And even when ECE is not triggered, one has still engaged in an empathic process—simply a rational empathic process rather than an emotive one. Consider a common scenario. None of us (I would venture to say) feel ECE toward all of the people we encounter on a daily basis who are in some kind of distress. Indeed, even with our friends and loved ones there might be days when we are feeling particularly self-absorbed or

⁴⁴ Depending on how far down the chain of cognitive empathic processing it takes for the ECE response to be triggered, I would either say that cognitive (but pre-rational) empathy leads to the automatic emotive response, or that RCE leads to such a response. As I described above, the automatic emotive response can be triggered at any stage, and thus could be triggered by the simple act of placing oneself in another's shoes (as the example with teaching children how to feel empathy suggested). However, it might also be triggered *after* one has made the logical step described above.

distracted, or perhaps irritated with the other person, and we fail to feel ECE when presented with that friend or loved one's distress. Yet, even in the absence of ECE we can still engage in RCE towards the person. Thus, we might help a friend even though we are annoyed with the friend and feel no ECE at the moment. I contend that at least part of our willingness to help people even when we are not feeling any ECE comes from the logical step embedded in RCE. Even without ECE we can still think to ourselves, "she would do the same for me," or, perhaps more precisely, "I would hope that someone would do the same for me." Of course, with friends and family there is likely to be more to the motivational story. We might help a friend who we are annoyed with, and whom we feel no current ECE toward, because we value the friendship and do not want to let our temporary lack of ECE ruin the relationship in the long term.

Yet even this reason contains elements of RCE. Our close relationships usually rely, at least in part, on the mutual expectation that we will help one another out when one of us is in distress. There is a duty that arises from our empathic connection with our friends and family, and that duty does not disappear when we fail to feel ECE on some occasion. As long as we continue to have RCE, the duty to aid remains, and we may still be held responsible for failing to uphold that duty. To illustrate, imagine that you and a friend have a fight. The next day, the friend gets sick and asks you to bring her chicken soup. Because you are angry with the friend you might very well not feel much ECE towards her. However, as long as your RCE is still operative, in other words, you recognize that you would want and expect your friend to help you if the situation were reversed, you still owe a duty to your friend to help her out—and your friend may justifiably be upset with you if you fail to do so. Indeed, if you told your friend, "I was

annoyed with you so I decided not to help you when you asked,” you would probably be considered quite cruel and your friend might well reconsider the relationship or at least be less willing to help you out in a future situation where she is annoyed with you and you need her help.

It is perhaps easier to understand this notion of a duty arising from the operation of RCE in the case of strangers, rather than friends or family, because there are less complicating factors associated with our relationships with strangers. Several paragraphs ago I suggested that there is an implicit recognition of RCE embedded in Singer’s argument that we owe a duty to help those in far off places. Of course, by that same token, RCE creates a duty for us to help those nearer to us as well. Singer’s argument could be seen to rest in part (though he does not use these words) on the fact that we are more likely to feel ECE towards those near to us than to those in far-off places, yet as long as our RCE is operative with respect to people in far-off places we still have as much duty to help them. RCE is at least partly responsible for the duty to help (or refrain from harming) others. As long as we are able to imaginatively project ourselves into another’s situation and to realize that we would want help (or to not be harmed) if we were in the other’s place, we have a duty to help (or not to harm) the other. Thus, helping (or lack of harming) behavior, à la Preston and de Waal, can and sometimes *should* arise from purely rational (cognitive, in their terms) empathic processes.

So, how does distinguishing between ECE and RCE relate to the psychopath's responsibility for her actions? To begin answering that question, let me return briefly to Murphy's description of the psychopath.

Murphy's View of the Psychopath as Animal

The common way of describing the psychopath—and the way in which I have described her up to this point—is to say that she lacks empathy. But what exactly does this mean? It will not do simply to say that the psychopath doesn't care about other people's feelings. Such a description begs the question: what distinguishes the psychopath's particular lack of feeling from a non-psychopathic individual's failure to take another's feeling into account? After all, we all fail to take the feelings of others into account on a daily basis. Some do this more than others, of course. But in general, it is safe to say that we do not walk around all day placing the feelings of others on equal footing with our own. We act selfishly, even if only in small ways such as deciding not to pick up a phone call from our mother because we can't be bothered to speak with her at the moment, or failing to say "hello" to a passing colleague in the hall. Murphy believes the distinction to lie in the psychopath's utter lack of capacity to feel empathy, which is different than the ordinary person's occasional lapse into selfish, un-empathic behavior. For Murphy, this makes the psychopath morally akin to an animal, as opposed to a human being. But again, this begs the question. What is it in the animal that distinguishes it from a human being? Animals often do have emotional responses to others' suffering—and researchers have confirmed that certain animals react to injustice similarly to humans.⁴⁵

⁴⁵ See Sarah F. Brosnan, *Justice-and Fairness-Related Behaviors in Nonhuman Primates*,

But surely this response in certain animals is more akin to ECE as I have previously described it. It is an empathic emotive response, unaccompanied by cognitive processes, at seeing another in distress. Viewed in this way, the animal is nothing like the psychopath, as Murphy would have it. Animals lack RCE. They are unable to form abstract concepts involving placing themselves in another's shoes (or paws), and they are certainly incapable of making the logical step from identifying with another animal's emotions to realizing that they would want help (or to not be harmed) in a similar situation. On the other hand, the psychopath is perfectly capable of these things. What the psychopath lacks is ECE. The psychopath never feels the sting of ECE that ordinary people feel upon identifying with another who is in distress. As such, the psychopath cannot be motivated to act by this feeling.⁴⁶ On this point, I agree with Morse and Murphy. It would be nonsensical to hold the psychopath accountable for not responding to a motivation that she is incapable of feeling. To do so would be to violate the dictum of *ought implies can*.

The Psychopath and the Rational Component of Empathy

However, the psychopath has no impairment in RCE. The psychopath is often described as a cold, calculating creature. This creature, it is said, will do anything for self-preservation, including sacrificing the rights of others. Yet the psychopath does not suffer from a deficit in the ability to imaginatively project herself into the place of another. As I've stated, the psychopath's deficit is in the ability to experience ECE—the automatic

110 PNAS 10416, 10416-423 (2013), for an overview of cases.

⁴⁶ This point is more fully fleshed out below.

affective response that is triggered upon seeing another in distress. Despite this lack of ECE, the psychopath is able to imagine herself in another's shoes and to make the logical step to recognize that she would not want to be harmed (or that she would want to be helped) if she were in the other's place. As I alluded to above, I think it is fair to say that a duty to help (in certain situations), or not to harm, attaches to an operative RCE. Thus, like the rest of us who have a duty to offer aid to those in need or to refrain from harming others regardless of whether we feel ECE on any given day, the psychopath has the same duty.

One might ask where this duty comes from. For my purposes, it isn't particularly important whether the duty is grounded in moral rationalism, moral sentimentalism, or some other moral system. As long as the psychopath has an operative RCE, she has the same duties towards others as do the rest of us *however those duties are grounded in moral theory*. I contend that regardless of which moral theory one ascribes to, RCE is necessary for moral responsibility to attach. Without RCE, one would be incapable of recognizing that one is part of a world of people, and other sentient beings, that share—at least to some degree—in feelings of sadness, happiness, pain, pleasure, etc. Without some ability to project oneself into another's shoes one would be a justified solipsist with respect to the feelings and rights of others.⁴⁷ Morse and Murphy appear to view the

⁴⁷ There is some evidence that certain highly autistic individuals fit into this category. See R. Peter Hobson & Jessica A. Hobson, *On Empathy*, in *EMPATHY AND MORALITY* 175-76 (Heidi L. Maibom ed., 2014). In terms of my position, these autistic individuals lack both ECE and RCE. With respect to the latter, they are unable to project themselves into someone else's shoes because they do not understand that "others exist *as persons*, that is, as beings with inner lives." Maibom, *supra* note 24, at 17. A person who is truly incapable of understanding that others have inner lives, and thus who is incapable of projecting herself into another's shoes and imagining how she would feel in that position,

psychopath in just this way. But as I have argued, the psychopath has no ECE but has operative RCE. And RCE is the component of empathy that is necessary for moral responsibility. To develop this point, let me consider Michael Slote's argument that empathy undergirds our moral distinctions.

Michael Slote's Ethic of Care and Empathy

Michael Slote builds an entire theory of morality around empathy and caring.⁴⁸ As I said, my argument that an operative RCE is necessary for moral responsibility does not depend on any particular moral framework. However, Slote's theory is illustrative of why some form of empathy is necessary for moral responsibility. According to Slote, actions are morally wrong only if they fail to exhibit empathy towards others.⁴⁹ Note that Slote does not say that actions are morally wrong if the actor fails to *feel* empathy towards others. Instead, for Slote, an action is not morally wrong if it exhibits empathy, even if the actor feels no empathy. Using the example of charitable contributions, Slote states, "even with [a] lack of empathy . . . it is within one's power to fulfill the obligation to make a substantial charitable contribution, and . . . if one *were* to do so, such action or behavior wouldn't *reflect* of *exhibit* (what is in fact) [a] deficiency of empath[y][.]"⁵⁰ Thus, for

cannot be held morally responsible for her actions based on my theory of moral responsibility, which requires RCE to be operative. For the sake of fairness, I have to point out that the kind of severe impairment discussed here is not the norm for individuals with autism. See TEMPLE GRANDIN & SEAN BARRON, *THE UNWRITTEN RULES OF SOCIAL RELATIONSHIPS* (2005), for self-reports from autistic individuals who talk about having concern for others' wellbeing.

⁴⁸ See generally SLOTE, *supra* note 35.

⁴⁹ *Id.* at 31.

⁵⁰ *Id.* at 32 (emphasis in original).

Slote, one can be held morally responsible for one's actions even if one is completely devoid of empathy.⁵¹

My position is somewhat in conflict with Slote's. Under the theory of empathy that I have outlined, it is necessary for one to have RCE in order for one to be held morally responsible. If RCE is inoperative, one cannot be held morally responsible—à la Morse and Murphy—for one's actions. Thus, making my position compatible with Slote's requires adding a caveat to his argument, namely, that one can be held morally responsible for one's actions if one is devoid of ECE but not RCE. I think that Slote would welcome this adjustment. One of the counterarguments that Slote anticipates is that his theory violates *ought implies can* because someone who is devoid of empathy, for example, the psychopath, cannot be motivated to act in accordance with empathy. As Slote's argument stands, I take this to be a viable objection that he does not adequately address.⁵² My distinction between ECE and RCE helps to shore up the response to this objection. I think it's fair to say that Slote is generally referring to ECE when he talks about those who lack empathy. As such, those who lack ECE but have operative RCE can be held morally responsible for their actions because the latter have it "within [their] power" to act in a way that exhibits ECE.⁵³ This is because those who lack ECE but have an operative RCE are still capable of being motivated to act in accordance with the duty that comes with being able to imaginatively project oneself into another's shoes and to

⁵¹ *See id.* at 33.

⁵² *See id.*

⁵³ *See id.* at 32.

understand that one would wish to be helped (or not to be harmed) if one were in the other's place.

My argument that the psychopath, having an operative RCE, is morally responsible for her actions rests, in part, on the idea that she is capable of being motivated to act in accordance with the duty that, I have argued, attaches to operative RCE. Thus, the psychopath's decision not to act in accordance with that duty does not constitute a responsibility-undermining factor. To more fully develop why this is the case, I will turn to the argument put forward by John Fischer and Mark Ravizza in their book, *Responsibility and Control*.⁵⁴

Fischer and Ravizza's Compatibilism – Guidance Control

Fischer and Ravizza begin by describing the concept of "guidance control." Put simply, "[g]uidance control of an action involves an agent's freely performing that action."⁵⁵ The key distinction between whether or not one has guidance control over one's actions rests on whether one is able to deliberate about one's course of action.⁵⁶ Further, this mechanism of deliberation must be, at least, relatively unimpaired. This second point is important because Morse and Murphy could argue that the psychopath does not have an unimpaired deliberative mechanism because she is incapable of being motivated by moral

⁵⁴ JOHN MARTIN FISCHER & MARK RAVIZZA, *RESPONSIBILITY AND CONTROL* (1998).

⁵⁵ *Id.* at 31.

⁵⁶ Guidance control is distinct from regulative control. The latter is what incompatibilists suggest is necessary for free will and consists of the ability to do otherwise—often referred to as the "principle of alternative possibilities" (PAP). For my purposes, there is no need to rehash the arguments between compatibilists and incompatibilists—I will be assuming that compatibilism is true.

reasons due to her lack of empathy. Fischer and Ravizza's concept of guidance control, however, defines an unimpaired deliberative mechanism fairly narrowly. As examples where an agent can be said to lack guidance control, they describe situations where an agent is coerced to act in a particular way. For example, a person who is told to rob a bank by a person holding a gun to her head cannot be said to have guidance control over her actions. Similarly, a person who is subject to irresistible psychological impulses—such as Alex from the preceding chapter—does not have guidance control, if, by irresistible impulses we mean feelings that one's will is being coerced. These are fairly uncontroversial cases.

The psychopath, as someone who is not subject to truly irresistible impulses, can be said to have guidance control. Yet guidance control is not sufficient for moral responsibility. After all, it's fair to say that at least some animals display guidance control—yet we do not hold animals morally responsible for their actions. Moral responsibility, in addition to requiring guidance control, requires the ability to respond to reasons.

Reasons Responsiveness

After introducing the concept of guidance control, Fischer and Ravizza lay out three possible schemas for determining whether someone's ability to respond to reasons makes her morally responsible for her actions. All three schemas shift the focus of moral responsibility away from the agent who engages in an action and towards the mechanism that leads to the action. When we talk about someone who has a truly irresistible urge, we can only mean that there is a mechanism in that person that fails to respond to reasons to do otherwise. It is the failure of this responsiveness, Fischer and Ravizza contend, that

justifies our unwillingness to hold a person who is subject to irresistible impulses morally responsible for her behavior. While this point benefits from further discussion, for my purposes I will leave it here and go on to discuss the three schemas described by Fischer and Ravizza.

The first is *strong reasons responsiveness* (SRR):

Suppose that a certain kind *K* of mechanism actually issues in an action. Strong reasons-responsiveness obtains under the following conditions: if *K* were to operate and there were sufficient reasons to do otherwise, the agent would *recognize* the sufficient reasons to do otherwise and thus *choose* to do otherwise and *do* otherwise.⁵⁷

Note the embedded concepts of receptivity and reactivity. Receptivity is the ability to recognize reasons, while reactivity is the ability to react to reasons once they have been recognized.

Hopefully it is clear that this first schema is too strong to be a necessary requirement for moral responsibility. To see why, imagine that a man, Ronnie, is considering using his position of trust to bilk vulnerable people out of their life savings. Ronnie recognizes that his action is immoral, and that it will ruin the lives of many people. He also recognizes that this moral reason, in the abstract, trumps any selfish gain he will receive. Thus, he is receptive to reasons. However, he decides to go through with the scheme anyway. Under an SRR schema, Ronnie would not be morally responsible for his act because, while he was receptive to the reasons to do otherwise, he failed to react to those reasons. Indeed, his failure to align his action with his recognition of the reasons that cut against that action would be seen as a kind of insanity under an SRR schema. And we do not hold insane people morally responsible for their actions.

⁵⁷ FISCHER & RAVIZZA, *supra* note 54, at 41.

Similarly, if Ronnie were receptive to some good reasons but not to others, he would not be morally responsible under SRR. For example, imagine that Ronnie learns that bilking a particular couple out of their life savings will cause the couple's daughter to die because the parents will be unable to afford her cancer treatments. Ronnie is receptive to this reason; in other words, he recognizes that it is a good reason not to bilk this particular couple out of their money. Imagine, instead, that Ronnie learns that bilking a particular couple out of their life savings will force the couple to move into state-run retirement housing. Ronnie is not receptive to this reason; in other words, he doesn't recognize it as a good reason not to bilk the couple out of their money. Under SRR, because Ronnie is receptive to one good reason but not receptive to another, he is not a morally responsible agent.

This schema is wildly out of sink with our intuitions about moral responsibility. We recognize that people often choose selfish acts over the *right* acts despite their recognition of the overwhelming reasons in favor of performing the right acts. A common example is people who chose to eat meat despite recognizing and agreeing with the reasons against doing so. It would be quite strange to say that all such people are insane. A more extreme example would be a Nazi concentration camp guard who did not agree with the treatment of the prisoners but chose not to help them escape for reasons of self-preservation. Surely, the guard is morally blameworthy. Yet, under an SRR schema, he would not be held morally accountable for his actions.

Thus, Fischer and Ravizza consider a second schema: *weak reasons responsiveness* (WRR). Under this schema, there must simply be a possible world in which there is a sufficient reason to do otherwise, the agent recognizes this reason,

chooses to do otherwise, and actually does otherwise.⁵⁸ To concretize this, reconsider the example of Ronnie. While Ronnie did not choose to avoid bilking vulnerable people out of their money in the example presented above—despite his recognition that doing so would be morally wrong and cause substantial harm—surely there is some reason, or set of reasons, that would cause Ronnie to abandon his plans. For example, he might refrain from his plans if he were sure that one of the victims would likely find out what he had done and would kill his family as revenge. Notice that under WRR, Ronnie is morally responsible for his actions just as long as there is some possible scenario, however remote, where he would respond to reasons.

It seems fair to say that WRR is necessary for attributions of moral responsibility. If the mechanism, *K*, were totally unresponsive to all reasons—no matter how clear and convincing—there is a strong sense in which the agent should be considered to be lacking in free will. Imagine the most extreme scenario where a person is confronted with innumerable reasons to refrain from some course of action, but only one small, insubstantial, reason to engage in the action. If there is no set of conditions that would cause the agent to refrain from the action, the agent is entirely lacking in reasons responsiveness—which could rightly be viewed as a form of insanity. From a psychological perspective (as opposed to a legal one)⁵⁹ this person would be truly alien. For, a hallmark of mature humans, in other words, the type of people whom we hold

⁵⁸ See *id.* at 44.

⁵⁹ From a legal perspective, a person's failure to be even weakly responsive to reasons would be impossible to prove. The best one could do is present a case history that demonstrates the person's failure to be responsive to reasons throughout her life. But this would not prove that there is *no* set of reasons to which the person would not be responsive.

responsible for their actions, is an ability to be responsive to reasons. We do not hold infants or people with severe mental disabilities responsible for their actions because we understand that they are incapable of recognizing reasons—in the sense of *rational* reasons. We do not hold insane people responsible for their actions because we understand that while they may be capable, in some sense, of recognizing reasons (i.e., they may be *receptive* to some reasons), they are incapable of conforming their behavior to the dictates of reason (i.e., they are not *reactive* to reasons).

However, the example of an insane person serves to highlight the problem with making WRR a sufficient condition for moral responsibility. Unlike infants or the severely mentally disabled, there is a possible world⁶⁰ in which the insane person responds to some reason or set of reasons. For example, it may be that a delusional person, Sarah, believes that she must kill the President in order to stop mass genocide. While she may not be responsive to reasons showing that her delusion is unfounded, she may be responsive to other reasons. If someone tells Sarah that her parents will be killed if she goes through with her plans, this may very well stop her from carrying out her plot. If WRR were sufficient for attributions of moral responsibility, Sarah would be morally responsible for killing the President even though she did so under a delusion where the murder was the only way to stop the President from engaging in mass genocide. The fact that Sarah is responsive to at least *some* reason would be sufficient to hold her accountable for failing to be responsive to other reasons, despite her delusional state.

⁶⁰ Note that the use of “possible world” here is somewhat restricted. We hold all things constant besides the existence of a set of reasons that would be recognized and acted upon. Thus, the fact that there is a more broad possible world in which an infant developed super intelligence and, therefore, became capable of responding to reasons is not a consideration here.

Interestingly, this is close to the restricted legal definition of insanity. Recall from the preceding chapter that the post-Hinckley legal definition of insanity excludes an inability to conform one's behavior to the dictates of the law (the volitional prong). The restricted definition only considers a person to be insane if she is incapable of recognizing the difference between right and wrong (the cognitive prong). A person who is incapable of recognizing the difference between right and wrong will, by definition, not recognize moral reasons—or will reliably misconstrue them. On the other hand, someone like Sarah is not legally insane because she recognizes the difference between right and wrong; she is simply delusional about the underlying facts in her case, namely, that the President is trying to kill her.

Moderate Reasons Responsiveness

Fischer and Ravizza make a *prima facie* case for why WRR might be sufficient for moral responsibility. Because they quickly abandon this approach, and because I find it unconvincing for the reasons presented in the preceding paragraph, I will not spend time addressing it here. Instead, let us move on to the final schema presented by Fischer and Ravizza: *moderate reasons responsiveness* (MRR). In formulating this third schema, Fischer and Ravizza argue that the concepts of receptivity and reactivity should be subjected to different tests in determining an agent's moral responsibility for her actions. They begin by refining the concept of receptivity—which, as we saw in the SRR schema is far too strong to be of use in actual ascriptions of moral responsibility, and in the WRR schema suffers from the opposite problem. To fix these problems, Fischer and Ravizza present a kind of Goldilocks solution. They argue that holding an agent morally responsible for her actions requires that the agent has an “understandable pattern of . . .

reasons-receptivity.”⁶¹ In other words, the fact that there is some, possibly highly idiosyncratic, world in which an agent would be receptive to a reason is no longer sufficient. Instead, there must be a pattern of possible worlds that, importantly, form a coherent pattern of reasons receptivity.

To help us understand this point, Fischer and Ravizza ask us to imagine an “imaginary interview” with an agent.⁶² To see if an agent has a minimally coherent pattern of reasons receptivity, the interviewer presents the agent with several scenarios and asks the agent what would count as a reason to act or refrain from acting in each scenario. Some of the scenarios involve taking on another’s perspective. A good example of an incoherent response by the agent is if an she says that a ticket costing \$1000 would be reason not to attend an event, but a ticket costing \$2000 would not be a reason not to attend an event.⁶³ However, if the same agent is able to take on another person’s perspective and values and say that both the \$1000 ticket and the \$2000 ticket count as reasons not to attend the event, then the agent displays, at least, a pattern of possible worlds—or possible viewpoints—that form a coherent pattern of reasons receptivity. As Fischer and Ravizza put it, “[r]egular reasons-receptivity, then, is reasons-receptivity that gives rise to a minimally comprehensive pattern . . . that takes into account the subjective features of the agent . . . but is also *not simply* the agent’s point of view.”⁶⁴ Thus, reasons receptivity does not require that the agent, herself, necessarily identify or feel moved by a

⁶¹ FISCHER & RAVIZZA, *supra* note 54, at 71.

⁶² *Id.*

⁶³ *Id.* at 72.

⁶⁴ *Id.* at 73 (emphasis in original).

particular reason, but only that she is able to adopt a viewpoint that displays a comprehensive pattern of reasons receptivity. This point will become important when we return to the psychopath and her operative RCE.

The requirement of an understandable pattern of reasons receptivity nicely deals with the example of the insane person who is receptive to certain reasons but resistant to others. While Fischer and Ravizza allow that a person's pattern of reasons receptivity will, to some extent, be subjective and idiosyncratic, it is a reasonable requirement that the pattern must be "at least minimally grounded in reality."⁶⁵ This requirement is a good one because it comports with our understanding of psychological states of delusion and insanity. A person suffering from a delusion, say, that she is the Queen of England, will often have an internally coherent set of reasons that justify her belief.⁶⁶ At the same time, she will be utterly immune to the influence of reasons that contradict her delusion. Again, under a WRR framework she would have sufficient control and reasons responsiveness to be held morally accountable for her actions. Yet it seems quite absurd to treat such a person as on morally equivalent footing as a non-delusional person. Moral opprobrium, after all, implies blame, and blame, in part, assumes that a person either consciously disregards a morally relevant reason or, at the very least, is negligent in failing to be aware of a morally relevant reason. I think it is fair to say that such blame ascriptions depend on a general understanding of the blameworthy actor as a rational person. The boundaries of what constitutes a rational person, for the purposes of moral

⁶⁵ *Id.*

⁶⁶ See ALISTAIR MUNRO, *DELUSIONAL DISORDER: PARANOIA AND RELATED ILLNESSES* 10-11 (1999).

blameworthiness, are best described by Fischer and Ravizza's third, MRR, schema because that schema best comports with our reactive attitudes to people suffering from delusions.

Interestingly, this last point is, paradoxically, best exemplified by many people's resistance to letting criminal defendants off the hook based on an insanity defense. The arguments put forward by these people rarely (if ever) suggest that people suffering from delusions should be held morally—and legally—responsible for their actions. Rather, the arguments take the form of questioning the defendant's diagnosis, or claiming that the defendant is faking her mental illness. The vociferousness of these arguments indicates, to me, that people fully recognize that mental delusions render a person morally blameless for her actions. As such, those who speak up against the insanity defense do so because they recognize the power of such a defense, and also recognize the great incentives for a defendant to fake mental illness.

Receptivity and Reactivity to Reasons

While MRR requires that an agent's pattern of receptivity to reasons be grounded in reality, it does not require that the pattern conform to some greater objective standard of appropriate reasons receptivity.⁶⁷ The fact that a person is extremely selfish and rarely accedes to reasons that the majority of us might consider objectively compelling does not mean that this person is unreceptive to reasons and, thus, not a proper subject for moral condemnation. Once again, this comports with our intuitions. Note how the issue of motivation is imbedded here. A person who chooses to live a life of selfish indulgence at

⁶⁷ FISCHER & RAVIZZA, *supra* note 54, at 73.

the expense of others should not be able to escape moral (and legal) blame for her actions simply by claiming that she was not motivated to act in response to reasons that the majority would accede to. Fischer and Ravizza make this point explicit when they refine the concept of reactivity (as apart from receptivity) to reasons.

Unlike receptivity, which Fischer and Ravizza argue (rightly, I believe) must be elevated above mere weak responsiveness while not fully approaching strong responsiveness, reactivity—they argue—can remain only weak and still be grounds for holding an agent morally responsible for her actions. As Fischer and Ravizza put it, “reactivity is all of a piece in the sense that the mechanism can react to all incentives, if it can react to one.”⁶⁸ To be quite clear, this is not the same as saying that if an agent is *receptive* to one reason, the agent can be receptive to all reasons. Instead, the point is that once an agent recognizes—or is receptive to—a reason, that agent *can* react to that reason insofar as we have determined that the agent can react to at least one reason. Furthermore, this does not require—as per SRR—that an agent always respond to reasons to which she is receptive. Again, if that were the case then any agent who was receptive to a reason but failed to act on it would not be considered morally responsible for her actions. Right off the bat, notice how this requirement would comport with the intuitions I have described in the preceding few paragraphs. Indeed, I don’t think it is a stretch to say that a quintessential part of our holding others (and ourselves) morally responsible for their actions involves assuming that a person is aware of reasons dictating against some course of action yet chooses to take the action anyway.

⁶⁸ *Id.* at 74.

MRR, then, fulfills two requirements of an acceptable theory of moral responsibility. First, it requires that a person's receptivity to reasons must represent a pattern. The pattern need not adhere to some objective standard of morality, but it must be grounded in reality. Second, it requires that a person who is moderately receptive to reasons (in the sense described above) is morally responsible for her actions if she is at least weakly reactive to reasons. That is, a person who recognizes reasons that conform to some coherent pattern is morally responsible for her actions if she is reactive to just one of those reasons.

The Emotive Component of Empathy Revisited

Let us now return to the case of the psychopath. Recall that the psychopath is generally considered to be devoid of empathic feeling. Also recall that, as I have argued, the concept of empathy can be divided into two components. The first is the emotive component—ECE. It hits us when we see someone in distress; it is automatic and generally leads to the accompanying act of helping (or refraining from harm). As I discussed earlier in this chapter, ECE is significant and represents the “face” of empathy, as it were. Often, when we say that we feel empathy for another person we do not engage in any intellectualization with respect to our feeling. Instead, we simply recognize a particular feeling as empathy. Unquestionably, the psychopath lacks ECE. The psychopath never feels an *emotion* that she can identify as empathy. On Fischer and Ravizza's model, the psychopath is not receptive to ECE as a reason because the psychopath is incapable of understanding it as a reason even if she takes on another's perspective, because the psychopath has no frame of reference for ECE. Consequently,

the psychopath cannot feel motivated to act in response to an emotion that she is wholly unreceptive to. On this point, Morse and Murphy are right about the psychopath.

Where Morse and Murphy get it wrong is in ignoring RCE. As I argued earlier in this chapter, this second component of empathy is the *rational component*, which involves the ability to go from imaginatively projecting oneself into another's shoes, followed by understanding that one would want help (or to not be harmed) if one were in the other's place, to realizing that, therefore, one should help (or not harm) the other. Indeed, when we explain empathy we tend to do so in terms of RCE. If someone asks us, "why do you empathize with such-and-such a person?," we generally will not provide the tautological response that we empathize with the person because we feel empathy for the person. Instead, we will seek to justify our feelings by describing the reasons why we empathize with the person, which include how we would feel in the other's situation. It's important to note that these reasons are not always post-hoc rationalizations (although they might be in a case where we are trying to explaining a purely emotive empathic response). Rather, they represent another reason, other than emotion, to respond to others in an empathic manner. In other words, the reasons we give for our empathic feelings are not only justifications for our feelings, but can also be reasons for why we have certain obligations towards people *in spite* of not experiencing ECE.

The Rational Component of Empathy and Moral Responsibility

Take a concrete example. If I see a person take a spill on her bike and she cries out in pain, I might have an immediate feeling of empathy for that person—ECE. At the same time, I will feel that I ought to do something to help her because she has hurt herself and I am in a position to help (or at least to call someone else for help), and I would surely

want someone to do the same for me. Consider how we would react if we witnessed, from a distance, a person falling off her bike and a nearby bystander stopping to look and then simply walking away. The question, as I have presented it, becomes this: would we hold the bystander accountable for her failure to *feel* ECE towards the biker, or would we hold her responsible for failing to act on good reasons to help the biker, independent of feelings?

We often conflate the two judgments. We might say, “how could she be so uncaring?,” which seems to reflect a judgment on her subjective feelings. But consider a scenario where the bystander does help the biker and we have the opportunity to question her as to why she did so. The bystander is likely to give us reasons that reflect RCE at work. She might say, “If I were in a similar situation I would hope that someone would help me,” or, “I could imagine myself falling off a bike and how much pain I would be in.” Of course, she might also give reasons that reflect ECE, such as, “I felt her pain.” Nevertheless, even without ECE, the reasons that the bystander provided with respect to RCE stand on their own. This makes sense because we generally do not form ought statements from the vantage point of emotion alone. In other words, we hold people responsible for their failure to act on reasons, not for their failure to act on emotions (although emotions can form the basis for reasons).

The Psychopath as Moderately Reasons Responsive

Indeed, Fischer and Ravizza’s theory of moral responsibility considers how receptive and reactive people are to reasons, and not to emotions. Thus, under MRR a person is morally responsible for her actions if she manifests a coherent pattern of receptiveness to reasons and if she is reactive to at least one of those reasons. As Fischer and Ravizza

convincingly argue—and as I have tried to bolster and elucidate here—this schema best captures the necessary conditions for holding a person morally responsible for her actions. So, the question becomes: does the psychopath fit into the MRR schema? Let me begin by saying that under SRR, the psychopath would unquestionably not be responsible for her actions. The psychopath is not receptive to all good reasons and does not react in response to all good reasons. But recall the weakness of that schema: *none* of us have such a perfect fit between our receptiveness to reasons and our propensity to act in response to reasons. A schema that does not hold ordinary people responsible for their actions is no better used to hold the psychopath responsible.

The psychopath does, however, manifest a coherent pattern of receptiveness to reasons that is grounded in reality. As noted earlier, even if the psychopath is not subjectively receptive to certain reasons, she is capable of taking on another perspective and picking out good reasons from that perspective. In this sense, the psychopath is like the agent in the imaginary interview who is able to recognize that, from another person's perspective, something might count as a good reason, even if the agent—here, the psychopath—does not, herself, recognize it as a good reason.

The Psychopath as Rationally Empathic Agent

Also, the psychopath is not driven by irresistible urges that literally overpower her will to do otherwise. The psychopath is reactive to many good reasons to do otherwise. Reconsider the case of Ted Bundy. During the time of his killings Bundy volunteered nights at a suicide prevention help line. In that capacity he worked alongside a woman, Ann Rule, who was—coincidentally—working on a book about a series of unsolved

murders that had taken place in Washington.⁶⁹ During the months they worked together, Bundy had ample opportunities to harm Rule. They worked late at night and into the very early morning, and Bundy would walk Rule to her car when their shifts were over. Yet, Bundy never harmed Rule. Indeed, the two were friends, and Rule—like so many others who knew Bundy—described him as a caring, conscientious, person. Writers and psychologists have spilled gallons of ink trying to make Bundy into a cold-hearted person, through and through. They have taken pains to interpret Bundy’s volunteerism at the suicide help line through that lens.

I do not dispute Bundy’s coldness. However, it seems clear that Bundy was receptive and reactive to reasons. I see no reason to think that he didn’t recognize that suicidal people were in need of help, and that he, too, could become suicidal and that he hoped that were he in that position he would want others to help him.⁷⁰ This line of reasoning could have motivated him to volunteer his time at the suicide help line.⁷¹ Note that I did not speak of Bundy’s compassion for others. To do so would imply that Bundy

⁶⁹ See ANN RULE, *THE STRANGER BESIDE ME* xxxix (Pocket Books 2008) (1981).

⁷⁰ It’s not important that I get the exact schema of Bundy’s motivation right. Rather, what’s important is that his motive was almost certainly not selfish, in the traditional sense of that term, i.e., at least part of his motivation derived from a desire to help others. While a psychological egoist might argue that even seemingly altruistic motives are selfish at their core, this does not negate a distinction between self-directed motives and other-directed motives—it is simply reconfiguring the distinction as between two kinds of ultimately selfish motives.

⁷¹ While this is my own—I think, perfectly reasonable—interpretation of Bundy’s actions in this situation, there is another—more banal—story from his own mouth that confirms that he was receptive and reactive to reasons. Bundy describes how after he had escaped from the jail in Colorado he broke into several cars to find a getaway vehicle. If he broke into a car and got the impression that the owner was not well off, he would leave the car and search for another. According to Bundy, he did this because he, too, had been poor. See RULE, *supra* note 69, at 319.

felt ECE—which he surely did not. Rather, I framed his reasons, and his actions, in terms of RCE. Now, it’s entirely possible that Bundy could have failed to be reactive to this set of reasons and not volunteered his time at the help line. Indeed, when it came to his victims he also appears—from his later interviews—to have been receptive to the reasons against killing them. Yet, he was not reactive to those reasons. Instead, he was reactive to his own set of reasons that, however, deranged they may seem to us, fit a pattern that was connected to reality. I say this to distinguish Bundy’s selfish, perverse reasoning from the “reasoning” of a truly insane person. Recall the case of Sarah, who has an intractable belief that she has to kill the President in order to stop a mass genocide. The reason we should not hold Sarah morally responsible for her actions is that she is not MRR. More specifically, while Sarah is receptive to some good reasons, her pattern of receptivity to reasons does not form a pattern that is even remotely connected to reality. For example, if I tell Sarah that the President is not planning a mass genocide and provide her with reams of proof to that effect, Sarah—as a paranoid schizophrenic—will not be receptive to these reasons. Rather, she will be convinced that I am simply naïve or stupid—or perhaps that I, too, am part of the conspiracy.

Contrast this with Bundy’s receptivity and reactivity to reasons. Bundy did not believe that he was killing his victims for some greater purpose. He knew full well that he was simply satisfying his own twisted inner urges. These urges were not irresistible and Bundy identified with them as being his own. Yet, it was not the urges alone—or Bundy’s identification with them—that led him to kill. The psychological literature makes it clear that many ordinary people have urges to kill—and they may even identify with these urges, to some extent—but their capacity for empathy, more precisely, their

capacity to experience both ECE and RCE, stops them from acting on their urges.⁷² Thus, it was a rare alignment of psychological drives, emotional coldness, and selfish desires that led Bundy to kill. But it is imperative to recognize that Bundy's world-view—his receptivity and reactivity to reasons—did not mean that he was destined, in any sense of the word, to kill. Bundy chose to kill—in the sense that he identified with his desire to kill and did not feel as though his will was being compelled by some outside force.

The Historical Nature of Agency in Fischer and Ravizza's Theory

This leads me to another aspect of Fischer and Ravizza's theory that is relevant to whether a person should be considered morally responsible for her actions. Recall that under all three of the schemas presented by Fischer and Ravizza, there is a requirement that the mechanism be the agent's own. In other words, the agent must feel—as most of us do—that her choices reflect her own will, in other words, that she isn't subject to a universe over which she has no control. In this sense, Fischer and Ravizza's theory is historical because it requires that an agent have a proper history of learning that her actions are the result of choices and that, due to this freedom of will, she is rightly held responsible for her actions.

To understand this point, imagine a world in which it was not the case. In this alternate world, a three-year-old child who hits her baby sister is not reprimanded; nor is she reprimanded when she misses the bus to preschool; and so on. Instead, the parents

⁷² This distinction between the psychopath and the ordinary person is exemplified by Jeffrey Dahmer and his father, Lionel Dahmer. After Jeffrey's arrest and imprisonment, Lionel wrote a moving memoir, *A Father's Story*, in which he described his own inner demons of murder and control. Unlike his son, Lionel never acted on his urges. The only difference between father and son appears to have been their respective capacities to experience ECE. See generally LIONEL DAHMER, *A FATHER'S STORY* (1994).

treat her actions as the result of deterministic fatalism—and one can hardly be reprimanded for obeying the laws of nature. In such a world, the child would never come to see herself as responsible for her actions. Indeed, if someone were—in middle school, say—to reprimand her for failing to do her homework, she would likely feel resentment rather than guilt. “Who is this person to judge me for things over which I have no control?,” she might think. So, in order for a person to feel that she is an agent, she must be subjected to suitable conditions that inculcate in her a feeling of agency.⁷³

Reactive Attitudes

Along with learning to feel that she is an agent, in order to meet the threshold requirement of being a morally responsible agent, a person must also come to learn that she is, at least sometimes, a fair target of other people’s reactive attitudes.⁷⁴ It might be more appropriate to say that learning that we are fair targets of other people’s reactive attitudes is *how* we learn that we are agents. Indeed, when we are young children one of the first things we learn is that the people around us—usually our parents—praise us when we act in ways they approve of, and blame us when we act in ways they do not approve of.

However, it’s not difficult to imagine a scenario in which the process could go wrong. If a parent habitually blames a child for things over which the child has no control—or, only minimal control—that child is more likely to feel that she is *not* a fair target of other people’s reactive attitudes. This is because the child will learn that

⁷³ See FISCHER & RAVIZZA, *supra* note 54, at 208.

⁷⁴ *Id.* at 211.

people's responses to her—their praise and blame—are not related to her actions as a free agent. Instead, she will feel that other people's reactive attitudes are capricious and unreliable. The child might well feel guilt when she is blamed—as we are all inclined to feel when we are reprimanded as young children. Yet, over time that feeling of guilt is likely to become tinged with resentment. Thus, it is possible to imagine a scenario in which a child was subjected to such arbitrary and capricious judgment that she would not feel herself to be a fair target of other people's reactive attitudes.⁷⁵

The Psychopath as Agent and Fair Target of Reactive Attitudes

What about the psychopath? Surely there are extreme examples of psychopaths who lived through such horrendous childhoods that it might be right to say that they do not consider themselves to be agents or fair targets of other people's reactive attitudes. But this is not a condition of psychopathy, per se. Such a disconnect from reality would be more characteristic of some other mental illness on top of psychopathy. Schizophrenics, as we have seen, sometimes feel that their wills are not theirs and, thus, it would be unfair for other people to hold reactive attitudes of moral blame towards this group. I want to emphasize that psychopathy, as a condition onto itself, is not characterized by the person feeling that she is not an agent with free will or, indeed, that she is not a fair target of other people's reactive attitudes.

⁷⁵ I imagine that such situations—where the child is subject to such an overwhelming assault of capricious judgment that she does not feel herself to be, even when she is acting freely, a fair target of other people's reactive attitudes—are very rare.

Consider my go-to example of Ted Bundy. Bundy was raised by a loving mother and a responsible, if somewhat ambivalent, stepfather.⁷⁶ He did well in school and did not get into serious trouble around the neighborhood. When he did well, he would be praised and rewarded; when he got into trouble, he would be punished. Bundy took pride in his academic accomplishments, and he certainly never felt that was not in control of his intelligence. Significantly, Bundy was a law and order man. If anything, he was hyper-conscious of the necessity for personal responsibility and he recognized the importance of reactive attitudes in ensuring that people who did wrong were punished. As a young man, Bundy worked on Governor Daniel J. Evans' re-election campaign and, while in law school, Bundy devoted his time and energy to conservative causes. When Bundy was finally caught and eventually confessed to his crimes, he was not indignant that someone would hold him responsible for the things he had done. He recognized that it was fair for him to be held accountable, even if he would personally prefer not to be held accountable. Whatever else one may say about Bundy, he was surely aware that he was a fair target of other's reactive attitudes.

Conclusion

In the end, the attributes of the psychopath are not such that she is immune from moral blame. While she cannot be motivated by ECE, the psychopath has no deficiency in her ability to imaginatively project herself into another's shoes and to realize that she would want to be helped (or not to be harmed) if she were in the other's place. Thus, she can be

⁷⁶ The woman that Bundy called his mother turned out to be, in fact, his older sister. Much has been written about how this affected Bundy's outlook on women. *See, e.g.,* RULE, *supra* note 69, at 51-52. But this troubling revelation is not relevant to Bundy's learned sense of agency.

motivated by RCE as much as the rest of us. Under the model of moral responsibility presented by Fischer and Ravizza, the psychopath meets all of the criteria. She understands herself to be an agent, and she understands that she is a proper target of other people's reactive attitudes. She is MRR: she is receptive to a coherent pattern of reasons that cohere with reality, which includes the ability to take another person's perspective and imagine what that person would recognize as a good reason (a hallmark of RCE), and she is reactive to at least some reasons that suggest that she shouldn't engage in certain acts. The fact that she is incapable of feeling ECE is no reason to discount her moral responsibility. At best, this lack of ECE is cause for us to pity the psychopath. The psychopath is not to blame for her inability to feel ECE. It is a biological condition that leaves her without a crucial component of what it means to be fully human.

In this chapter and the last I have focused on people who have a diminished capacity for ECE. I have discussed the psychopath as the quintessential case. Indeed, the psychopath has more than a diminished capacity for ECE; in the psychopath, that component is utterly lacking. Notwithstanding this deficit, I have argued that the psychopath is morally responsible for her actions because she suffers from no deficit in RCE. Another purpose in these two chapters has been to present the case that RCE, but not ECE, is necessary for moral responsibility. It follows, and I have argued, that diminished ECE, on its own, is not relevant to ascriptions of moral responsibility.

In the following chapters I will make what I believe to be a more controversial claim. Namely, I will argue that an *enhanced* capacity for ECE, unlike a diminished capacity for that component, *is* relevant to our ascriptions of moral responsibility.

Chapter Five

ENHANCED EMPATHY

THE POSSIBILITY OF ENHANCING EMPATHY

Introduction

In the previous two chapters, I provided an argument for why people with a diminished capacity for empathy—specifically, a diminished capacity for ECE, as I defined it—are morally responsible for their actions. In this chapter and the next, I will turn to the opposite side of the coin and consider the moral status of people who have an enhanced capacity for ECE.

In this chapter I will argue that we live in an era where the possibility of artificial enhancement of empathy (along with other brain / mind states) is quite real. In these chapters, I hope to present an analysis of the moral status of the enhanced empath *before* such beings come into existence. It is my hope—perhaps a naïve hope—that an ex ante analysis of the moral status of enhanced empaths will provide for less confusion, both moral and legal, when these beings first make their appearance on the scene. In this chapter, I will discuss the feasibility of artificial enhancement of the human empathic response. While not necessary for a purely theoretical argument, I hope to ground my work here in practical concerns. A brief overview of the state of the science regarding brain / mind enhancement will provide such a ground. In the following chapter, I will present my analysis of the moral status of the enhanced empath.

The Need For Enhanced Empathy

Recall that in Chapter Two I provided an argument—relying, in part, on Singer's metaphor of the “expanding circle”—for why it might be desirable, or perhaps even morally required, to artificially enhance human empathic response. To summarize that argument:

P₁: We are morally obligated to help others to the extent of our abilities and means.

P₂: Our willingness to help others is determined, as least in part, by our motivation to help others.

P₃: Our motivation to help others is determined, at least in part, by our feelings of empathy towards others.

C₁: All else being equal, an increase in empathic response will lead to an increase in motivation, and thus an increase in willingness, to help others.

P₄: Evolution, insofar as it has imbued us with feelings of empathy at all, has limited the reach of those feelings to a fairly small group.

P₅: Evolution's haphazard path does not necessarily align with human aspirations.

P₆: We cannot rely on natural evolution to expand the sphere of empathic feeling.

P₇: There exist safe and effective means by which to artificially enhance empathic response.

C₂: We ought to artificially enhance our empathic response.

In fact, in Chapter Two I defended P₁-P₆, but I did not address P₇.¹ Indeed, I will not be defending P₇, as written, in this chapter either. Instead, I will argue that P₇ will be true in the not-so-distant future based on an overview of currently existing technologies and the current state of brain science.

The Current State of Brain / Mind Enhancement Technologies

There is good reason to believe that P₇ will be true in the not-to-distant future. Even ignoring theoretical speculations about the coming “Singularity,” where human beings will be able to meld ourselves with machines to become superhuman and immortal,² there is more prosaic—yet still rather incredible—evidence that the potential to enhance specific aspect of our brains / minds is not far off.

Deep Brain Stimulation

Consider deep brain stimulation (DBS). DBS is a method of neurostimulation that uses electrical signals to stimulate specific areas of the brain in order to gain some beneficial response.³ DBS can be seen as a refined version of electroconvulsive therapy (ECT). Yet, the two technologies could not be more different from a scientific, and pragmatic, perspective. With ECT, which is still in use today, there is little understanding of the mechanism by which a large pulse of electricity directed over an undifferentiated portion

¹ Note the hidden premise of *ought implies can*. We cannot be obligated to enhance our empathic response if we do not have the means to do so.

² See RAY KURZWEIL, *THE SINGULARITY IS NEAR* (2005), for a detailed account of this fanciful future.

³ See WALTER GLANNON, *BIOETHICS AND THE BRAIN* 136 (2007).

of the brain—indeed, over the entire brain—causes beneficial responses in some patients.⁴

On the other hand, DBS is relatively well understood, even in its infancy. It began as a therapy for Parkinson's disease, which is a disease that causes electrical abnormalities in specific areas of the brain. Clinicians can identify the areas of the brain affected. If the affected area is not too close to a particularly important part of the brain, clinicians can implant an electrode with a great deal of precision. The implanted electrode is attached to a wire that is surgically implanted so that it travels over the skull, down the neck, and into a “brain pace maker” placed under the skin near the shoulder and upper chest.⁵

Clinicians fine-tune the pace maker to send measured jolts of electricity to the implanted electrode. The effects can be quite dramatic.⁶ Patients who experience debilitating parkinsonian tremors are suddenly brought to a state of calm. The targeted jolts of electricity have the effect of calming the electrical storm that rages out of control in the parkinsonian brain. The treatment has been recognized as a breakthrough for Parkinson's patients whose symptoms are not responsive to other, less invasive, forms of treatment.⁷ This is not to say that DBS for the treatment of Parkinson's is not without its

⁴ See Roar Fosse & John Read, *Electroconvulsive Treatment: Hypotheses About Mechanisms of Action*, 4 FRONTIERS IN PSYCHIATRY 1 (2013).

⁵ Researchers have already developed wireless systems that would obviate the need for a surgically implanted wire, but these have yet to be put into use outside of the laboratory. See, e.g., R Chen et al., *Wireless Magnetothermal Deep Brain Stimulation*, 347 SCIENCE 1477 (2015).

⁶ See GLANNON, *supra* note 3, at 136.

⁷ *Id.*

risks. If the electrode must be placed in a sensitive area or if the clinician gets the precise placement wrong, there can be unwelcome side effects.⁸

DBS has also been used, though far less regularly, in the treatment of intractable depression. In 2005, Mayberg et al. published a paper in which they identified that the subgenual cingulate region of the brain (otherwise known as “Brodmann area 25”) is overactive in treatment-resistant depression.⁹ Brodmann area 25 is a narrow band of brain tissue deep inside the brain.¹⁰ Its characteristics of being rich in serotonin receptors and acting as a “governor” of the amygdala, which regulates mood, and the frontal cortex, which regulates higher order thinking, makes it a reasonable candidate for being implicated in depression.¹¹

In their study, Mayberg et al. took detailed MRIs of the patient’s brains, which they used to guide the implantation of an electrode in the Brodmann area 25 region.¹² When the electrode was correctly placed and turned on, the patients reported remarkable

⁸ However, side effects from DBS generally pale in comparison to those from drug treatments. For example, there is a well-known phenomenon of Parkinson’s sufferers becoming compulsive gamblers after taking the drugs pramipexole and ropinirole. *See id.* at 138.

⁹ Helen S. Mayberg et al., *Deep Brain Stimulation for Treatment-Resistance Depression*, 45 NEURON 651 (2005).

¹⁰ *Id.*

¹¹ Other researchers have targeted different areas of the brain for DBS treatment of depression. These include the ventral anterior internal capsule (the area targeted in OCD treatment), the nucleus accumbens, the inferior thalamic peduncle, and the lateral habenula. *See* Cristina Cusin & Darin D. Dougherty, *Somatic Therapies for Treatment-Resistance Depression: ECT, TMS, VNS, DBS*, 14 BIOLOGY OF MOOD & ANXIETY DISORDERS 1 (2012).

¹² Mayberg et al., *supra* note 9.

and instantaneous transformations. The patients described “‘sudden calmness or lightness,’ ‘disappearance of the void,’ sense of heightened awareness, increased interest, ‘connectedness,’ and sudden brightening of the room[.]”¹³ Following the surgery, the patients were sent home with the implant set to deliver DBS continuously at a set voltage.¹⁴ The initial effects that occurred upon first stimulation continued as the patients resumed their normal lives. They experienced all of the feelings that are generally associated with remission of depression: increased energy and interest in activities, renewed pleasure in social activities, and an ability to make plans and complete tasks that had previously seemed impossible.¹⁵

DBS has also recently been approved for limited use in the treatment of obsessive-compulsive disorder (OCD).¹⁶ OCD symptoms have been linked to the anterior limbs of the internal capsule—part of the structure of the striatum, which is a critical part of the brain’s reward system.¹⁷ A three-year follow-up study in 2006 found that OCD patients who received the DBS treatment showed significant improvement in their

¹³ *Id.* at 652.

¹⁴ *Id.* at 653.

¹⁵ *Id.* at 654.

¹⁶ United States Food and Drug Administration, *FDA News Release: FDA Approves Humanitarian Device Exemption for Deep Brain Stimulator for Severe Obsessive-Compulsive Disorder*, FDA.GOV (Feb. 19 2009), <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm149529.htm>.

¹⁷ GLANNON, *supra* note 3, at 133.

symptoms, including comorbid symptoms of depression and anxiety.¹⁸ However, the positive effects for OCD patients were not as significant as those seen in Parkinson's and depression patients.¹⁹

There are videos available that show DBS in action in the case of Parkinson's patients and patients with treatment resistant depression.²⁰ These videos—in conjunction with the scientific papers demonstrating the efficacy of DBS for these patient populations—provide compelling evidence that DBS has the potential to work wonders that appear almost magical in their immediacy and strength. One question that naturally arises is why, given the remarkable results from so many documented cases, DBS isn't a first-line treatment for Parkinson's and depression? The answer lies in the invasive nature of the technology. All major surgery is risky, and when the brain is involved it becomes especially dangerous to have someone cutting into the fragile material and streaming electricity into a highly interconnected system. With great risk comes great expense. Thus, insurance companies are reluctant to approve the use of DBS, especially in patients who have not tried less invasive—and less costly—methods of treatment.²¹

¹⁸ Benjamin D. Greenberg et al., *Three-Year Outcomes in Deep Brain Stimulation for Highly Resistance Obsessive-Compulsive Disorder*, 31 NEUROPSYCHOPHARMACOLOGY 2384, 2390 (2006).

¹⁹ *Id.*

²⁰ E.g., Andrew Johnson, *The Effects of DBS on the Motor Systems of Parkinson's Disease*, YOUTUBE (June 12, 2013), <https://www.youtube.com/watch?v=uBh2LxTW0s0>; Stephan Hamman, *Deep Brain Stimulation for Depression CNN Presents Special with Dr. Sanjay Gupta*, YOUTUBE (May 6, 2012), <https://www.youtube.com/watch?v=Lq5rIIcVgA>.

²¹ However, this appears to be changing. See, e.g., BlueCross BlueShield of Alabama, *Deep Brain Stimulation*, BCBSAL.ORG (Nov. 2014),

Another reason that DBS is not a first-line treatment for Parkinson's and depression is that, despite the remarkable results seen in test populations, not every case of these respective diseases manifests in a predictable way within the brain. It is well understood that Parkinson's affects the basal ganglia—part of the forebrain that, *inter alia*, controls motor functions—with primary concentration in the substantia nigra.²² The disease affects the dopaminergic system in these areas, causing the death of dopamine neurons and thus disrupting the electrical system of neuronal communication.²³ DBS targets specific areas of the basal ganglia depending on the individual patient's symptomology.²⁴ However, there are limitations to the specificity of placing an electrode in a parkinsonian brain. Clinicians may have difficulty identifying the precise location in the chain of neuronal signals where DBS will cause the appropriate amount of inhibitory function.²⁵ Even if a fairly precise location for the disturbance is identified, there is a risk that the inhibitory function of DBS will have unwanted effects in other areas of functioning. Piasecki and Jefferson identified a case where a man who received DBS for parkinsonian symptoms began to display strong cognitive side effects approximately six weeks after the initiation of treatment. The man showed up at the emergency room with

<https://www.bcbsal.org/providers/policies/final/347.pdf> (approving coverage DBS for numerous diseases).

²² C.A. Davie, *A Review of Parkinson's Disease*, 86 BRIT. MED. BULL. 109, 110 (2008).

²³ *Id.*

²⁴ See Gilberto Pizzoloato & Tomasz Mandat, *Deep Brain Stimulation for Movement Disorders*, 6 FRONTIERS IN INTEGRATIVE NEUROSCIENCE 2, 2 (2012).

²⁵ See Shelley D. Piasecki & James W. Jefferson, *Psychiatric Complications of Deep Brain Stimulation for Parkinson's Disease*, 65 J. CLINICAL PSYCHIATRY 845, 846 (2004).

severe agitation and odd thought patterns.²⁶ Most alarmingly, the man described a combination of increased goal-directed activity and sexually inappropriate behavior.²⁷ He described a desire to harm his daughter and stated, “I have never had a stronger sense of suicide.”²⁸ Similar manic-type reactions have been reported in several other studies.²⁹

Clearly, side effects such as those described above are reason for pause. Even when tempered with descriptions of the black cloud of depression suddenly lifting, from depressed patients who underwent DBS, the real possibility of danger looms large with DBS. Of course, sufferers of Parkinson’s, depression, and other ailments that demonstrate resistance to less invasive forms of treatment, might very well choose the risk of adverse side effects when the alternative is continuing misery and disability. Recognizing this fact, the FDA has approved the use of DBS under the Humanitarian Device Exemption, which allows for very limited patient populations to receive the treatment for disorders, such as OCD, where the clinical evidence would not otherwise meet FDA standards for medical use.³⁰ This special exemption by the FDA represents a recognition that DBS is a potentially power tool—in many ways, uniquely powerful—in combatting certain otherwise intractable maladies. At the same time, it represents a recognition that the

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.*

²⁹ *See id.*

³⁰ *See* United States Food and Drug Administration, *supra* note 16.

technology is still in its relative infancy in terms of researchers' and clinicians' ability to predict the outcomes of its use on a specific individual.

Transcranial Magnetic Stimulation

In response to the concerns surrounding DBS—including the invasiveness of the procedure—researchers have developed another technology, transcranial magnetic stimulation (TMS), that utilizes a similar concept. TMS is non-invasive because it makes use of magnetic signals that can penetrate the brain from the outside.³¹ A magnetic coil, sometimes contained in a helmet, is placed on the patient's head and the magnetic signal is targeted towards an area of the brain directly underneath the current. One of the drawbacks of TMS, as compared to DBS, is that the TMS signal can only penetrate approximately six centimeters into the brain.³² As such, treatment is limited to patient's who have brain abnormalities that are relegated to the outer portions of the brain—a small portion of the brain as a whole.³³ For certain migraine sufferers, TMS is effective in reducing symptoms. The visual abnormalities associated with migraines have been traced to unusual activity of the cortex.³⁴ The location of the cortex—forming the outer layer of the brain—makes it a good target for TMS.³⁵ The left prefrontal cortex has also been

³¹ GLANNON, *supra* note 3, at 138.

³² See Mark S. George et al., *Image-Guided Brain Stimulation*, in NEUROBIOLOGY OF MENTAL ILLNESS 216 (Dennis S. Charney et al., eds., 4th ed. 2013).

³³ See *id.* at 215.

³⁴ See Filippo Brighina et al., *Reduced Cerebellar Inhibition in Migraine with Aura: a TMS Study*, 8 CEREbellum 260, 260 (2009).

³⁵ See *id.*

implicated in certain forms of depression.³⁶ Thus, the FDA has given approval for the use of TMS in major depressive disorder.³⁷ Similar to the outcomes of DBS, one woman who was treated with TMS reported that she “felt completely well for the first time in 3 years.”³⁸ Nevertheless, the overall efficacy of TMS appears to be quite minimal. Several meta-studies have found that TMS has a statistically significant, but very small, effect compared to placebo.³⁹

This is in comparison to the extremely positive efficacy that has been demonstrated with DBS.⁴⁰ At the same time, the cost of TMS, while significant, at up to \$12,000 per round of treatment, is far less than the price tag of DBS—which can easily exceed \$400,000 when follow-up visits are taken into account.⁴¹ Perhaps not surprisingly, insurance companies have shown resistance to reimbursing the cost of both TMS and DBS treatment. With TMS, insurance companies have provided coverage in a few cases where clinicians deemed TMS necessary. With DBS, there are so far no cases where insurance companies have been willing to foot the enormous bill for treatment.⁴²

³⁶ Mark S. George et al., *Daily Repetitive Transcranial Magnetic Stimulation (rTMS) Improves Mood in Depression*, 6 NEUROREPORT 1853, 1854 (1995).

³⁷ Caroline Cassels, *FDA Clears TMS Device for Resistant Depression*, MEDSCAPE (Jan. 9, 2013), <http://www.medscape.com/viewarticle/777357>.

³⁸ Mark S. George et al., *supra* note 36.

³⁹ Cristina Cusin & Darin D. Dougherty, *supra* note 11, at 3.

⁴⁰ *Id.* at 6.

⁴¹ *Id.* at 2, 6.

⁴² *See id.* at 7. On the other hand, Medicare and private insurance companies will pay for DBS for treatment of Parkinson’s disease in most cases. *See generally* Andrew K. Chan

DBS and the Alteration of Personality

As described in some of the preceding case studies, DBS—along with TMS—can have profound effects on mood. In several of the noted cases, the patients reported positive feelings such as joy and renewed interest in activities. There were also several cases where DBS induced negative feelings, such as the man who reported a sudden impulse to physically harm his daughter and an overwhelming desire to commit suicide. There are other cases where use of DBS, or other technology used to stimulate the brain, has led to sudden and striking changes in personality. The protocol for implantation of the DBS electrode requires that the patient remain awake so that she can report any effects as the electrode is positioned, and so clinicians can make note of any behavioral changes.⁴³ At this point in time, clinicians cannot make perfectly accurate predictions about what behavioral and personality changes will accompany placement of the DBS electrode.⁴⁴ The following cases study illustrates the effect that DBS can have on an individual's personality.

In an early case of DBS, researchers noted that the patient's personality changed significantly after the DBS electrode was activated.⁴⁵ Prior to activation, the patient displayed the typical attitude of a person about to undergo major surgery: he was anxious and reserved, and his conversation was generally limited to concern about the procedure.

et al., *Disparities in Access to Deep Brain Stimulation Surgery for Parkinson Disease*, 71 JAMA NEUROLOGY 291 (2014).

⁴³ See Andrew Koivuniemi & Kevin Otto, *When "Altering Brain Function" Becomes "Mind Control,"* 8 FRONTIERS SYS. NEUROSCIENCE 1, 2 (2014).

⁴⁴ *Id.*

⁴⁵ This case is adapted from E.S. VALENSTEIN, *BRAIN CONTROL* (1973).

When the DBS electrode was turned on, the patient's outward display suddenly changed from anxious and reserved to euphoric and verbosely friendly. His speech was interspersed with jokes and loud, jovial, singing. The change in outward personality was repeated upon each application of DBS.

Another case of personality change involved a patient who was suffering from morbid obesity.⁴⁶ Clinicians had targeted an area of the brain that was associated with eating disorders—the anterior forix. When the DBS electrode was activated, the patient reported a vivid memory flashback. While DBS apparently had no effect on the patient's desire for food, clinicians were interested in the flashback that the patient had experienced. In a series of follow-up visits, they tested the patient's verbal memory prior to and following DBS activation. After each pulse of DBS, the patient's verbal memory showed significant improvement, leading the clinicians to recommend a study of DBS for the treatment of dementia.

A third case of personality change involved a woman who was receiving DBS for treatment of depression.⁴⁷ As indicated from previous research, clinicians implanted an electrode in the patient's subgenual cingulate gyrus. The patient did not experience relief from her depression. However, in addition to depression, the patient also suffered from severe anorexia nervosa. Following administration of DBS, the patient stopped having

⁴⁶ This case is adapted from C. Hamani et al., *Memory Enhancement Inducted by Hypothalamic/Fornix Deep Brain Stimulation*, 63 ANNALS OF NEUROLOGY 119 (2008).

⁴⁷ The case is described in M. Israel et al., *Deep Brain Stimulation in the Subgenual Cingulate Cortex for an Intractable Eating Disorder*, 67 BIOLOGICAL PSYCHIATRY e53 (2010).

symptoms of anorexia nervosa. This led the clinicians to recommend a study of DBS for treatment of anorexia nervosa.⁴⁸

One take-away from the preceding cases is that it is clear that DBS has the potential for alteration of behavior outside of a therapeutic context. Andrew Koivuniemi and Kevin Otto have warned that these kinds of cases signal that DBS could be used for purposes of mind control.⁴⁹ However, they argue that the criteria of mind control would only be satisfied if the patient did not give informed consent for the procedure.⁵⁰ Thus, if an individual requested that DBS be used to improve his verbal recall or to make him into a more jovial person, the nefarious connotations of mind control would not be a concern.⁵¹

DBS: Unintentional Behavior vs. Intentional Behavior

Without getting into the weeds of the therapy / enhancement distinction, it is possible to distinguish between the use of DBS to change unintentional behavior and intentional behavior. DBS for treatment of parkinsonian symptoms falls neatly into the former category. Parkinson's sufferers have no conscious control over the muscle spasms that cause their bodies to twist and jerk in unpredictable ways. In the treatment of Parkinson's, DBS alters a reflexive action over which the patient has no conscious control. Arguably, depression falls into this category as well. People who suffer from

⁴⁸ *Id.*

⁴⁹ See Andrew Koivuniemi & Kevin Otto, *supra* note 43.

⁵⁰ *Id.* at 5.

⁵¹ There is much more that could be said about this issue, but it is beyond the scope of my dissertation to discuss it in detail.

depression generally feel that their moods are not subject to conscious control—their moods are, in a sense, reflexive.⁵² One way to look at DBS for treatment of depression is that it changes the reflexive attitudes, moods, and behaviors associated with depression. A depression sufferer who is used to waking up every morning in a black mood and feeling no interest in daily activities might suddenly feel energized in the morning and motivated to engage in activities. To the depression sufferer and the person with Parkinson's, the relief experienced from DBS would represent escape from an intolerable existence.

Several of the case studies I mentioned in earlier paragraphs fit more neatly into the latter category. For example, consider the man who went from being anxious and reserved to jovial and outgoing. In that case, DBS affected an intentional behavior, since, at least to some degree, we have control over how outgoing we are. Or, at least, outgoingness and joviality are the kinds of traits we might ascribe to a person's will, unlike parkinsonian tremors. We might even say the same about the man who experienced increased verbal memory after receiving DBS for an eating disorder. While memory may not be under intentional control to the extent of a personality trait like outgoingness, there is still a sense in which we would ascribe *normal* memory (the same cannot be said of someone who suffers memory loss as a result of Alzheimer's) to a

⁵² Of course, depression sufferers do have conscious control over certain factors that might change their mood, e.g., exercise, eating habits, etc. However, even these factors can feel out of conscious control during major depressive episodes. See Thomas Fuchs, *Corporealized and Disembodied Minds: A Phenomenological View of the Body in Melancholia and Schizophrenia*, 12 PHIL., PSYCHIATRY, & PSYCHOL. 95, 99 (2005) (“[In depression] the body loses the lightness, fluidity, and mobility of a medium and turns into a heavy, solid body that puts up resistance to the subject's intentions and impulses.”).

person's will in a way that we would not feel comfortable in doing with parkinsonian tremors.

Both of these cases—the sudden joviality and the improved memory—represent putatively positive personality changes. But negative personality changes can fit into the category of changes to intentional behavior as well. Recall the man who received DBS for parkinsonian symptoms and had sudden and profound changes in his personality, including a desire to physically harm his daughter, increased aggression, and uncontrollable sexual impulses.⁵³ Obviously, these personality changes were not changes that the man would have decided to engage in on a conscious, intentional, level.

My purpose in making the above distinction is to highlight the fact that DBS is not relegated—even as this early point in its history—to the treatment of physical or mental disease. Rather, DBS has already demonstrated a potential to alter underlying aspects of a person's personality, at the level of both unintentional and intentional behaviors. These personality changes range from capacities associated with intelligence, for example, improved memory, to alteration of a person's core traits, for example, openness, conscientiousness, extraversion, agreeableness, and neuroticism.⁵⁴

The latter category—called the Big Five Personality Traits by psychologists—is especially interesting because these traits are generally considered to remain stable throughout adulthood.⁵⁵ Indeed, research indicates that even significant life events that

⁵³ Piaseck & Jefferson, *supra* note 25.

⁵⁴ See generally John M. Digman, *Personality Structure: Emergence of the Five-Factor Model*, 41 ANN. REV. PSYCHOL. 417 (1990).

⁵⁵ See R.R. McCRAE & P.T. COSTA, PERSONALITY IN ADULTHOOD 63 (1990).

we might assume would result in personality changes (a belief bolstered by fictional depictions of human beings, from Shakespeare to Stephanie Myers), rarely result in lasting changes in the Big Five.⁵⁶ Vito Agosti and Patrick J. McGrath have conducted a study using two selective serotonin reuptake inhibitors (SSRIs) to determine if changes in depression correlate with changes in core personality traits.⁵⁷ Agosti and McGrath used the Hamilton Rating Scale for Depression (Hamd-D) and the Temperament and Character Inventory (TCI) to determine whether a correlation exists.⁵⁸ The researchers did not find a statistically significant change in core personality traits, as measured by the TCI, in patients who experienced significant improvement in depression, as measured by the Hamd-D.⁵⁹

At this point, the evidence that DBS can cause statistically significant changes in core personality traits is largely anecdotal. There has not been a study that specifically tries to answer this question. However, the anecdotal evidence is quite strong and, thus, worth attention. Consider the difference in effect on personality between SSRIs and DBS. Although the Agosti and McGrath study turned up no statistically significant changes in core personality traits amongst SSRI users, we know from decades of reports that many people do experience relief from depression symptoms when treated with SSRIs.

⁵⁶ D. Cobb-Clark & S. Schurer, *The Stability of Big-Five Personality Traits*, 115 ECON. LETTERS 11, 15 (2012).

⁵⁷ Vito Agosti & Patrick J. McGrath, *Comparison of the Effects of Fluoxetine, Imipramine and Placebo on Personality in Atypical Depression*, 71 J. AFFECTIVE DISORDERS 113 (2002).

⁵⁸ *Id.* at 115-16.

⁵⁹ *Id.* at 118-19.

Whether or not the changes reflect statistically significant changes in core personality traits, the people who have found relief in SSRIs frequently report changes in affect and behavior. For example, it is well known that people who benefit from SSRIs sometimes experience increased sociability and renewed interest in activities. Yet, these changes are typically mild in their demonstration.⁶⁰ In other words, a person who takes an SSRI is unlikely to suddenly develop significant personality changes.⁶¹ On the other hand, the evidence with respect to DBS suggests that sudden and drastic changes in personality are possible.

There are rare cases where people receiving SSRIs suffer severe side effects, including full blown manic symptoms.⁶² These side effects might be likened to the case I described earlier where the previously withdrawn man suddenly burst into song. The difference is that with DBS, the effect can be removed instantaneously, and clinicians can continue to probe for new areas in the surrounding tissue that provide the desired effect. With SSRIs, the side effects sometimes take weeks to go away, and occasionally leave lasting psychological damage.⁶³ My point here is that the nature of DBS makes direct,

⁶⁰ See Irving Kirsch et al., *Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration*, 5 PLOS MED. 260, 265 (2008) (“Using complete datasets (including unpublished data) and a substantially larger dataset of this type than has been previously reported, we find that the overall effect of new-generation antidepressants [SSRIs] is below the recommended criteria for clinical significance.”).

⁶¹ See *id.*

⁶² See Antonella Benvenuti et al., *Treatment-Emergent Mania/Hypomania in Unipolar Patients*, 10 BIPOLAR DISORDER 726, 726 (2008).

⁶³ See Lut Tamam & Nurgul Ozpoyraz, *Selective Serotonin Reuptake Inhibitor Discontinuation Syndrome: A Review*, 19 ADVANCES IN THERAPY 17, 20 (2002).

deliberate, and specific alteration of personality traits much easier, and safer, than would be similar attempts with SSRIs.

Current Limitations of DBS and New Frontiers in Brain Science

Currently, the arguably greatest limitation of DBS is our knowledge of the function of brain structures. We have a fairly good idea of the large-scale distribution of function across the brain. For example, we know that the occipital lobes are largely responsible for visual processing; the temporal lobe is involved, *inter alia*, in language comprehension; the corpus callosum allows for communication between the two hemispheres of the brain. We also understand the roles that certain areas of the brain play in disease and dysfunction. As noted earlier, Parkinson's disease largely affects the substantia nigra region of the midbrain; Alzheimer's disease causes degeneration of neurons and synapses in the cerebral cortex—though the precise mechanism is not well understood.⁶⁴ With respect to personality disorders, a recent meta-analysis by Yaling Yang and Adrian Raine provides strong evidence that psychopathy is associated with specific regions of the prefrontal cortex—the dorsolateral and orbitofrontal regions—as well as the anterior cingulate cortex.⁶⁵ Certainly, our understanding of the brain has advanced by leaps and

⁶⁴ See Alistair Burns & Steve Iliffe, *Alzheimer's Disease*, 338 BRIT. MED. J. 467, 468-69 (2009).

⁶⁵ See Yaling Yang & Adrian Raine, *Prefrontal Structure and Functional Brain Imaging Findings in Antisocial, Violent, and Psychopathic Individuals: A Meta-Analysis*, 174 PSYCHIATRY RES. 81, 86 (2009).

bounds since the late nineteenth century, when Paul Broca first hypothesized that the brain was divided into specific regions that regulate specific functions.⁶⁶

The Current State of Brain Science and The BRAIN / HPB Initiatives

Despite the progress that researchers have made in understanding the brain, some researchers have lamented the state of brain science. Anthony Zador has argued that our understanding of very large structures in the brain, such as the four lobes of the occipital cortex, and the very small structures in the brain, such as neurons, does not provide us with much information about how the brain actually works.⁶⁷ According to Zador, what we really need is a “wiring diagram” of how the brain works at the resolution of individual neurons.⁶⁸ In 2013, President Barack Obama announced an initiative with just that goal. The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative is a \$100 million partnership between public and private institutions with the goal of using emerging technologies such as nanoscience, information technology, and biotechnology, among others, to drastically increase our understanding

⁶⁶ See ERIC R. KANDEL, JAMES H. SCHWARTZ, & THOMAS M. JESSEL, *PRINCIPLES OF NEURAL SCIENCE* 10-11 (4th ed. 2000).

⁶⁷ Anthony Zador, *The Connectome as a DNA Sequencing Problem*, in *THE FUTURE OF THE BRAIN: ESSAYS BY THE WORLD’S LEADING NEUROSCIENTISTS* 40 (Gary Marcus & Jeremy Freeman eds., 2015).

⁶⁸ *Id.* Philosophers have argued that even a precise understanding of the brain’s structure and function will not help us solve the “hard problem”—why and how we have subjective experience instead of just being zombies. See, e.g., DAVID J. CHALMERS, *THE CHARACTER OF CONSCIOUSNESS* 6-7 (2010). For my purposes this discussion is not particularly relevant. Whether we understand the nature of subjective experience does not, and has not, stopped us from altering the parameters of that subjective experience.

of the brain.⁶⁹ The hope is that a combination of advanced imaging technology and new understandings of how the brain processes information “will open new doors to explore how the brain records, processes, uses, stores, and retrieves vast quantities of information, and shed light on the complex links between brain function and behavior.”⁷⁰

The United States is not alone in its desire to better understand the brain. The announcement of Europe’s Human Brain Project (HBP) came just months before the United States’ BRAIN Initiative was revealed. Based in Geneva, the HBP is a ten-year, \$1 billion project to develop, inter alia, a complete, functional, computer model of the human brain.⁷¹ The ultimate goals of both projects may be ambitious but, like the push to get to the moon and to fully map the human genome, the hope is that a commitment to the end goal will result in many breakthroughs along the way that will lead to practical results such as new treatments for currently ill-understood brain disorders like Alzheimer’s Disease.⁷² One impetus of the initiatives is a theoretical move away from understanding brain disorders as “chemical imbalances” and towards understanding them as malfunctions in the electrical communication systems in the brain.

The first results of the BRAIN Initiative were published in May 2015. Researchers were able to control the behavior of mice using a synthetic brain messenger

⁶⁹ *Fact Sheet: BRAIN Initiative*, WHITE HOUSE OFFICE OF THE PRESS SECRETARY (Apr. 2, 2013), <https://www.whitehouse.gov/the-press-office/2013/04/02/fact-sheet-brain-initiative>.

⁷⁰ *Id.*

⁷¹ *See An Overview of the Human Brain Project*, HUMAN BRAIN PROJECT, <https://www.humanbrainproject.eu/discover/the-project/overview> (last visited Jan. 14, 2016).

⁷² *See id.*; *Fact Sheet: BRAIN Initiative*, *supra* note 69.

system that integrates with the natural neuronal system and allows individual neurons to be turned on and off.⁷³ The ability to target the behavior of individual neurons gives researchers the tools to study the effect that single neurons can have on the operation of the brain as a whole.⁷⁴ While there are obviously ethical barriers to applying this technology to human beings, there is actually good reason to think that it is a safer alternative to current drug therapies and alternative therapies such as DBS. The technology requires a genetically engineered receptor to be attached to neurons. The receptors are designed to accept only a specific synthetic chemical that must be administered by researchers. The receptor and the synthetic chemical are biologically inert on their own. Neither one, alone, causes any changes in biology or behavior. When the chemical is introduced into the body, it binds to the receptor and either turns the neuron “on” or “off” depending on the specific make-up of the receptor.

The Advantages of Synthetic Systems

There are myriad advantages to this synthetic system over a naturally occurring one. Consider dopamine-based therapies. When dopamine is introduced into the brain it binds to naturally occurring receptors. Increased dopamine leads to modulation of mood and behavior. The effect can be quite extreme. For example, in the 1970s, Oliver Sacks used dopamine to bring patients out of decades-long catatonic states.⁷⁵ However, the

⁷³ See *Souped-Up Remote Control Switches Behavior On-and-Off in Mice*, NATIONAL INSTITUTES OF HEALTH (Apr. 30, 2015), <http://www.nih.gov/news/health/apr2015/nimh-30.htm>.

⁷⁴ *Id.*

⁷⁵ See generally OLIVER SACKS, *AWAKENINGS* (1974).

limitations of dopamine-based therapies manifest themselves quite quickly. The brain's natural response to a flood of dopamine, or other chemicals that bind to specific receptors, is to cull the number of receptors.

This effect is perhaps most obvious in drug addiction. Cocaine causes dopamine to flood the brain, resulting in feelings of euphoria. The brain quickly compensates by reducing the number of dopamine receptors.⁷⁶ Thus, as one continues to ingest cocaine the effect is two-fold. First, it takes an ever-increasing amount of cocaine to achieve the same high. When dopamine enters the brain, not every molecule will bind to a dopamine receptor. Some of it will simply miss the mark. As the number of receptors is decreased, there is an increasingly low chance that any individual dopamine molecule will come in contact with a receptor.⁷⁷ Thus, it takes an ever-greater amount of dopamine to get the same cumulative effect as when there were more receptors available. Second, the lower number of receptors means that it becomes more and more difficult to feel happiness or pleasure without ingesting cocaine. And it can take months or years for the brain to create new receptors; frequently, the number of receptors never fully returns to the pre-cocaine state.⁷⁸ A similar process took place in Sacks' patients. The amount of dopamine required to bring the patients out of catatonia resulted in a massive loss of dopamine receptors,

⁷⁶ See *Drug Use Changes the Brain Over Time*, U. UTAH HEALTH SCI., <http://learn.genetics.utah.edu/content/addiction/brainchange/> (last visited Jan. 14, 2016).

⁷⁷ See *id.*

⁷⁸ See *id.*

which led to an increased need for the drug (L-DOPA) as well as number of intolerable side effects.⁷⁹

The combination of genetically engineered receptors and synthetic drugs potentially avoids the problems associated with the brain's natural system of adaptation. As described, the receptors and drugs used by researchers in the mouse study were designed to be biologically inert. In other words, their only function was to turn neurons on or off when used in conjunction. Unlike chemicals, such as dopamine and serotonin (and their associated receptors), which have a pervasive and complicated role in regulating not only the brain but also the entire body, the biologically inert receptors and chemicals designed by researchers had only one function. As inert substances, the body's system of adaptation is not called into action when the receptors are diminished or a flood of synthetic chemical is introduced. Thus, even if the receptors were reduced or the drug was not introduced, there would be no adverse effect on the brain. This idea can be difficult to accept since we are so used to thinking of the brain—and the body as a whole—as an interconnected system where changes in one system, say, an introduction of SSRIs into the brain, can have a substantive effect on other bodily systems. Of course, the entire purpose of turning individual neurons on or off is to have an effect on the brain. Presumably, SSRIs indirectly cause neurons to turn on or off as well. The difference is that the genetically engineered receptors and synthetic chemical exist entirely outside the interconnected system of response and adaptation within the body. Thus, unlike traditional drug therapies that, in a sense, attempt to “trick” the brain's natural systems into causing a change in the system, genetically engineered receptors and synthetic

⁷⁹ See SACKS, *supra* note 75.

chemicals ignore the brain's natural systems altogether. Instead of asking a complex and highly non-specific chemical such as serotonin to do the work, these artificial systems talk directly to neurons.

These artificial systems also have advantages over DBS therapies (at least, in their current state). Using DBS, clinicians are able to target fairly specific areas of the brain. Some DBS electrodes are as small as one millimeter in diameter. Yet, consider that one cubic millimeter of cerebral cortex contains, on average, 50,000 neurons.⁸⁰ Thus, even the most well-placed electrode will implicate an enormous number of neurons, some of which are surely not involved in the behavior or mood that is being targeted by DBS. Recall the man who received DBS that successfully reduced his parkinsonian symptoms but had the unfortunate side effect of causing him to feel an urge to harm his daughter and an overwhelming desire to commit suicide. In that case, clinicians tried several methods to reduce the unwanted side effects. They adjusted the position of the electrode by a small amount and they lowered the voltage of the signal from the pacemaker. The latter adjustment worked to eliminate the side effects, but only at the cost of also eliminating the intended effect.⁸¹ The genetically engineered receptors and synthetic chemicals designed by the scientists in the “remote control mouse” study are able to target individual neurons—meaning that the potential therapeutic resolution would be improved by 50,000 times. If the clinicians in that case had access to this synthetic chemical technology, they would have been able to switch individual neurons on or off

⁸⁰ RODNEY DOUGLAS, HENRY MARKRAM, & KEVAN MARTIN, NEOCORTEX 500 (2004).

⁸¹ *Id.* at 846.

until they determined precisely which neurons were implicated in the positive or negative effects.

Targeting Specific Brains

This brings up another issue with the current state of therapies targeted directly at the brain. As noted, we currently have a decent understanding of the large-scale topography of the brain. Yet, even at relatively large scales there can be significant differences between individuals.⁸² For example, the auditory processing center is fairly localized within the brain—yet the specific areas of the auditory cortex that allow for specific functions having to do with auditory processing may be distributed more heavily in a particular area in one person’s brain than in another.⁸³

Even aside from genetic factors that influence an individual brain’s specific topology, the brain is, to a certain extent, elastic. Thus, if one area of the brain is damaged other areas of the brain can take over the function of the damaged area.⁸⁴ In other words, researchers have found that the classical conception that specific areas of the brain are dedicated to specific functions is not wholly true. Moriel Zelkowsky et al. illustrated this fact by studying individuals who had severe damage to their dorsal hippocampus—the area associated with Pavlovian fear response.⁸⁵ If a Pavlovian

⁸² See GLANNON, *supra* note 3, at 141.

⁸³ See generally Christiana M. Leonard et al., *Normal Variation in the Frequency and Location of Human Auditory Cortex Landmarks. Heschl’s Gyrus: Where Is It?*, 8 CEREBRAL CORTEX 397 (1998).

⁸⁴ See GLANNON, *supra* note 3, at 61.

⁸⁵ Moriel Zelikowsky et al., *Prefrontal Microcircuit Underlies Contextual Learning After Hippocampal Loss*, 110 PROC. NAT’L ACAD. SCI. U.S. 9938 (2013).

response is extinguished it can be renewed in less time than it took to develop the response in the first place, because the brain circuits underlying the response maintain their “memory,” in a sense.⁸⁶ Zelkowski et al. found that the response could be renewed even in patients who had severe damage to the dorsal hippocampus. They discovered that alternative brain structures were recruited to compensate for the loss of function in the dorsal hippocampus.⁸⁷ This is just one small example of how the brain is far more elastic than has been thought in previous decades.⁸⁸

Because of the potential for large variation between individual brains, clinicians typically make use of MRI and other brain imaging technologies to develop a picture of a specific patient’s brain before treatment with DBS.⁸⁹ Yet, brain-imaging technologies are currently limited in a number of ways. Perhaps most significantly, not enough brain scans have been performed. Without a large sample of brain images from the general population it is impossible for clinicians to make accurate predictions about the correlation between a specific person’s brain and the likelihood that aspects of the scan

⁸⁶ See *id.* at 9941 (“Following extensive . . . damage, the brain can invoke alternate circuitry to compensate . . . , restoring some of the essential elements of context-sensitive learning and memory.”).

⁸⁷ *Id.* at 9942.

⁸⁸ See generally Alvaro Pascual-Leone et al., *The Plastic Human Brain Cortex*, 28 ANN. REV. NEUROSCIENCE 377 (2005).

⁸⁹ See, e.g., Rajesh Pahwa, Kelly E. Lyons, & Jules M. Nazzaro, *Deep Brain Stimulation for Parkinson’s Disease*, U. KAN. MED. CTR. 14, <http://www.kumed.com/~media/Imported/kumed/parkinson/dbs.ashx> (last visited Jan. 16, 2016).

represent pathology.⁹⁰ Consider the preceding paragraph where I discussed the elasticity of the brain. If clinicians were to find a patient with a specific kind of activity in a region of her brain, and DBS targeted to that area resulted in a cessation of symptoms, they might be tempted to think that another patient with the same localized brain activity would receive similar benefits. Yet, the elastic nature of the brain means that the same area of the brain in two individuals could be responsible for wildly different functions. Until we have a large sample of brain scans, clinicians will not be able to make statistically significant predictions.

Thankfully, obtaining a large sample of brain scans is a relatively easy goal to achieve. There are no theoretical hurdles to overcome; the only question is one of resources. And one of the BRAIN Initiative's goals is to scan more brains in order to create a database of statistically significant correlations.⁹¹ In addition to scanning *more* brains, the BRAIN Initiative hopes to develop better methods to scan the brain. Awards for funding in this area have already been handed out. As just one example, researchers out of the California Institute of Technology (Caltech) have received close to \$1 million to study a brain mapping technology that would represent an enormous step forward in our understanding of how individual neuronal firings influence the operation of the

⁹⁰ See Ed Bullmore & Paul Fletcher, *The Eye's Mind: Brain Mapping and Psychiatry*, 182 BRIT. J. PSYCHIATRY 381, 381-82 (2003).

⁹¹ See Karen Lazo, *Huge Brain Database is Revealing Secrets of the Mind*, LIVE SCIENCE (July 10, 2015), <http://www.livescience.com/51523-massive-brain-scan-database-goes-online.html>.

brain.⁹² The method is related to that used in the “remote mouse” study already described. The idea is that neurons will be genetically engineered to emit an artificial molecule (an emitter molecule) when they send out signals, and a different kind of molecule (a receptor molecule) when they receive signals. Thus, when one neuron communicates with another neuron, both artificial molecules will be released. The emitter and receptor molecules “fit” together and when they come together they express a fluorescent protein. Using high-resolution imaging, these fluorescent bursts can be monitored, allowing researchers to track the movement of signals throughout the brain at the resolution of two neurons.⁹³

This technology could lead to highly precise, context-specific, brain maps for individuals. Currently, if a person goes under an MRI, the technology can only monitor blood flow to somewhat specific regions of the brain.⁹⁴ While such images give some insight into how the brain responds to stimuli, compared to the technology being developed by the Caltech researchers it is like images taken of Pluto by earth-based satellites compared to the fly-by images generated by the New Horizons Orbiter. The new technology will allow researchers to see how individual neurons light up in response to stimuli *and* trace the enormously complex map of signals sent and received by neurons throughout the brain.

⁹² See *NIH BRAIN Awards*, NATIONAL INSTITUTES OF HEALTH, Project Title: Tracing Brain Circuits by Transneuronal Control of Transcription, <http://braininitiative.nih.gov/nih-brain-awards.htm> (last visited Oct. 13, 2015).

⁹³ *Id.*

⁹⁴ See GLANNON, *supra* note 3, at 46.

What It All Means

So far in this chapter, I have described the current state of brain science—including therapies directed at diseases of the brain—as well as up-and-coming advances. As I have described, current technologies allow clinicians to identify fairly specific regions of the brain and to alter behavior by targeting those regions with electrical stimulation. The United States’ BRAIN Initiative, and Europe’s HBP, has already produced research that has the potential to exponentially increase our ability to map the brain and to treat the causes of brain disease at the neuronal level. Of course, along with the possibility of treatment comes the possibility for human enhancement.

Some such technologies are already in the works. Even without the advances being produced by the BRAIN Initiative, researchers have developed technologies that allow subjects to control robotic arms using their minds, in other words, by willing the movement of the robotic arm, which is wired to the subject’s nerve endings.⁹⁵ While this technology allows a human mind to reach out to control a robotic arm, researchers have also developed technologies that allow the “sensations” experienced by a robotic hand to reach into a human subject’s mind.⁹⁶ It’s not difficult to imagine therapeutic, as well as enhancement, purposes for a combination of the two technologies. Paraplegics would no longer be limited to lifeless artificial limbs; instead, they might be able to have cybernetic limbs that are fully wired to their brains—making the artificial limbs nearly indistinguishable from the real thing.

⁹⁵ See Zackary Canepari, Drea Cooper, & Emma Cott, *Prosthetic Limbs, Controlled by Thought*, N.Y. TIMES (May 20, 2015), <http://www.nytimes.com/2015/05/21/technology/a-bionic-approach-to-prosthetics-controlled-by-thought.html>.

⁹⁶ *Id.*

With these developments in brain science, the near future also holds the possibility of enhancing aspects of our personalities. My concern in this dissertation is with empathy; so let me focus on that aspect of personality. The first step in understanding how empathy works in the brain would be to develop a large sample of individual brain maps during times when people are experiencing feelings of empathy. Using the new brain mapping technology developing through the BRAIN Initiative, it will be possible to see specifically what neuronal circuits are at work when we feel empathy. From there, a technology similar to the “remote mouse” combination of genetically engineered receptors and synthetic drugs could be used to trigger these circuits. Alternatively, more sensitive DBS could be used to trigger a similar response without the need for introducing synthetic biological devices into the brain. Obviously, there are all kinds of ethical issues that surround the implementation of a technology that could so radically change a person’s personality. For example, if the “remote mouse” technology were used, it would likely require parents to decide to have the genetically engineered receptors integrated into their children’s DNA at a very early stage of development. Furthermore, the specific synthetic drugs used in that technology only last for approximately one hour,⁹⁷ meaning that anyone who wanted the continuing benefit of enhanced empathy would be dependent on continual doses of the synthetic drug.

However, my purpose here is not to defend or admonish the technology that might give rise to enhanced empathy. I have only described how such technologies are likely to be available in the not-too-distant future. It is my belief that if these technologies are shown to be safe and effective, as well as reversible, they will be embraced. Regardless

⁹⁷ See *Souped-Up Remote Control Switches Behavior On-and-Off in Mice*, *supra* note 73.

of my own beliefs, it is important to consider the ethical implications not only of the technology (which ethicists have debated for decades), but of the changes in how we might evaluate people who are the recipients of such technology. That is my project for the rest of this dissertation.

The World of the Enhanced Empath

Imagine the future, not too long from now. Revolutions in our understanding of the brain have led to dramatic reductions in brain disorders such as Alzheimer's and Parkinson's, as well as dramatic reductions in mental disorders such as depression and anxiety. Treatment regimes for brain / mind disorders no longer involve a individual trial-and-error approach where patients are forced to make the agonizing choice between barely tolerable side effects of some medication and the mild relief from a devastating disorder. Instead, clinicians are able to map individual patient's brains at the resolution of single neurons. Additionally, a combination of DBS and biologically inert synthetic receptors and drugs, integrated into the brain's natural biology, allows for adjustment of brain signals at the level of individual neurons.

Along with the miraculous therapeutic benefits that accompany these new technologies, there is the possibility of enhancing areas of the brain beyond their natural states. Certain ethicists have urged that these technologies herald the possibility for a new evolution in how humans relate to one another.⁹⁸ Human beings finally have the

⁹⁸ See generally JOHN HARRIS, ENHANCING EVOLUTION: THE ETHICAL CASE FOR MAKING BETTER PEOPLE 19-35 (2010) (Chapter title: *Enhancement is a Moral Duty*); NICHOLAS AGAR, LIBERAL EUGENICS: IN DEFENSE OF HUMAN ENHANCEMENT (2004); Julian Savulescu, *Procreative Beneficence: Why We Should Select the Best Children*, 15 BIOETHICS 413 (2001).

opportunity to escape the confines of natural selection that have blindly driven human development for millennia. According to these ethicists, if human beings wish to achieve truly lasting change, not just in one or a few generations but in the species as a whole, it is necessary—even morally obligatory—to use the technological means at our disposal to alter our brains so as to increase our sense of ECE towards our fellow human beings.⁹⁹ Doing so will short-circuit (so to speak) the natural human tendency to ignore the suffering of others in far-off places. Finally, as I argued in Chapter Two, we may be able to achieve the goal of a fully expanded circle of care and concern for our fellows.

Conclusion

Increased ECE will allow people to come to terms with their naturally ingrained tendencies that place small personal pleasures and benefits over the potential for enormous harm reduction for people all over the planet. As Descartes wrote in the *Discourse on Method*: “the mind is so dependent on the temperament and on the arrangements of the organs of the body that, if it is possible to find some means that generally renders men more wise and more capable than they have been up to now, I believe that we must seek for it[.]”¹⁰⁰ Given that we have now discovered the means, it is

⁹⁹ Julian Savulescu lists empathy—alongside imagination, sympathy, fairness, and honesty—as one of the “all purpose goods,” i.e., moral capabilities whose enhancement would be good in nearly every circumstance. See Julian Savulescu, *Justice, Fairness, and Enhancement*, 1093 ANNALS N.Y. ACAD. SCI. 321, 333 (2006). See also ALLEN BUCHANAN ET AL., FROM CHANCE TO CHOICE (2000) (“[G]eneral-purpose means are capabilities that are broadly valuable across a wide array of life plans and opportunities pursued in a society like our own.”). I will explain why ECE but not RCE would likely be enhanced in the next chapter.

¹⁰⁰ RENÉ DESCARTES, DISCOURSE ON METHOD 49 (Richard Kennington trans., Focus Publications 2007) (1637).

possible to achieve Descartes' dream of overcoming the limitations placed on human nature *by* nature.

Let us assume for a moment that a certain portion of the population chooses to artificially enhance their ECE. They feel that there is a general moral duty to take advantage of ways to become a better person. In the past, this moral duty might have manifested in listening to the teachings of religious leaders or ethicists who talked about the importance of helping the less well-off among us. They also feel that the easier some means of bettering oneself is, the stronger the moral duty to take advantage of the opportunity. With the technological capabilities of the new brain science, enhancing one's sense of ECE is quick, easy, and devoid of side effects (and given the direct neuronal manipulation involved, entirely reversible if some unknown side effect were to occur). Thus, these moral pioneers feel compelled to receive the enhancement.

Quite suddenly, there is a new breed of creature walking the planet. These beings—human in genetic and cultural heritage—would be different from the rest of us in important ways. Significant ethical questions would be raised by the presence of these enhanced empaths. In which specific ways would they be different from ordinary humans? For example, would they have free will? And how would we assess these enhanced empaths from moral and legal standpoints? I will attempt to answer these questions in the next chapter.

Chapter Six

ENHANCED EMPATHY

MORAL AND LEGAL RESPONSIBILITY

Introduction

In the preceding chapter I gave an overview of the current state of brain / mind science and our current ability to manipulate emotion and personality. I presented evidence to bolster my assertion that the ability to artificially enhance empathy is not far off. In this chapter I turn to the enhanced empath herself. Specifically, I am interested in how the enhanced empath's actions might be judged from moral and legal perspectives. In order to address this issue, I begin with a discussion of the status of the enhanced empath's free will. I then turn to an analysis of duty as it relates to moral and legal obligations, placing the enhanced empath within that framework.

Part I: Is the Enhanced Empath Free?

One of the threshold questions with respect to the enhanced empath is whether her actions can be seen as the result of free will. There is undoubtedly something about artificial enhancement that makes ascriptions of freedom and responsibility *seem* trickier. Recalling my discussion of free will from Chapter Four, there is an intuitive folk appeal to the notion that free will is predicated on the actor being in *full control* of her decision-making process. However, as I emphasized in that same chapter, this folk understanding of free will does not stand up to intellectual scrutiny. After all, none of us are in control of the specific biological processes or neuronal firings that influence (or even dictate) our

behavior. Yet, few save the adamant incompatibilist would hold that such control is necessary for free will. The question becomes: is the enhanced empath's situation, with respect to the underlying mechanisms of her behavior, so different from our own?

The Essential Freedom of the Natural

Obviously, if the enhanced empath's brain were being remotely controlled by another person there would be substantial concern that her will was not her own. One might argue that the enhanced empath's will would not be free even in a case where she was not under the direct control of another person but where she had chosen to receive a procedure that overrides her "natural" responses. Note that this argument is distinct from the argument for *the essential good of the natural* from Chapter One. The argument here holds not that everything natural is for the best but that only natural mechanisms can be considered *one's own* in the sense that one can take ownership of them, in other words, point to them and say, "yes, that is me—that is mine." We might call this argument *the essential freedom of the natural*. As the argument goes, when an artificial element is introduced into the system, one can no longer claim full ownership of the mechanisms behind one's actions.

The problem with this argument is that it ignores the fact that people frequently use "artificial" means to alter the "natural" mechanisms that underlie their behavior. For example, millions of Americans use anti-depressant chemicals to influence their moods and behaviors. Most of these anti-depressants work by blocking the reuptake of serotonin.¹ Each drug uses a patented synthetic chemical to interfere with the brain's

¹ This is the hypothesis, anyway. In truth, researchers do not understand the effect that SSRIs have on the neurobiology of the brain. See Stefanie C. Altieri et al., *Serotonergic*

“natural” response to serotonin. The entire purpose of these drugs is to change mood and behavior. Even if we concede that these drugs are “therapy” rather than enhancement, in the sense that they bring the brain into conformity with some baseline standard of how brains are supposed to work, this doesn’t change the fact that an artificial element is being introduced into the body and influencing the person’s will. If it’s true that only that which is natural can be free, then we would have to accept that anti-depressants—along with every other mood and behavior altering drug—rob people of their free will.²

A proponent of the *essential freedom of the natural* might try to avoid this argument ad absurdum by arguing that there is a threshold at which artificial elements overwhelm free will. Accordingly, drugs like anti-depressants do not sufficiently influence behavior such that a person under their influence is somehow less free. On the other hand, when we consider a scenario like the one I presented in Chapter Five—with a complicated array of genetically engineered receptors, synthetic drugs, and electrical impulses—we have crossed the threshold to where there is intuitive appeal to the notion that free will has somehow been compromised. I see two problems with this line of thinking. First, it isn’t really an argument. Instead, it is a kind of rule-like intuition heuristic. Without more, it suggests that we can rely on our intuition to evaluate a

Pathways in Depression, in NEUROBIOLOGY OF DEPRESSION 156 (Francisco Lopez-Munoz & Cecilio Alamo eds., 2011).

² Alternatively, one might argue that mood (or personality) altering drugs (or medical procedures) rob people of their authenticity, if not their will. For example, imagine a husband tells his wife that he no longer loves her, then goes in for a round of electroshock therapy only to return home and tell his wife that he does, in fact, still love her. The wife might complain that the husband’s sudden change of heart is not authentic. For a discussion of authenticity in the context of mood enhancing drugs, see Ron Berghmans et al., *Scientific, Ethical, and Social Issues in Mood Enhancement*, in ENHANCING HUMAN CAPABILITIES 158 (Julian Savulescu, Ruud ter Meulen, & Guy Kahane eds., 2011).

complex question of free will. Second, and more significantly, it begs the question: just what is it that defines a person's will as free?

What Makes the Enhanced Empath Free

In Chapter Four I presented Fischer and Ravizza's argument for moral responsibility, which includes the requirement that an agent have guidance control, in other words, free will. For Fischer and Ravizza, if an agent is capable of deliberating about a course of action and she does not feel coerced by some irresistible impulse, she has free will. I argued that the psychopath has free will because she feels that her will is her own,³ in other words, that she is not under the influence of an irresistible force or impulse. Moreover, I argued that the psychopath's lack of ECE does not render her lacking in free will (and, thus, in moral and legal responsibility) because she is still able to imaginatively project herself into another's place and realize that she would not want to be harmed if she were in the other's shoes; from there, the psychopath is able to form conscious, uncoerced, intentions about how to act. We can apply a similar argument, with a bit of a twist, in the context of the enhanced empath.

Case Study: Enhanced Empathy as a Means of Aligning First and Second-Order Desires

Let us begin with an easy case. Lisa is an adult who voluntarily decides to have her capacity for empathy enhanced. Prior to undergoing the procedure, she is already a fairly empathic person. Her RCE is intact, and she frequently thinks carefully about the effect

³ One might wonder about the case of a person who is brainwashed. If an agent's mind is altered such that she suddenly desires to engage in entirely different kinds of acts but, due to the brainwashing, she still feels able to deliberate and does not feel coerced, is her will free? I will address this question in some detail later on in this chapter.

her actions will have on others. However, while she feels the pull of ECE in many situations, she is unable to emotionally connect with the plight of certain people in particular circumstances. For example, she is unable to feel ECE with malnourished children in the Congo. She is not sure why she has this problem; after all, she recognizes through RCE that the plight of Congolese children is no different from that of the malnourished children from her hometown, to whom she feels a strong emotionally empathic connection, and she can imaginatively project herself into their place and realize that she would want someone to help her. Because of this recognition, she donates some time and money to the Congolese children, but not as much as she would if she felt motivated by ECE.

Lisa believes that enhancing her ECE will allow her to feel more motivated to act in accordance with her rational beliefs. In this sense, prior to undergoing the procedure, her first-order desires—what she feels motivated to do—do not align with her second-order desires—what she feels that she *ought* to feel motivated to do, and does in fact feel rationally motivated to do. Lisa’s hope is that undergoing the enhancement procedure will allow her to align her first-order desires with her second-order desires.

Lisa undergoes the empathy enhancement procedure and is delighted to find that she feels a strong emotionally empathic connection with people to whom she previously felt only a rationally empathic connection. This emotional connection makes it easier for her to feel motivated to give time and money to the causes she rationally believes to be worthwhile. Lisa recognizes that her newfound ECE is a product of an artificial process, but she also identifies with the process in a way that she never identified with the “natural” process that made it difficult for her to feel ECE. She never feels like her will is

being manipulated by an outside force or a force over which she has no control. As such, Lisa has guidance control and her will is free.

To argue otherwise, one would have to rely on a strange notion of coercion—that Lisa’s *biology* is being coerced, and that that coercion translates to her will because biology underlies decision-making. This view would strip Lisa of her agency. It would imply that what she believes and feels to be her own free choices are, in fact, not her own—that she is being hoodwinked by a biological trick.⁴ Recall the case of Alex from Chapter Four. Under the influence of a brain tumor, Alex developed compulsions to view child pornography. Also recall my conclusion that Alex’s will was not his own while he was acting under the influence of the brain tumor. One could try to argue that the same logic applies here: Lisa’s will is not her own as long as she is under the influence of the empathy enhancement procedure. There is a dis-analogy between the two cases, however. Alex never identified with his compulsions. Indeed, he felt that he was suddenly being bombarded by an outside force that took over his first-order desires. Yet he never lost his second-order desire, which was to *not* look at child pornography. This is in direct contradiction to Lisa’s case. Where Alex’s “outside force” (the brain tumor) moved his first-order desires away from his second-order desires, Lisa’s “outside force” (the empathy enhancement procedure) did the opposite.

⁴ For my purposes, I assume that if one is acting according to one’s free will, then one is acting authentically. See Ron Berghmans et al., *supra* note 2 (“A person might take such a drug voluntarily, fully aware of its personality-changing effects. If so, concern about alienation [authenticity] becomes less acute; there is no reason why he should be prevented from choosing to fashion himself into a different person and instead be encouraged to preserve his old personality.”).

One might accept the conclusion of the preceding paragraph and still believe that Lisa is being coerced through her biology, even though this argument is difficult to sustain once one has accepted that Lisa's enhanced ECE is in line with her second-order desires. Considered briefly, the argument would be that Lisa's biology is being directed against the course it would take if left on its own. Thus, even if the outcome corresponds to Lisa's second-order desires, her will cannot be said to be free. Again, this argument runs up against the *essential freedom of the natural*. If we are concerned with freedom of will, what does biology have to do with it? The proponent of the argument might point to certain cases of biological intervention that might, at first glance, seem to result in a loss of freedom of will. I consider one such case below.

Case Study Two: Coerced Empathy Enhancement

Consider a scenario where a person, Jill, is forced to undergo the empathy enhancement procedure. Imagine that prior to the procedure, Jill does not have a second-order desire to increase her ECE. In fact, Jill believes that there is something valuable in feeling a stronger emotionally empathic attachment to people in her relative inner circle. She believes that this in-group preference is valuable and that giving it up would be giving up something uniquely human. However, following the procedure Jill's enhanced sense of ECE feeds back into her second-order desire and leads her to believe that her previous belief was wrong. She is now glad that she had the procedure. While her previous second-order desire would not have been in line with her newly enhanced sense of emotional empathy, her newly formed second-order desire is in such alignment. In her new state, Jill would not choose to revert back to her pre-enhancement self. She still feels that it was wrong to force the procedure on her, and she would oppose the procedure

being forced on others, but she is now an advocate for voluntary adoption of the procedure.

This is a difficult case. Jill currently feels that her will is her own, and she does not feel that she is subject to an irresistible force. In this sense, Jill's will appears to be free. Yet, there is still something troubling about the case. The concern might be that there is something pernicious about ascribing free will to someone who has had her mind altered against her will. If we were willing to ascribe free will in these cases, it would be just a small step to justifying forced enhancement procedures.

I think the correct response to this situation is simply to admit that Jill's post-enhancement will is, in fact, free, while acknowledging that forcing her to undergo the procedure was wrong. In other words, we can accept that Jill still has free will without having to condone forced enhancement. As an analogy, consider a woman who is raped and becomes pregnant. After giving birth the woman loves and cares for the child and is glad that the child was born. Clearly, none of these things—loving the child or being glad that the child was born—justify the rape. We can accept that the woman's feelings of love for the child are genuine and that the child makes her happy without having to abandon our condemnation of the rape.⁵ Similarly, the fact that a person is glad to have an enhanced sense of empathy after undergoing a forced enhancement procedure does not justify forced enhancement.

⁵ Thanks to Elizabeth Brake for suggesting this analogy.

When Coerced Empathy Enhancement May be Justified

Perhaps the most significant problem with forced enhancement is that the justification for it would inevitably have to rely on a determination, by someone other than the individual undergoing the procedure, of what the most desirable outcomes are for society. I need not linger on the dangers of such an enterprise. Aldous Huxley, among others, has covered that ground.⁶ Suffice to say that allowing governments to engage in forced mind control is not likely to be a prospect that any reasonable person would accept. However, draconian worries aside, there may be limited situations where forced enhancement might be viewed as acceptable.

Consider the psychopath. Recall from Chapter Three that psychopaths make up approximately 12% of the prison population. In certain institutions, the percentage is even higher. For example, in some states judges are allowed to commit particularly dangerous sex offenders to indefinite terms of civil confinement after they have served their criminal sentences.⁷ In these civil institutions, the rate of psychopathy is significantly higher than in the general prison population.⁸

⁶ See generally ALDOUS HUXLEY, *BRAVE NEW WORLD* (1931).

⁷ Twenty states currently allow for indefinite civil confinement of sex offenders. See Adam Deming, *Sex Offender Commitment Programs: Current Practices, Characteristics, and Resident Demographics*, 36 J. PSYCHIATRY & L. 439, 441 (2008).

⁸ For example, a 2010 study found that 25.8% of civilly committed men in Wisconsin qualified as psychopaths based on their scores on the PCL-R. See Richard W. Eldwood, Dennis D. Doren, & David Thornton, *Diagnostic and Risk Profiles of Men Detained Under Wisconsin's Sexually Violent Person Law*, 54 INT'L J. OFFENDER THERAPY & COMP. CRIMINOLOGY 187, 193 (2010).

Psychopathy is a good predictor of recidivism, and Hare's PCL-R is used specifically to help parole boards in deciding whether to release a particular offender.⁹ In Minnesota, which has a robust system of indefinite civil confinement for certain sex offenders,¹⁰ not a single offender has been released because parole boards have not found any individuals to be sufficiently rehabilitated and not at risk of reoffending.¹¹ It may very well be the case that some of these psychopathic individuals would voluntarily undergo an empathy enhancement procedure. As I noted in Chapter Four, at least some psychopaths—for example, Ted Bundy—experience mental anguish and distress as a result of their inability to experience ECE. These individuals would likely jump at the chance to become “normal” human beings. At the same time, it would be naïve to think that all psychopaths would want to undergo an empathy enhancement procedure.¹² But what if such a procedure were offered to them as an alternative to remaining behind bars? Surely, a large number would opt for freedom and accept the procedure. The ethics of

⁹ See Robert D. Hare, *Forward*, in *THE CLINICAL AND FORENSIC ASSESSMENT OF PSYCHOPATHY* 3 (C.B. Gacono ed., 2d ed. 2015) (“Some commenters assert that a high score on the PCL-R is *the* primary reason for denying parole[.]”).

¹⁰ In 2015 a federal district judge found Minnesota's system to be unconstitutional. *Karsjens v. Jesson*, No. 11-3659, 2015 WL 3755870 (D. Minn. June 15, 2015).

¹¹ David Post, *Minnesota's Egregious Sex Offender Confinement Statute Held Unconstitutional*, VOLOKH CONSPIRACY (June 18, 2015), <https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/06/18/minnesotas-egregious-sex-offender-confinement-statute-held-unconstitutional/>.

¹² There are surely psychopaths who have both first and second order desires to harm others. Due to this alignment, these individuals would be unlikely to voluntarily choose to undergo the empathy enhancement procedure. This raises the question of whether a person who is forced to undergo the procedure could be considered free. I address this question in brief below. However, my primary interest in this chapter is with people who voluntarily undergo empathy enhancement. Thus, I do not address every facet of the issue with forced empathy enhancement.

forced empathy enhancement in this situation are far less troubling than the idea of a generally mandated forced enhancement regiment. The concerns that attend the latter prospect do not apply—or, at least, apply to a far smaller degree—in the former case. In the case of a generally mandated forced enhancement program, the gravest concern is that the government would have to rely on some set of norms that it determined were superior to the views held by whoever in the population believed that enhancement was not in the best interests of individuals or society overall. It is difficult to imagine that there would be no reasonable disagreement as to the wisdom of enhancement. Even as someone who believes that enhancement (of some form) may well become a necessary (or even morally required) path if human beings are to thrive in the future, I realize that there are strong arguments to the contrary. It would unquestionably be barbaric for me, or anyone else, to enforce my will on someone who disagreed with me about the wisdom of enhancement. In the case of violent psychopaths, however, this argument holds far less water.

Enhanced Empathy as an Alternative to Incarceration

It hardly seems like hubris to state that the world would be a better place for everyone if people who harbored violent tendencies had enough empathy to stop them from acting out their fantasies.¹³ It is hard to imagine a reasoned argument that it is better for society to have violent psychopaths wandering the streets. One could argue that this is a straw man because we already have an alternative: lock them up indefinitely. However, this argument has a paternalistic bent to it that says to the psychopath: “we believe that you

¹³ Of course, it might also be possible to use our newly emerging knowledge of the brain to disrupt or eliminate the violent impulses themselves, rather than using enhanced empathy as a stopgap measure to stop people from acting on those impulses.

are too dangerous to be walking the streets, so we're going to lock you up forever. We have a technology that could allow make you less dangerous and able to walk free, but we think it would be dehumanizing to give you the choice of using the technology in order to walk free when, if you were making the choice freely, you would not otherwise choose to use the technology.”¹⁴

We can find analogies to the choice between confinement and empathy enhancement in the use of psychotropic drugs in conditions of release.¹⁵ Typically, when a parole board decides that a person is capable of being released back into society, they place a number of conditions on the person's release. Common conditions include abstaining from illegal drugs and alcohol, and not associating with other criminals.¹⁶ For people who suffer from psychiatric disorders such as schizophrenia, there may be additional conditions that require individuals to stay on their prescribed medications and to make regular visits to a psychiatrist. If the person refuses these conditions, she can be denied parole; and if she violates the conditions after release she can be sent back to prison.¹⁷

¹⁴ Jeffrie Murphy addresses this same point when he states, “What I find bizarre . . . is the claim that . . . leaving the inmate to languish in some terrible place because of some terrible condition somehow preserves respect for his rights, dignity and autonomy.” JEFFRIE MURPHY, *RETRIBUTION, JUSTICE, AND THERAPY* 189 (1979).

¹⁵ Another example would be coerced chemical castration of certain sex offenders.

¹⁶ *See generally* BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, *PROBATION AND PAROLE IN THE UNITED STATES, 2014* (2014), <http://www.bjs.gov/content/pub/pdf/ppus14.pdf>.

¹⁷ *Id.*

Many of the psychotropic drugs that parolees are required to take have an enormous influence on mood, personality, and behavior. And those are just the “positive” effects. The side effects of psychotropic drugs—especially those drugs that are used to treat schizophrenia—are notorious. Frequently, patients who take medication for schizophrenia or manic depression feel that the drugs take away their true personality and leave them as a shell—a complacent shell, but a shell nonetheless.¹⁸ On the flip side, some patients who are forced to take medication find that it does them enormous good and with their new, clear, perspective they realize that their old feelings and behaviors were the result of disease.¹⁹ These patients (who might become parolees) might end up thanking the very system that forced them to take drugs that their previous selves did not want.

In general, there is not widespread horror and opposition to the idea of presenting potential parolees with the dilemma of choosing between prison and psychotropic medication.²⁰ So, why should there be greater opposition to offering potential parolees

¹⁸ See *Schizophrenia Therapy*, THRIVEWORKS, <https://thriveworks.com/austin-schizophrenia/> (last visited Jan. 18, 2016).

¹⁹ See *People’s Experiences with Involuntary Treatment*, SCHIZOPHRENIA.COM, <http://schizophrenia.com/family/involuntaryb.htm> (last visited Jan. 19, 2016).

²⁰ Some might argue that all therapeutic interventions in the context of prisons are inherently coercive because the promise of release presents the patient (i.e., the prisoner) with a dilemma: either go along with a therapy that she would not, in other circumstances, agree to, or spend more time behind bars. Jeffrie Murphy has addressed this problem by arguing that whether a therapeutic intervention is coercive or not depends on how it is presented to a prisoner. See MURPHY, *supra* note 14, at 188. Murphy presents two cases. In the first case, a dangerous prisoner, Jones, is told that he will be released only if he consents to psychosurgery. In the second case, Jones is told that he can only be released if he is found to no longer be dangerous, and psychosurgery has the best chance of securing that outcome. See *id.* According to Murphy, the difference between the two cases lies in the reason for the prisoner’s release. In the first case, Jones is released

the choice between prison and empathy enhancement? If anything, empathy enhancement would seem to be a better option than psychotropic medication. After all, the enhancement procedures that I described in Chapter Five would be entirely without side effects. The only effect would be an increase in ECE.²¹ If the patient came out on the other side of the procedure feeling that her previous feelings and perspective on the world were entirely a result of her lack of empathy, then so be it.²²

because he consented to surgery. In the second case, Jones is released because he is no longer dangerous. Murphy likens the second case to other conditions of parole such as vocational programs and checking in with a parole officer. *See id.* at 189. Many parolees surely resent having to adhere to such conditions of release, but the purpose of these conditions is to ensure that parolees potential dangerousness is kept in check, and surely we would not want to hold that conditions of release are unfairly coercive simply because parolees must adhere to them in order to remain free. *See id.* *See also id.* at 193 (“My very strongly tempting you is simply not the same as my coercing you or putting you under duress.”). For Murphy (and I am in agreement with him here), with respect to coercion we ought to be concerned not with coercion per se, but with coercion that is unfair. *Id.* at 194. Requiring a prisoner to undergo a procedure that will ensure that she is not a danger to society (assuming it is the least invasive procedure required to obtain such an outcome) before she is released is not unfair.

²¹ It is entirely possible that there could be unintended effects of having one’s empathy enhanced, however. As an extreme example, one might become overwhelmed with the amount of suffering in the world and commit suicide.

²² One might ask whether the empathy enhancement procedure would result in a more fundamental change than is produced through treatment with psychotropic drugs. This seems unlikely given that researchers have found that SSRIs can result in significant changes to the core “Big Five” personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. *See Tang et al., Personality Changes During Depression Treatment: A Placebo-Controlled Trial*, 66 ARCHIVES GEN. PSYCHIATRY 1322, 1322 (2009). Empathy enhancement may have similar—or even greater—effects on these core traits, yet the difference would be in degree rather than in kind.

Coerced Empathy Enhancement as Distinct from True Mind Control

There is even an argument to be made that the empathy enhancement procedure would be less coercive than psychotropic medication. Consider that the enhancement procedure provides the individual with a new experience—the experience of greater empathy for one’s fellow beings. While it is likely that this new experience will influence attitudes and behavior, it does not force the individual to change her behavior. Indeed, a previously psychopathic individual who derived pleasure from watching others suffer would still be free to engage in sadistic acts. The new perspective on life gained from the empathy enhancement procedure would likely (and hopefully) sway the person from engaging in the same harmful behavior, but it would not be like a switch that made it impossible for the person to act on her desires—the essence of true mind control. Instead, the empathy enhancement procedure would be more akin to other kinds of life-changing experiences that influence the way a person thinks about and interacts with the world.

Both positive and negative experiences can have a lasting impact on a person’s life and permanently alter a person’s perspective on the world and her attitude towards other people. Sometimes people are forced to experience things that alter their personality in a negative manner. For example, a person might be subjected to the horrors of war and have to watch her entire family killed. The impact of such an experience is likely to last throughout the person’s life. If the person develops negative character traits following such an experience, we are likely to mitigate our judgment of the person based on her past experience. However, we are unlikely to let the person off the hook entirely. We are almost certainly unlikely to say that she has been deprived of her free will. We would respond similarly if the experience of watching her family being killed brought out

positive character traits that were previously absent. We are not going to downplay positive traits such as courage, newfound compassion for others, etc., that result, by saying that the continuing effects of the negative event have deprived the person of her free will. On the contrary, we will praise the resilience of the person and commend her for any positive life changes she makes as a result of some insight gained from the negative experience.

The example above is akin to the resulting impact of forced empathy enhancement. A psychopath who is forced to undergo the empathy enhancement procedure would not lose her free will. She would still be able to act in cruel and uncaring ways if she so chose; and she would be blameworthy for these acts. By the same token, she would be praiseworthy for engaging in kind, caring, acts. The fact that the new behavior was due to the (previous) psychopath's new perspective on the world as a result of the empathy enhancement procedure would not take away from her praiseworthiness.

Thus, if a person were subjected to forced empathy enhancement, the worst that could be said of the perpetrators is that they committed a wrong against the person by violating her autonomy. This is not to downplay the wrong. Autonomy is rightly considered to be one of the most important bioethical principles.²³ All I am arguing is that it does not follow from a violation of autonomy that has a resulting effect on a person's mind that the person loses her free will.²⁴ If I am right about this point then my argument

²³ See TOM L. BEAUCHAMP & JAMES F. CHILDRESS, *PRINCIPLES OF BIOMEDICAL ETHICS* 57-104 (5th ed. 2001) for an in-depth discussion of the importance of respecting patient autonomy.

²⁴ One might be concerned that a violation of autonomy that is intended to bring about a certain result and, in fact, brings about that result, is an assault on free will. For example, imagine a gang that rapes women with the intent of instilling continuous fear and terror in

that the person who *voluntarily* undergoes the empathy enhancement procedure also maintains her free will should be easier to accept.

Part II: The Effect of Enhanced Empathy on Responsibility

If a person who has an enhanced sense of empathy maintains her free will, the next question is what effect, if any, empathy enhancement has on moral and legal responsibility. Given that the enhanced empath has free will (or so I have argued) it should be clear that the enhanced empath should not be subject to *diminished* moral or legal responsibility. In the legal context, it is difficult to imagine how a person with enhanced empathy could use that enhancement to argue that she is unable to form the *mens rea* necessary for criminal liability.

However, there is another sense in which enhanced empathy might lead to a difference in how we evaluate a person from both a moral and a legal perspective. We might find that acts that are supererogatory for the non-enhanced empath become

the women. If a woman is raped by the gang and ends up feeling terrified for the rest of her life, there's intuitive pull to the idea that her free will, in addition to her autonomy, has been violated. However, I think this case is distinct from a case of forced empathy enhancement. In the former case, we can imagine that the terrified woman wishes that she could rid herself of the effects of terror, e.g., jumpiness. Yet, she is unable to quell her jumpiness despite her desire to do so. In a Frankfurtian sense, she is not free; her jumpiness acts as a kind of irresistible impulse. In the case of the psychopath who is forced to undergo empathy enhancement, I don't think the same can be said. In that case, the psychopath has a new experience of ECE—one that she has not had before—and this will hopefully influence her to stop harming others. But the experience of ECE does not seem like an irresistible impulse. As I noted in the main text above, the (previous) psychopath would still be able to choose to harm others if she wished. Thus, it seems that there is a difference between a violation of autonomy that causes a person to feel that her will is compelled and one that causes a person to feel that her will is influenced by a new experience of the world.

obligatory for the enhanced empath. Another way of looking at this, as I'll argue, is to say that the enhanced empath might have a wider range of *special obligations* than the ordinary empath. Let me begin by discussing supererogation.

Supererogation

Put simply, supererogatory acts are those acts that are praiseworthy but not morally obligatory. For example, imagine that you are standing next to the president and you see someone nearby pull a gun out of her coat and take aim. If you were to jump in front of the president, thus taking the bullet, you would surely be praised for your sacrifice. However, at the same time, choosing not to jump in front of the president in order to take the bullet would not be a morally blameworthy act. Consider another familiar example, slightly tweaked. Imagine that you are standing on a platform above a railroad track with a group of other people. You look below and see that there are five people tied to the track about fifty feet beyond the platform. Everyone on the platform notices that a train is approaching and that there is no time to free the five individuals below. In order to stop the train, you hurl yourself off the platform and onto the track between the train and the five trapped individuals. You would surely be hailed as a hero. At the same time, the other people on the platform who choose not to sacrifice their lives would not be blamed for their failure to sacrifice their individuals lives to save fives lives. Consider one more example. Imagine that you win the lottery, but instead of taking the prize for yourself you allow it to be distributed evenly among one thousand people in need. Few people would argue that you had a moral duty to disown your property interest in the lottery winnings. Thus, while you would be praised for your generosity, few people would blame you if you decided to keep the majority of the money for yourself.

Contrast the preceding cases with cases of moral obligation. In cases of moral obligation, there is a duty to perform an action, which is (usually) accompanied by an attendant sense of blameworthiness for a failure to perform the action. Of course, there are also actions that have no moral content, in other words, whether one performs them or not does not matter from a moral perspective. J.O. Urmson codifies the categories of action discussed thus far as follows:

1. Actions that are good to do but bad not to do;
2. Actions that are neither good to do nor bad not to do;
3. Actions that are bad to do and good not to do;
4. Actions that are good to do but not bad not to do.²⁵

The fourth category defines supererogation.²⁶

This four-category schema, while helpful for placing supererogatory action within the context of other actions, simplifies the complex moral nature of actions. The biggest problem is how we go about defining some action as supererogatory rather than morally obligatory. In the first paragraph of this section I gave a few examples of what might be considered supererogatory actions. The first two examples I provided are arguably prototypical. Yet, even within these examples we can come up with wrinkles that make

²⁵ See generally J.O. Urmson, *Saints and Heroes*, in *ESSAYS IN MORAL PHILOSOPHY* (A.I. Melden ed., 1958). For the four-category schema as presented here, see R.M. Chisholm, *Supererogation and Offence: A Conceptual Schema for Ethics*, 5 *RATIO* 1, 1-14 (1963).

²⁶ Peter Singer has argued that the category of “actions that are good to do but not bad not to do” is much narrower (if not non-existent) than ordinarily thought. See Peter Singer, *Famine, Affluence, and Morality*, 1 *PHIL. & PUB. AFF.* 229 (1972). For example, Singer notes that when a well-off person gives money to the poor it is generally considered charity, i.e., supererogatory. Yet, Singer argues that the well off have a moral obligation to provide aid to the poor as long as doing so does not require sacrificing something of greater moral value, and thus providing “charity” is not, in fact, supererogatory. See *id.* at 235, 240-41. For my purposes, Singer’s argument will not be particularly important. I will be arguing that even where there is a moral duty to act, certain individuals can have a *greater* moral duty depending on their relationship to the person in need.

assessment of their moral nature more difficult. Imagine that instead of the president, the person whom the person with the gun intends to shoot is your young son or daughter. In that case, it is much more difficult to say that your failure to jump in front of the bullet would not be morally blameworthy. Thus, it would seem that at least a portion of what makes an act supererogatory depends on the relationship between the actor and the person who benefits from the act.

Supererogation and Obligation

Interestingly, most United States jurisdictions—to an extent—codify this intuition that the status of a person’s relationship to another person is relevant to whether we consider certain acts to be moral obligatory or supererogatory. In general, there is no legal obligation to affirmatively assist another person who is in need. For example, if you see a stranger drowning out past the breakers in the ocean, you have no legal duty to save her—even if you are a strong swimmer. However, the law generally recognizes four scenarios where a person has a legal duty to help another person in need:

[F]irst, where a statute imposes a duty to care for another; second, where one stands in a certain status relationship to another; third, where one has assumed a contractual duty to care for another; and fourth, where one has voluntarily assumed the care of another and so secluded the helpless person as to prevent others from rendering aid[.]²⁷

For my purposes, I will focus on the second and fourth scenarios. I will refer to the second scenario as the “status relationship” scenario and the fourth scenario as the “voluntary care” scenario. I focus on these scenarios for the following reasons. The voluntary care scenario is illustrative for how a supererogatory act can be transformed

²⁷ Jones v. United States, 308 F.2d 307, 310 (1962).

into a morally obligatory act. The status relationship scenario is also illustrative of this transformation and it further demonstrates, as I will argue, that it is empathy that does the transformative work.²⁸ Relatedly, as I will argue, the legal treatment of status relationships can be seen as reflecting moral judgments that track degrees of ECE.

Let's begin with the voluntary care scenario to see how a supererogatory act can be transformed into a morally obligatory act. As I already stated, as a general legal rule you do not have a duty to help someone who is drowning, even if you are a strong swimmer.²⁹ However, if you are surrounded by a group of people who also see the person

²⁸ It's possible that empathy has a role in all four scenarios, but I see the link between empathy and statutory duty or contractual duty as more tenuous. There is probably a stronger argument that empathy plays a role in the voluntary care scenario, but for my purposes this is not particularly important. Michael Slote would likely want to say that all four scenarios are covered by empathy in some way. However, unlike Slote, I am not committed (though I also do not necessarily oppose) a moral theory that places empathy at the center of *all* moral judgments (and legal judgments that rely, at least in part, on moral evaluations). See MICHAEL SLOTE, *THE ETHICS OF CARE AND EMPATHY* 1 (2007) (“[I] seek[] to show that a care ethical approach [based around empathy] makes sense across the whole range of normative moral and political issues that philosophers have sought to deal with.”).

²⁹ Although there is generally no legal duty to render aid (in the United States at least), a few states have passed “Good Samaritan” laws that require a person to offer “reasonable assistance” to a person in need if such assistance “can be rendered without danger or peril.” VT. STAT. ANN. tit. 12, §519(a). See also MINN. STATE. ANN. 604A.01; R.I. GEN. LAWS §11-56-1; WIS. STAT. ANN. §940.34; JOSHUA DRESSLER, *UNDERSTANDING CRIMINAL LAW* 109 (7th ed. 2015). Good Samaritan laws have a different character than voluntary care and status relationship scenarios. In the latter two scenarios we can see how the specific scenario—starting to render aid such that others do not feel it necessary to do so, or being in a close relationship with someone—transforms what might be considered supererogatory acts into morally, and legally, obligatory acts. This is not true in the case of Good Samaritan laws. Instead, the language of those laws seems to codify what is already a moral duty to act. If we look at the language of the Vermont statute, for example, we see that, in essence, it states that one may not stand by and watch another person die when rendering aid, or calling for help, would not put one in danger. This is a very minimal standard, and it is difficult to imagine that a person who stood by in such a scenario would not be morally (though not legally, in most jurisdictions) condemned. Thus, there is no supererogatory element at play in Good Samaritan laws.

drowning, and you take the initiative to jump into the ocean and swim out to help, you have now assumed a legal duty to render aid—assuming you remain able to do so. If you jump into the ocean to help the drowning person, you signal to the other people on shore that they do not need to act; you have assumed responsibility for aiding the drowning swimmer. There may have been other people on shore who were willing to help, but decided to wait on shore as you made the rescue. If you reach the drowning person and suddenly decide that you don't feel like rescuing her after all, you have at least radically reduced the potential that someone else will be able to reach her before she drowns. Thus, what was initially a supererogatory act—your decision to risk your life to save the drowning person—transforms into a legal duty, and arguably a moral duty as well. After all, the legal duty that attaches in this case clearly mirrors a moral intuition that it would be wrong to assume responsibility for helping someone, such that others think the person no longer requires help, only to abandon one's efforts on a whim.

Status Relationships and Supererogation

Let's move on to status relationships, where I will argue that empathy provides the moral impetus for transforming what would be a supererogatory act for one person into a morally obligatory act for another. Status relationships describe familiar relationships such as between a parent and child,³⁰ or between a married couple,³¹ or between two people who happen to have a very close relationship.³² If your teenage child falls down a

³⁰ DRESSLER, *supra* note 29, at 107.

³¹ *Id.*

³² *See, e.g., State ex. rel. Kuntz v. Thirteenth Judicial District*, 995 P.2d 951 (Mont. 2000) (couple in a long-term relationship owe each other same duty as owed between spouses).

flight of stairs while the two of you are out in public, you have a legal duty to render aid. Perhaps more fundamentally, you have a legal duty to feed and house your child.³³ Of course, you do not have a legal duty to feed and house other people's children.³⁴ You do not even have a *legal* duty to render aid to someone else's child when doing so would not inconvenience you. A one-hundred year-old example from James Fitzjames Stephen sums up this standard: "A number of people who stand round a shallow pond in which a child is drowning, and let it drown without taking the trouble to ascertain the depth of the pond, are, no doubt, shameful cowards, but they can hardly be said to have killed the child."³⁵

³³ See generally *People v. Beardsley*, 113 N.W. 1128 (Mich. 1907).

³⁴ Unless you have held yourself out as the child's caretaker. See *State v. Sherman*, 266 S.W.3d 395 (Tenn. 2008).

³⁵ JAMES F. STEPHEN, *A HISTORY OF THE CRIMINAL LAW OF ENGLAND* 10 (1883). While one would not be said to have been legally responsible for the child's death (at least in jurisdictions that do not have "Good Samaritan" laws), the issue is not as clear-cut from a moral perspective. There are two issues here. First, even if we were willing to say that the person who let the child drown did not *kill* the child, we would surely say that she is morally responsible for the child's death. This tracks Peter Singer's argument that where saving a drowning child's life requires no great sacrifice on our part, we have a moral duty to do so. See Singer, *supra* note 26, at 231. Second, intuitively, it may seem that even if we hold that a person is morally responsible for letting the child die, there is still some moral distinction between killing and letting die. For example, this intuition is codified in laws that allow physicians to refrain from providing life-saving care to a patient and to prescribe life-ending medication (in the three states that allow for such) but do not allow physicians to administer drugs that will actively end a patient's life. See, e.g., *Death With Dignity Act: Frequently Asked Questions*, OREGON.GOV, <https://public.health.oregon.gov/ProviderPartnerResources/EvaluationResearch/DeathwithDignityAct/Pages/faqs.aspx#present> (last visited Jan. 12, 2016) ("A physician may be present if a patient wishes it, as long as the physician does not administer the medication him/herself."). Yet, many philosophers have argued that there is no moral distinction between killing and letting die, assuming the intention behind the act (or lack thereof) is the same. For example, Heidi Malm asks us to imagine two cases. In the first case, Smith finds a child inside a machine and presses a button that will cause the machine to crush the child, because he wishes to see how flat a person can be. In the second case, Jones

So what transforms shameful cowardice into criminal liability in the case of a parent who similarly lets her child drown?³⁶ As with the strangers, there is no affirmative act on the part of the parent. Rather than killing her child, she simply lets her die. Yet even in a room full of shameful cowards, there are those who can rightly be looked upon as especially worthy of moral condemnation. As I will argue, it is this moral condemnation, which tracks degrees of ECE, that best explains the law's treatment of status relationships. However, before I turn to that argument let me discuss the philosophical literature surrounding *special obligations*, of which legally recognized status relationships are a subset.

finds a child in a machine that is already on and about to crush the child. Jones could easily stop the machine by pressing a button but chooses not to because he wishes to see how flat a person can be. See Heidi Malm, *Killing, Letting Die, and Simply Conflicts*, 18 PHIL. & PUB. AFF. 238, 240 (1989). As Malm sees it, there is no moral difference between Jones' action and Smith's omission. See *id.* at 241. Malm's argument relies on all considerations such as the agents' motives, their lack of prior intentions, etc., being equalized. See *id.* at 240-41. In reality, it is usually easier to determine the intentions of someone who performs a positive action as opposed to someone who refrains from performing some act. I will return to this point later in this chapter.

³⁶ While a parent has a legal duty to look after her child and render aid when her child is in need, there are limits to this duty. In the famous case of *McFall v. Shimp*, 10 Pa. D. & C. 3d 90 (1978), McFall sued his cousin, Shimp, after Shimp refused to donate his bone marrow in order to save McFall's life. Shimp was the only person who could save McFall, and McFall sought to have the court force Shimp to give up his marrow. Judge Flaherty in the case refused to order Shimp to help his cousin, stating that "[f]orceable extraction of living body tissue causes revulsion to the judicial mind." *Id.* at 92. However, the judge also made it clear that Shimp's decision was "morally indefensible." *Id.* at 91. It is almost certain that the same ruling would apply to a case where a mother refused to donate marrow, or any other body part, to save her child. See ANGELA HOLDER, *LEGAL ISSUES IN PEDIATRICS AND ADOLESCENT MEDICINE* 171 (2d ed. 1985) for a discussion. All of this is simply to point out that there are limits to what the law will require a person to do to help another, even if one owes a moral duty of care. None of this diminishes the fact that we would assign greater moral judgment to a parent who refuses to donate marrow to save her child than we would to a cousin or a neighbor or a stranger.

Special Obligations and the Emotive Component of Empathy

In the case of the drowning child, there is surely a moral obligation on everyone's part to render aid—as Stephen implies when he labels the onlookers as “shameful cowards.” But it seems fair to say that there is an even greater moral obligation on the part of the parent's child. Similarly, commonsense morality seems to tell us that we have greater moral obligations to our friends than we do to strangers.³⁷ One might ask where these greater moral obligations come from. A number of philosophers have wrested with the issue of special obligations that result from status relationships (or roles).³⁸

One strategy taken by philosophers is to attempt to place special obligations within a consequentialist framework.³⁹ According to this view, there are no special obligations per se (i.e., obligations that can run counter to the general obligations to promote the general good) but only special cases where an agent acting *as though* she had special obligations actually contributes to the general good. Thus, under this view, you only have duties to promote the general good—but one way of doing that is through extending special concern to your friends and family. The problem with this view is that it ignores the feeling that most of us have that we have special obligations to our friends

³⁷ Note that this is not the same as saying that we have greater moral responsibility to those *near* us, i.e., to people in our town or country as compared to people in far-off places. Thus, there is not necessarily a conflict between this commonsense view and Singer's argument that we owe the same duty of assistance to those in distant countries as we do to our own neighbors. It's possible to hold Singer's view and still believe that we have special obligations to our friends and family. See Diane Jeske, *Special Obligations*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY 4-5 (Jan. 5, 2014), <http://plato.stanford.edu/entries/special-obligations/>. I will return to this point later in this chapter.

³⁸ For an overview, see generally *id.*

³⁹ See, e.g., HENRY SIDGWICK, *THE METHODS OF ETHICS* 434 (Hackett 1981) (1907).

and family *even if* fulfilling those obligations goes against the general good.⁴⁰ Thus, it seems that consequentialism cannot provide an adequate account of the special obligations that commonsense morality tells us that we have.

If we abandon consequentialism as a way to explain away special obligations we can return to the comfort of commonsense morality. But we still must ask ourselves from where special obligations derive their moral power. For some philosophers, special obligations are simply part of what it means to be a flourishing human being. According to Michael Hardimon, “our roles as family members . . . are the source of some of the deepest and most important bonds we have,”⁴¹ and without special obligations to our families and friends those relationships would take on an entirely different character. Thus, it seems plausible that special obligations derive their moral power from being an integral part of institutions, such as family and friendship, that we value. Yet, this explanation feels incomplete. We might still want to ask what it is about certain relationships, such as family relationships and friendships, that cause us to see them as having greater value—and thus instantiating special obligations in us—than other

⁴⁰ Another way of looking at this is to grant that we might on occasion give preference to our friends and family to the detriment of the general good, but because special relationships are necessary to fulfill the general good, this occasional act against the general good (i.e., this occasional *wrong*) is acceptable. See Peter Railton, *Alienation, Consequentialism, and the Demands of Morality*, 13 PHIL. & PUB. AFF. 134 (1984). The difficulty with this view is that most of us don’t want to have to believe that giving our friends and family special preference is a wrong. See Jeske, *supra* note 37, at 4.

⁴¹ Michael O. Hardimon, *Role Obligations*, 91 J. PHIL. 333, 353 (1994). Hardimon focuses on what he calls *institutional roles*, and since he does not consider friendship to be an institution he excludes friendships from his analysis. However, I think it’s fair to say that whether or not friendship is considered as an institution, we can extend the analysis of special obligations between family members to friendships as well. In any case, I will be assuming that here.

relationships. I think that Michael Slote has provided the most compelling answer to this question.

Empathy and Special Obligations

Slote argues that the best way to understand special obligations is through the lens of empathy.⁴² Consider that in all of the contexts where special obligations exist, feelings of empathy are likely to be stronger. For example: we feel more empathically connected to our own parents than to someone else's parents; parents feel more empathically connected to their children than to someone else's children; we feel more empathically connected to our friends than we do to strangers. These are all cases where we intuitively feel that we have special obligations. For Slote, this is no coincidence. He argues that distinctions of empathy undergird all moral distinctions.⁴³ While I am not willing, at this point, to accept this broad conclusion, I agree with Slote that empathy is at least intimately connected to the moral distinctions that we intuitively make. I discussed this point briefly in Chapter Four, but let me expand upon it here.

As Slote points out, relying on the psychological literature about empathy, for people with a normal, mature, sense of empathy (broadly construed), the strength of empathic feeling towards others increases with familiarity, salience, proximity, and immediacy.⁴⁴ As I already discussed in Chapter Four with respect to Slote's theory, it's

⁴² Slote does not actually use the term "special obligations," but he is clearly talking about the same concept.

⁴³ SLOTE, *supra* note 28, at 8.

⁴⁴ See *id.* at 23. See also Stephanie D. Preston & Frans B.M. de Waal, *Empathy: It's Ultimate and Proximate Bases*, 25 BEHAV. & BRAIN SCI. 1, 1 (2002).

important to distinguish between RCE and ECE here. RCE will not be subject to the strengthening criteria mentioned above.^{45, 46} Recall the example of Lisa earlier in this chapter. Lisa has an operative RCE with respect to suffering children in far-off countries but wishes to undergo the empathy enhancement procedure so that she may also engage her ECE and thus feel more motivated to help the children. If only ECE is subject to the strengthening criteria mentioned above, then only ECE is relevant to special obligations. ECE is the component whose strength varies depending on our relationship to different people.

Thus, tentatively, it seems fair to say that the stronger one's ECE towards a particular person, the stronger one's moral obligations are to that person. This applies to situations involving friends and family, as opposed to strangers, as well as situations involving witnessing someone in pain, as opposed to merely hearing about someone in pain. As Slote puts it: "Turning away from someone we see [in pain]. . . seems *worse* than ignoring someone [in pain] whom one knows about only by description."⁴⁷ Where Slote's argument falls short is in explaining *why* this is the case. Why is it worse to turn away from someone we see, or a family member, or a friend, as opposed to ignoring someone in a far-off place, or a stranger?

Again, consider a familiar example. We are faced with a drowning child in front of us and a starving child in a far-off country. Assume, as Singer does, that it is just as

⁴⁵ An exception might be in a case of a truly incomprehensible being—say, a non-biological alien life form—with whom we have difficulty in engaging our rational component of empathy.

⁴⁶ I will discuss this point further later in this chapter.

⁴⁷ SLOTE, *supra* note 28, at 23 (emphasis in original).

easy to save a starving child in another country (through a donation to Oxfam) as it is to save the drowning child. The drowning child in front of us will (assuming we are normally developed) engage our ECE more than a starving child in a far-off place. As such, we will feel a stronger compulsion to act to save the drowning child than to donate to Oxfam to save the starving child. ECE gives us a greater feeling of *motivation* to act. Thus, if we fail to save the drowning child despite feeling this greater motivation—perhaps because we are late to an important appointment—we have had to actively resist what we knew and felt to be right in a way that we do not have to do in the case of the starving child, where our ECE is not as engaged and we, as a result, do not feel as much motivation to act. It is this issue of motivation and active resistance that I contend explains *why* we judge someone who chooses not to help the drowning child as worse than someone who chooses not to help the starving child. Put another way, this issue explains why ECE is relevant to moral responsibility—and why we have special obligations.

This same argument helps explain why we see a parent as having a greater moral obligation to help her drowning child than does a stranger. I think it is helpful to think of the two scenarios—parent versus stranger—in terms of the respective roles of RCE and ECE. In the case of the child drowning in a puddle, there is obviously a moral duty on the part of each person near the child to render aid if they can. This is true even if they do not feel ECE. However, while everyone in that situation has a moral duty—not a supererogatory choice—to help the drowning child, I think it is fair to say that the child's parents have a *stronger* moral duty to render aid. There is an expectation that parents will feel ECE to a stronger degree than the rest of the strangers in the crowd, and thus the

parents ought to feel more motivated to render aid to their child. On an intuitive level, the parents' connection to their child is expected to produce a greater degree of ECE towards the child. Thus, given the argument I provided in the preceding paragraph, we are justified in judging the parent more harshly than the stranger.

In the example of the drowning child we can say that both the parents and the stranger have a moral duty to help the child since doing so requires no risk, even if the parents have a *greater* moral duty to help. If we want to see the transformation of supererogatory action into morally obligatory action at work, however, we need to consider a different case where the stranger would have little to no moral obligation to offer aid. Imagine that sixteen year-old girl has had several drinks at a party and calls up her mother to ask for a ride home since all of her friends are drunk. Surely the mother has a moral obligation to pick up her daughter to stop her from getting in a car with a drunk driver. If the girl called up a stranger, or a distant acquaintance, that person would have far less—if any—of a moral obligation to give the girl a ride. In any case, if the stranger did give the girl a ride her action would likely be hailed as supererogatory. In this case we can see how an action transforms from being supererogatory to morally obligatory depending on the relationship between the parties. A parent's relationship to her child attaches a greater moral obligation to the parent with respect to certain actions due to the child; and these moral obligations exceed those due from people other than the child's parents. As such, what might be supererogatory in one situation becomes morally obligatory if the person is the child's parent rather than a stranger. Thus, we can now view special obligations in terms of a transformation of supererogatory acts into morally obligatory acts depending on the relationship between the parties. And this change in the

status of a person's moral obligations is at least in part due to our assumption (or expectation) that parents have stronger ECE towards their children.

It's interesting to note that in the case of the parent's moral obligations to her children we can only infer that greater ECE is present. And we are willing to judge her accordingly, even if we have no proof as to her actual state of mind.⁴⁸ This leads us back to legally recognized status relationships. It is now possible to see the legal take on status relationships as a heuristic for determining degrees of ECE. We *expect* that people in certain status relationships with others will have greater ECE.

The Principle of Reflective Acceptability

Before moving on to an explicit treatment of the enhanced empath's responsibility let me address a concern about special obligations that is especially worth addressing if we tie special obligations to degrees of ECE. The concern is that even if we accept that special obligations exist we might still worry about the reach and extent of those obligations. For example, most people accept that we have special obligations to our parents, namely, to take care of them as they get older. But what about someone whose parents were abusive when she was a child? It would seem monstrous to impose a special obligation on this person to take care of her parents simply because they are her parents and she has a special obligation towards them. Hardimon responds to this concern with the *principle of reflective acceptability* (PRA), which holds that "role obligations are not morally binding

⁴⁸ In legal language, it could be said that we hold these people responsible under a "reasonable person" standard that considers not the actual fact of the matter but what we, as members of a society, can reasonably expect from our fellows.

unless the roles to which they attach are reflectively acceptable.”⁴⁹ For those of us who were raised in relatively healthy family environments, we probably feel that we have a special obligation to take care of our parents as they age, and if we reflect on that obligation we will likely decide that the obligation is acceptable. For someone who was raised in an abusive family, it does not seem reflectively acceptable to impose a special obligation on the person to care for her aging parents. Thus, while children who were raised in a relatively healthy family environment have a moral obligation to care for their parents as they age, children who were raised by abusive parents have no such moral obligation.

The PRA is especially important, I think, if we accept that special obligations depend on degrees of ECE, based on Slote’s (and my own) view that our ascriptions of moral blameworthiness track (or often track) distinctions of empathy. The concern, again, is that a person might feel strong ECE in a context where it would seem unfair to hold that person to a higher degree of moral responsibility. Using the example in the preceding paragraph, a person might feel strong ECE towards her parents despite their abusiveness. Yet, do we really want to say that strength of the person’s ECE in such a case corresponds with a greater moral responsibility to care for her parents? I think not. In this case, the PRA tells us that the person should not be held to a higher standard even though she feels great ECE towards her parents. The PRA will remain important as we consider the enhanced empath’s degrees of responsibility.

⁴⁹ Hardimon, *supra* note 41, at 350.

Status Relationships as Heuristics for Expected Empathy

Recall that in the preceding section I said that legally recognized status relationships operate as a heuristic for determining degrees of ECE, which track ascriptions of moral blameworthiness. In the case of ordinary empathys we have expectations about the degree of ECE a person will feel in specific relationships, for example, a parent's relationship with her children, so when we see a parent and child together we *infer* that there is a strong empathic connection between the two. Contrast this with the case of the enhanced empath. With the enhanced empath, there can be no doubt as to her state of mind. The procedure that she undergoes ensures that she will feel an enhanced sense of ECE. We would not need to infer that a strong empathic relationship exists; we could be certain that one exists. Thus, there would be as much (if not more) reason to judge the enhanced empath harshly in situations where there is an expectation of heightened ECE.

I presented the examples above to provide grounding for the claims that supererogatory acts can become morally obligatory acts, and that this relates to ECE; and special obligations derive their moral power from degrees of expected ECE. But the examples I provided rely on common intuitions about what situations have the appropriate ingredients to work the moral alchemy of transmuting supererogation into moral obligation. A parent's relationship to her children is a strong and easily agreed upon example where moral alchemy can take place. However, as we move now into the reality of the enhanced empath, the intuitive landscape is likely to become distorted.

Empathy Enhancement Would Target the Emotive Component of Empathy

The enhanced empath has undergone a procedure that magnifies the power and scope of her ECE. Thus far, I have not been especially specific about that magnification. I have

already argued that empathy can be broken down into two components: a rational component and an emotive component. Based on my argument in Chapter Four, with respect to RCE, it seems that RCE would not be the focus of an empathy enhancement procedure. I do not mean that there would be some technical choice on the part of the clinician to focus on one component of empathy or the other. Rather, I argue that the enhancement procedure would necessarily target ECE. Recall that in my discussion of RCE, I emphasized that that component simply involves the ability to imaginatively project oneself into another's shoes and to realize that one would want help (or to not be harmed) if one were in the other's place.

As I've argued, even the psychopath is capable of this feat. To say that a person had an unusually strong RCE would be more akin to saying that the person had greater overall intelligence, rather than saying the person was especially emotionally empathic or especially motivated to act in empathic ways. This is because RCE is associated with making logical connections. Perhaps one could argue that Peter Singer has stronger RCE than the average person—but this is simply a product of his intelligence and his willingness to follow logical leads wherever they may go. Thus, enhancing RCE would involve something more like enhancing one's overall intelligence, rather than targeting the amygdala or other area of the brain that is associated with ECE.

The Enhanced Empath's Expanded Sphere of Special Obligations

Thus, I think it's fair to say that the empathy enhancement procedure would have its primary effect on ECE. If this is the case, it is possible that there are at least certain situations where acts that would be supererogatory for the non-enhanced person would become morally obligatory for the enhanced empath. Recall my discussion of

supererogation and special obligations above. I argued that special obligations are dependent (at least in part) on degrees of ECE; in other words, special obligations derive their moral force from the empathic connection that exists in certain relationships. I further argued that special obligations can be seen as situations where acts that would be supererogatory for one person, generally, become morally obligatory if the person is in a certain relationship with another person (e.g., a parent-child relationship). I also argued, following Slote, that other special obligations (construed broadly) are best understood through the lens of empathy. For example, degrees of ECE help explain why we would judge someone who let a baby drown in front of her more harshly than someone who failed to donate a small amount of money to Oxfam—even if the outcomes and effort are the same. If special obligations track degrees of ECE then I think it is fair to say that the enhanced empath would find herself subject to a new set of special obligations. Or, put in terms of supererogation, the enhanced empath would find that certain acts that were supererogatory prior to her receiving the empathy enhancement procedure become morally obligatory following the procedure.⁵⁰ To help understand this point, let me return

⁵⁰ At this point in my argument, someone might raise an objection with respect to my treatment of empathy when it comes to psychopaths versus my treatment of empathy when it comes to enhanced empaths. In Chapter Four, I argued that psychopaths should not be excused from moral responsibility. I argued that the psychopath's RCE is functional and, thus, she is morally responsible for her actions. The fact that the psychopath lacks ECE is not a good reason to say that she has a lesser degree of moral responsibility. Yet, in this chapter I have argued that the enhanced empath—who undergoes a procedure that enhances her ECE—should be (*prima facie*) held to a higher moral standard. At first glance, there might appear to be a contradiction between these two stances. But this seeming contradiction disappears when we consider that moral duties attach even in the absence of special obligations, but that the two kinds of duties rest of different aspects of empathy. Consider Slote's acknowledgment that "we have moral obligations to help strangers and people we only know about" even where we feel little to no empathy towards those people. SLOTE, *supra* note 28, at 33. It is difficult to

to the original setup I provided for why someone might undergo the empathy enhancement procedure.

In Chapter Two, I considered Singer's argument regarding our moral duty to help those in need even if they are not within our immediate social or geopolitical sphere. I have posited that this "expanding circle" could best be achieved through artificial empathy enhancement that would allow people to feel ECE towards people in far flung places around the globe, in other words, it would allow people to align their ECE with their RCE. And in this chapter, I have argued that part of what motivates us to hold people in certain kinds of status relationships to a higher degree of moral responsibility is that we expect people in those kinds of relationships to feel a stronger sense of ECE. And, again, with the enhanced empath we do not have to rely on expectation—we can *know* that she feels a stronger sense of ECE. She underwent a procedure specifically to obtain that outcome. I have also argued that one of the—if not *the*—strongest motivating

square Slote's account of empathy with obligations to those towards whom we feel no empathy. This is where my account of empathy as a two-part concept helps, and it is how we can clear up the seeming contradiction between saying that a lack of ECE is not relevant to moral responsibility while a surplus of ECE is relevant to moral responsibility. As I have attempted to show, RCE exists in nearly all people. The exceptions that I have mentioned include people with severe forms of mental illness such as schizophrenia and people with severe autism. Psychopaths have an operative RCE. RCE is the ground of moral responsibility towards all people—it forms the basis of the moral obligations that Slote, and myself, believes we have towards all people. Thus, one can be entirely devoid of ECE and still owe moral obligations to people. But such a person would owe no *special obligations* to people because special obligations depend on degrees of ECE. Since psychopaths, for example, have no (or highly diminished) ECE they cannot, under my view, be held to a higher degree of moral responsibility with respect to particular relationships they have with other people. On other hand, for those people who *do* have operative ECE, their moral responsibility to specific people is dependent on the strength of their ECE with respect to each person. Their responsibility to individuals can never drop below the baseline provided by RCE, but it can exceed that baseline depending on the degree of ECE.

factors for receiving the empathy enhancement procedure is to be able to feel more motivated to act on what one knows to be morally right. Let's consider how this would apply to the case of Lisa from earlier in this chapter.

Easy Cases

After the empathy enhancement procedure Lisa suddenly feels the same degree of ECE towards children that she knows are suffering in far-off places as she does towards children that are suffering in her hometown. As such, Lisa feels motivated to help the far-off children in the same way she feels motivated to help the children in her hometown. This is precisely the outcome that Lisa was hoping for. Lisa feels more at ease with herself now that her ECE is in line with her RCE. Given Lisa's new circumstances, how are we, from the outside, to judge her?

If what I have argued so far about the connection between moral responsibility and degrees of ECE is correct, it would tentatively seem that Lisa now has a new (or at least a stronger) moral obligation that she did not have before. With Lisa's new state, our intuition that people have a stronger obligation to help those close to them breaks down. We can no longer say that Lisa would have a lesser moral obligation to help those in far-off places due to feeling a lesser degree of ECE towards those people. Instead, if moral obligation tracks degrees of ECE, Lisa now has an equivalent moral obligation to help suffering children in far-off places as she does to help children in her hometown.⁵¹ I

⁵¹ This is a somewhat loose way of talking, which I think is justified for making my larger point. However, to avoid any objection let me be a bit more precise. In Singer's argument he is careful to say that we have the same obligation to children in far-off places as we do to children nearby *assuming* that all else is equal. In other words, if we can save a life (or reduce suffering in a life) by helping a drowning child in front of us and an equivalent effort, through a small donation to Oxfam, would save a life (or reduce

noted that this new moral obligation is tentative. This is because we must apply the PRA in order to determine if it is *fair* to hold Lisa to this new responsibility. This is a relatively easy case, I think. Here, Lisa underwent the empathy enhancement procedure specifically because she wanted to feel motivated to help children in far-off places. Lisa already believed, through her RCE, that she should help these children. From the outside, we would not have said that Lisa had an equivalent moral obligation to help those in far-off places as she did to help those in close to her; in other words, we would have considered her efforts to help those in far-off places to be (at least in a sense) supererogatory. But the very fact that we would have judged her actions as supererogatory means that we see value in helping those in far-off places. As such, Lisa's newfound ECE towards children in far-off places aligns with something that we value and, thus, saying that Lisa now has a moral obligation—as opposed to a supererogatory choice—to help those children is reflectively acceptable.⁵² Given the importance of the PRA here, and to special

suffering in a life) of a child in a far-off place, our moral obligation is the same to both lives. If we assume that Lisa now has this equivalent moral obligation, given her enhanced empathy, it's still possible to say that Lisa would have a greater moral obligation to children in her hometown if some element of her proximity to these children makes her better able to help them. But this does not undermine my point; it simply clarifies it.

⁵² Slote does not spend much time talking about those with enhanced empathy. Indeed, he spends no time talking about *artificially* enhanced empathy. However, Slote briefly mentions people with “an unusually high degree of empathic concern for others,” and concludes that these people's actions would be considered supererogatory. I find this statement baffling given that Slote argues throughout *The Ethics of Care and Empathy* that moral judgments track degrees of empathy. SLOTE, *supra* note 28, at 34. As I see it, Slote's own argument leads to the conclusion that someone with a greater degree of empathy would have a greater degree of moral responsibility (or, at least, an expanded sphere of moral responsibility). Thus, it would be inappropriate based on Slote's argument to say that the enhanced empath acts supererogatorily. I think my analysis of the enhanced empath's moral responsibility provides another helpful addendum to Slote's overall argument.

obligations in general, it's important to make an explicit principle out of what I have so far only alluded to: prima facie, degrees of ECE track moral responsibility (à la Slote), but the ascription of moral responsibility can be overridden by the PRA. Looking at a few more cases will make the desirability of this principle more clear.

As I mentioned, I think that Lisa's case is a relatively easy one; in other words, it is relatively easy to see that holding her to a new degree of moral responsibility with respect to suffering children in far-off places is reflectively acceptable. On the other side of the coin, there are cases that are easy to see as not being reflectively acceptable. We've already encountered such a case in the discussion of the ordinary empath's moral responsibility. Recall the example of the child who feels ECE towards her abusive parents. Despite this ECE, we concluded that it would not be reflectively acceptable to hold the child responsible for caring for her parents in old age. Similarly, we could say that a person who underwent the empathy enhancement procedure and suddenly felt a newfound sense of ECE towards her abusive parents would not suddenly become morally responsible for taking care of them in old age. The two cases discussed so far are what I would call easy cases. Let's move on to a few cases where it is more difficult to determine whether holding the enhanced empath to a higher moral standard would be reflectively acceptable.

Hard Cases

Reconsider the example of the teenage child who is drowning out in the ocean. If someone other than the child's parent risks her life and limb by swimming out to save the child, this act would likely be considered supererogatory. On other hand, as I've argued, there is an expectation that a parent will risk life and limb to save her child, so when she

swims out to save her child she has fulfilled a moral duty rather than engaged in a supererogatory act. Again, the expectation that the parent will have a greater sense of ECE with her child is at least a partial driver of this distinction. But now imagine that one of the onlookers on the shore has undergone the empathy enhancement procedure. As she sees the child drowning out in the ocean she feels the same strength of ECE towards the child as does the child's parent. As such, if this enhanced empath does not act to save the child, we might want to hold her in moral disdain in the same way that we would hold the child's parent in moral disdain for letting her child drown. But would this judgment be justified?

Like in the case of suffering children in far-off places, we clearly place value on the rescue of nearby children in peril. And we allocate moral responsibility for nearby children in peril (at least in part) on the basis of degrees of ECE between the child in need and the would-be rescuer. Thus, there is a *prima facie* case for saying that the enhanced empath would have a moral responsibility—not just a supererogatory choice—to save the drowning child in the example above (assuming the enhanced empath were physically capable). The question is whether this particular case of holding the enhanced empath morally responsible is reflectively acceptable. Here, it is not so clear. There are several factors at play in this case that distinguish it from the two previous cases. First, there are considerations that come into play when considering the relationship between a parent and a child that are not present in the relationship between two strangers. The specific issues involved are beyond the scope of this dissertation, but I want to recognize that there are significant issues that might cause a parent to feel that it is inappropriate for the enhanced empath to take on for herself an equal moral responsibility for the child. In a

sense, the enhanced empath might be viewed as an *empathic interloper*. In the Conclusion of this dissertation I will attempt to sketch an answer to this concern, but for now I simply want to gesture towards the concern and recognize it as such.⁵³

Second, and somewhat relatedly, there might be practical concerns with holding the enhanced empath morally responsible for rescuing the drowning child. Here, I do not mean practical concerns such as the difficulty of determining whether the enhanced empath actually felt sufficient empathy for the child (I have assumed for the purposes of my argument that she would); what I mean is that holding the enhanced empath (or several enhanced empaths) responsible for rescuing the child might have an all-things-considered bad effect on the child. Of course, if the child's parents do not save the child it would be hard to argue that having the enhanced empath save the child would be all-things-considered bad. But if we assume that both the child's parent and the enhanced empath are in a position to save the child there might be something of value lost for the child if we say that the parent who saves the child was acting on a moral duty that is not unique between that parent and the child. In other words, if the child knows that the parent shares her moral responsibility with non-parents such as the enhanced empath the child might end up feeling like the relationship between her and her parent is less special—that it lacks the special value she previously thought it had. As with the first concern I mentioned, I do not intend to fully flesh out this second concern. Rather, I

⁵³ For an overview of some of the issues involved in the allocation of rights and responsibility between parents and children, see Elizabeth Brake, *Parenthood and Procreation* § 5, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Oct. 28, 2013).

simply want to acknowledge that the status of the enhanced empath's moral duty to others will not always be clear-cut.⁵⁴

Conclusion

In this chapter I have made two broad claims. First, that the enhanced empath would not suffer from a lack of free will; and, second, that the enhanced empath would have an expanded sphere of special obligations such that acts that were supererogatory prior to the person undergoing the enhancement procedure become, *prima facie*, morally obligatory following the procedure. I have acknowledged several difficulties with the

⁵⁴ One might also wonder whether my account of the enhanced empath's situation runs up against Bernard Williams' "integrity objection." In brief, the integrity objection is concerned with utilitarianism's requirement that individuals give up their subjective lived experience and take on "the point of view of the universe" when making moral decisions. See HENRY SIDGWICK, *supra* note 39, for a discussion of this requirement of utilitarianism. According to Williams, this requirement is as undesirable as it is impossible. It is simply not possible for any individual to dissociate herself entirely from her lived experience and take a completely objective moral view of the world. See BERNARD WILLIAMS, MAKING SENSE OF ETHICS 169-70 (1995). At the same time, even asking that someone attempt to do so is objectionable because it asks that individuals give up their individual agency and align their interests with some "super agent." See J.J.C. SMART & BERNARD WILLIAMS, UTILITARIANISM: FOR AND AGAINST 116-17 (1973). Thus, the question becomes whether the enhanced empath is placed (whether by choice or otherwise) in the position of giving up her individual agency and becoming simply a conduit for some objective morality. I think not. I have not argued in this dissertation that the enhanced empath would automatically become a utility machine—concerned only with increasing the overall happiness in the world. Rather, I take the view—advocated by Williams—that the enhanced empath would simply be subject to an enhanced "commonsense impartiality" (i.e., the way that most of us frequently think about what others are thinking and feeling but without systemizing this feeling into an ethical theory that requires us, at all times, to abstract from our individual experiences and take "the point of view of the universe"). See Sophie Grace Chappell, *Bernard Williams*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Spring 2015), <http://plato.stanford.edu/archives/spr2015/entries/williams-bernard/>, for a discussion of this point. Indeed, I would posit that the enhanced empath would find herself at odds with the dictates of utilitarianism, which abstracts away from individuals, whereas the enhanced empath would be greatly concerned with individual lives rather than lives as some abstract utility tally.

latter claim. In the Conclusion of this dissertation I will take a stab addressing one of these difficulties, namely, that the enhanced empath might be considered an empathic interloper.

CONCLUSION

I have had four principle aims in this dissertation. First, I have argued that in order to achieve the goal of an “expanding circle” of care for all human beings, it will be necessary to use some form of artificial empathy enhancement. Evolution has simply not provided us with the capacity to extend the reach of our ECE such that we feel motivated to help those in far-off places to the same degree we feel motivated to help those near to us. Within this context, I have tried to show that empathy enhancement is 1) a reasonably foreseeable possibility within the next few decades or so, and 2) morally defensible. Of course, this line of argument depends on the fact that ECE motivates us in ways that RCE—or reason alone—simply cannot.

Thus, second, I have attempted to show that empathy is best thought of as a two-part process. The first component is what I have called the rational component of empathy. RCE is necessary for moral responsibility as it allows us to put ourselves into another’s shoes and to realize that we would want help (or not to be harmed) if we were in the other’s place. The second component of empathy is what I have called the emotive component. ECE is usually an automatic response to witnessing others in distress, though it can be triggered by cognitive processes. Expanding on Michael Slote’s view, I have argued that it is ECE that varies in strength depending on our relationship to specific people.

Third, relying on the distinction between RCE and ECE I have tried show how philosophers who argue that psychopaths are not morally responsible for their actions are mistaken. As I see it, these philosophers have erred in treating empathy as a singular

concept and concluding that because psychopaths lack empathy they cannot be held morally responsible for their actions. I have argued that the distinction between RCE and ECE allows us to say that psychopaths lack one component of empathy—ECE—but are still responsible for their actions because they clearly have a functional RCE.

Fourth, I have tried to paint a tentative portrait of the landscape of responsibility with respect to the enhanced empath. In that vein, I have argued that the enhanced empath would be subject to an expanded sphere of special obligations such that acts that were supererogatory prior to the person undergoing the enhancement procedure become, *prima facie*, morally obligatory following the procedure.

Without getting back into the specific arguments that I have presented throughout this dissertation, I would like to use this concluding chapter to take a tentative stab at addressing one of the concerns I raised in Chapter Six, namely, that the enhanced empath might be seen as a interloping empath. I will explain what I mean by that term in what follows. As a warning, I will not be presenting anything like a knockdown argument here. Rather, I hope to briefly develop what I believe could be a satisfactory response to the concern regarding the interloping empath.

To frame the issue of the interloping empath, let me begin with a hypothetical. Any soldier who has served in a wartime situation will tell you that there is a powerful and distinct bond that forms between members of a squad or platoon.¹ Indeed, the armed forces go out of their way to instill this bond through the rigorous training that soldiers

¹ See LEONARD WONG ET AL., STRATEGIC STUDIES INSTITUTE, U.S. ARMY WAR COLLEGE, WHY THEY FIGHT: COMBAT MOTIVATION IN THE IRAQ WAR 2 (2003) (quoting S.L.A. MARSHALL, MEN AGAINST FIRE 42-43 (1947)) (“When soldiers were asked what kept them going during the war . . . the second most common response and the primary combat motivation . . . referred to the strong group ties that developed during combat.”).

have to endure before heading to the theater of war. The shared experience of suffering combined with overcoming enormous obstacles creates a tightly knit group of individuals.² It's not uncommon to hear soldiers say that the friends they made during their time in the armed services are like "brothers" and that other relationships pale in comparison.³ When a fellow soldier is killed by an enemy, there is an enormous sense of loss.⁴

I don't think it's going out on a limb to say that the training, suffering, overcoming of obstacles, and facing death on a daily basis while fighting an enemy imbues members of a squad or platoon with an incredibly strong sense of ECE towards one another. Members of a squad are willing to lay down their lives for one another quickly and easily, in part, because they see themselves as part of a collective and the suffering of their fellows becomes their own suffering in turn. In this vein, we often hear stories of soldiers who throw themselves onto grenades or other explosive devices, sacrificing themselves to save the other members of their squad.⁵ And while there is no law that requires one soldier to sacrifice her life to save another, there is surely a shared

² See *id.* at 12 ("Once deployed, soldiers spent more time together training. As one soldier noted, . . . 'it is like we are brothers. Suffered through it all together.'").

³ *Id.* See also LORI HOLYFIELD, VETERANS' JOURNEY HOME: LIFE AFTER AFGHANISTAN AND IRAQ 187 (2011).

⁴ For a compelling take on the sense of loss felt when a fellow soldier dies in combat, see RESTREPO (Nat'l Geographic 2010).

⁵ E.g., Sgt. Herbert J. Thomas, Jr. (WWII); Private First Class James Anderson, Jr. (Vietnam); Master-at-Arms Michael A. Monsoor (Iraq); Corporal Kyle Carpenter (Afghanistan). See *Falling on a Grenade — Notable Examples*, WIKIPEDIA, https://en.wikipedia.org/wiki/Falling_on_a_grenade#cite_note-1 (last visited Jan. 21, 2016).

expectation within the group that no soldier will leave her fellow squad members behind. Imagine if a grenade landed next to a particular soldier, and she were the first to notice, but instead of jumping on the grenade to save other squad members in the vicinity she ran quickly away leaving her fellow soldiers who were unable to react as quickly to die. The soldier who ran away would almost surely feel a sense of guilt after the incident, and her fellow soldiers may even look down upon her.

Unquestionably, if a soldier flees a battle situation, leaving her fellow soldiers to fend off the enemy without her, she would be subject to harsh moral rebuke. Of course, she would also be subject to court martial, but I think it's safe to say that the condemnation she received from her fellow squad members would come primarily from the fact that she violated her moral obligation based on a shared empathic bond within the group, and not from the fact that she violated military code.

Counter this with a journalist who has been recently embedded with a squad. She would not share the same empathic bond with the group, and the soldiers would be unlikely to find her morally lacking if she were to flee in the face of attack. And if she did stay and help the soldiers defend their position, this act would be considered supererogatory and she would likely be lauded as a hero. Significantly, this isn't necessarily due to her status as a non-soldier. If the journalist were embedded with the squad for several months and shared in the empathic bond of the group, she may still be called a coward if she used her non-soldier status to justify abandoning the group when they came under fire.

Now imagine that an enhanced empath is introduced into the squad. Despite not being a part of the historical pattern that has led to the empathic bond of the group, her enhanced capacity for ECE gives her an immediate in. She feels instantaneously immersed in the shared sense of ECE that permeates the squad. Unlike the embedded reporter, who might initially feel aloof from the group or in awe of their bond that appears to be stronger than friendship or even family, the enhanced empath would feel herself at home in a way that she probably had not felt before, because the nature of the squad has brought out a sort of natural empathy enhancement within each of the members. The primary difference between the members of the squad and the enhanced empath is that the squad's enhanced ECE is restricted to that domain—it is a product of that domain. When the soldiers return home they do not carry that feeling of enhanced ECE with them to the other domains of their lives. And, indeed, the same conditions that lead to the squad members' shared bond creates a collateral kind of *diminished* ECE with the enemy.

In any case, the enhanced empath would fit into the group in a way that other outsiders would not. As such, it might seem fair—and even proper—for the other members of the squad to hold her to the same moral standards to which they are all bound. She could be expected to not abandon the squad when the enemy attacked. If a grenade fell into the camp and she ran instead of warning the others, she could be subject to the same opprobrium as if she were a fellow soldier. In other words, like the other members of the squad, acts that would be supererogatory for outsiders become, for her, morally obligatory.

I'm sure that some will scoff at this idea. The counterargument might be that the soldiers would not know of the enhanced empath's unique exaptation and thus would not hold her to the moral standards of the group. Or that even if the soldiers did know that she was an enhanced empath they would consider her an interloping empath whose shared feeling of empathy is inappropriate given her lack of historical connection with the squad.

I take these criticisms seriously. However, I have tried to make a strong argument that a significant part of the basis for how we evaluate whether an act is supererogatory or morally obligatory is the empathic connection between subjects. If a shared sense of ECE is what ties the soldiers in the squad together, and what transforms acts that would otherwise be supererogatory into morally obligatory acts, then it would be irrational to not hold the enhanced empath to the same standards if she shared in the empathic connection of the group. I propose that if the soldiers were reluctant to accept the interloping empath, it would only be due to the soldiers' unfamiliarity with an artificially enhanced human. Indeed, this is likely to be generally the case when enhanced empaths first make their appearance on the scene.

Recall the example of the enhanced empath witnessing a teenage child drowning at sea from Chapter Six. As I noted, there may very well be a similar feeling that the person is an empathic interloper. But, again, I think that this skepticism is likely to be a result of peoples' lack of experience with enhanced empaths. We are accustomed to thinking of a child's parents as having a special empathic relationship with the child; thus, we feel comfortable holding the parents to a higher degree of moral responsibility with respect to their children. The enhanced empath would represent a new quantity in

our understanding of how we should allocate degrees of moral responsibility. The simple fact that enhanced empathes are new to the scene, however, should not be a reason to discount our attributions of moral responsibility as we have in the past (as I have argued) according to degrees of ECE. To deny the enhanced empath the same moral standing—even if it is a standing that requires *more* of the person—as those with naturally enhanced empathy (e.g., parents, soldier in a squad, etc.) would be to engage in a form of bigotry.

Consider an analogy. Transgender men and women in the United States have, in the past and into the present, faced hatred, violence, and discrimination by a significant portion of the population.⁶ Indeed, transgender individuals have faced opposition not only from the straight community but, at times, from the gay and lesbian community as well.⁷ One explanation for this bigotry towards transgender individuals is that those individuals are viewed as interlopers by those whose gender identity matches their birth sex assignment. Jody Norton has argued that the reaction to transgender individuals results from a feeling that these individuals “threat[en] . . . the . . . stability of hegemonic gender.”⁸ The bigot might feel that the transgender individual has no right to the accouterments of a gender other than that which conforms with the individual’s assigned birth sex. Indeed, we see this kind of talk from bigots. They talk about how transgender

⁶ See generally Rebecca L. Stozar, *Violence Against Transgender People: A Review of United States Data*, 14 *AGGRESSION & VIOLENT BEHAV.* 170 (2009).

⁷ See GENNY BEEMYN, *TRANSGENDER HISTORY IN THE UNITED STATES* 23 (2016) (“Transgender women often faced rejection . . . from members of lesbian organizations . . . many of whom viewed them not as “real women” but as “male infiltrators.”).

⁸ Jody Norton, “*Brain Says You’re a Girl, But I Think You’re a Sissy Boy*”: *Cultural Origins of Transphobia*, 2 *J. GAY, LESBIAN, & BISEXUAL IDENT.* 139, 140 (1997).

women have no right to dress like women or be treated like women, and vice versa.⁹ They talk about how transgender men have no right to the benefits—or the determinants—of being a man.¹⁰

We can find parallels with this kind of attitude in other contexts as well. White Americans in the past (and some still in the present, unfortunately) condemned African Americans as claiming rights to things which were solely within the province of white Americans. Part of that rhetoric stemmed from a feeling that African Americans were not part of the historical story that was the story of the white race—a story of progress and society and property ownership, etc.¹¹

With transgender individuals, African Americans, and other historically oppressed groups, part of society's coming to accept them as equals is society coming to accept that their experiences—even if disconnected from the historical context that typifies “traditional” notions of gender identify, racial histories, etc.—are equally valid. A large part of coming to this understanding is simply allowing these groups to have equal

⁹ For a particularly misinformed example of this kind of talk, see Matt Walsh, *Calling Bruce Jenner a Woman is an Insult to Women*, THE BLAZE (June 2, 2015), <http://www.theblaze.com/contributions/calling-bruce-jenner-a-woman-is-an-insult-to-women/> (“[A] man can never lay claim to womanhood.”).

¹⁰ See Tarynn M. Witten & A. Evan Eyler, *Hate Crimes and Violence Against the Transgendered*, 11 PEACE REV. 461, 467-68 (1999).

¹¹ Indeed, this view was supported by the United States Supreme Court in *Dred Scott v. Sandford*, where Chief Justice Taney stated that “negroes” are “so far inferior, that they had no rights which the white man was bound to respect.” 60 U.S. 393, 407 (1856). If not quite explicitly, Taney made clear that African Americans had no right to the benefits conferred upon whites by the Constitution when he said, “No one of that race [African American] ever migrated to the United States voluntarily; all of them had been brought here as articles of merchandise.” *Id.* at 411. “It is obvious that they were not even in the mind of the framers of the Constitution when they were conferring special rights and privileges upon the citizens of a State[.]” *Id.* at 411-12.

rights—and equal responsibilities—within society. As African Americans were granted more rights, the number of people who viewed them as inferior, or not deserving of America’s benefits, diminished.¹² We have seen a similar phenomenon with transgender individuals—and certainly with gay and lesbian individuals.¹³

Returning to the experience of the enhanced empath, it would not be surprising to see—as I previously described—people saying that the enhanced empath is an interloper. As someone who is not privy to the historical relationships and connections that form the building blocks of our willingness to hold certain people to higher moral standards, the enhanced empath might be seen as someone who is claiming ownership of something that she does not deserve, namely, the right to be held to a higher moral standard. I hope that my analogy to transgender individuals and African Americans provides support for my claim that the enhanced empath has just as much right to her feelings of enhanced empathy as does a mother towards her child or a soldier towards her squad.

As noted earlier, I have only tried to provide a sketch of a response to the concern of the interloping empath here. A more fully developed response would require, at least, addressing each of the concerns associated with the importance of parental rights and responsibilities and considering their respective importance as compared to the

¹² E.g., public support for equal rights for African-Americans increased from 49% to 60% within one year after the passage of the Civil Rights Act of 1964. See GEORGE GALLUP, THE GALLUP POLL: PUBLIC OPINION 1999 237 (2000). When a new voting rights act was proposed for 1965, 76% of the public approved of the measure. See Andrew Kohut, *50 Years Ago: Mixed Views About Civil Rights but Support for Selma*, PEW RES. CTR. (Mar. 5, 2015), <http://www.pewresearch.org/fact-tank/2015/03/05/50-years-ago-mixed-views-about-civil-rights-but-support-for-selma-demonstrators/>.

¹³ See Susan Page, *Poll: Support for Gay Marriage Hits High After Ruling*, USA TODAY (July 1, 2013), <http://www.usatoday.com/story/news/politics/2013/07/01/poll-supreme-court-gay-marriage-affirmative-action-voting-rights/2479541/>.

importance of empathic connection in determining responsibility. Only then could we be fully secure that holding the enhanced empath to a higher standard of responsibility in these situations would be reflectively acceptable. But that is a project for another time.

Hopefully, if philosophers such as Slote, among others, continue to focus on the importance of empathy in moral theory and responsibility we will come to a more complete understanding of the world of the enhanced empath before that world comes into being.

REFERENCES

- 1982 World Charter for Nature, G.A Res. 37/7, U.N. Doc. A/RES/37/7 (Oct. 28, 1982).
- Adam Deming, *Sex Offender Commitment Programs: Current Practices, Characteristics, and Resident Demographics*, 36 J. PSYCHIATRY & L. 439 (2008).
- Albin Eser, *The Principle of "Harm" in the Concept of Crime: A Comparative Analysis of the Criminally Protected Legal Interests*, 4 DUQ. L. REV. 345 (1965).
- ALDOUS HUXLEY, *BRAVE NEW WORLD* (1931).
- Alexandra Minna Stern & Howard Markel, *The History of Vaccines and Immunization Familiar Patterns, New Challenges*, 24 HEALTH AFFAIRS 611 (2005).
- Alice G. Walton, *The Chicken Pox Party*, THE ATLANTIC (Nov. 25, 2011), <http://www.theatlantic.com/health/archive/2011/11/the-chicken-pox-party-parents-caught-infecting-kids-with-virus/248768/>.
- Alistair Burns & Steve Iliffe, *Alzheimer's Disease*, 338 BRIT. MED. J. 467 (2009).
- ALISTAIR MUNRO, *DELUSIONAL DISORDER: PARANOIA AND RELATED ILLNESSES* (1999).
- ALLEN BUCHANAN ET AL., *FROM CHANCE TO CHOICE* (2000).
- ALLEN E. BUCHANAN, *BEYOND HUMANITY?: THE ETHICS OF BIOMEDICAL ENHANCEMENT* (2011).
- Alvaro Pascual-Leone et al., *The Plastic Human Brain Cortex*, 28 ANN. REV. NEUROSCIENCE 377 (2005).
- Alvin I. Goldman, *Interpretation Psychologized*, 4 MIND & LANGUAGE 161 (1989).
- American Psychiatric Association, *American Psychiatric Association Statement on the Insanity Defense*, 140 AM. J. PSYCHOL. 681 (1983).
- AMERICAN PSYCHIATRIC ASSOCIATION, *DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS* (1952).
- AMERICAN PSYCHIATRIC ASSOCIATION, *DSM V: DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS* (2013).
- AMERICAN PSYCHIATRIC ASSOCIATION, *DSM-II: DIAGNOSTIC AND STATISTICAL MANUAL OF MENTAL DISORDERS* (1968).

AMERICAN PSYCHIATRIC ASSOCIATION, *STATISTICAL MANUAL FOR THE USE OF INSTITUTIONS FOR THE INSANE* (U. Mich. 2008) (1918).

An Overview of the Human Brain Project, HUMAN BRAIN PROJECT, <https://www.humanbrainproject.eu/discover/the-project/overview> (last visited Jan. 14, 2016).

Andrew Johnson, *The Effects of DBS on the Motor Systems of Parkinson's Disease*, YOUTUBE (June 12, 2013), <https://www.youtube.com/watch?v=uBh2LxTW0s0>.

Andrew K. Chan et al., *Disparities in Access to Deep Brain Stimulation Surgery for Parkinson Disease*, 71 JAMA NEUROLOGY 291 (2014).

Andrew Kohut, *50 Years Ago: Mixed Views About Civil Rights but Support for Selma*, PEW RES. CTR. (Mar. 5, 2015), <http://www.pewresearch.org/fact-tank/2015/03/05/50-years-ago-mixed-views-about-civil-rights-but-support-for-selma-demonstrators/>.

Andrew Koivuniemi & Kevin Otto, *When "Altering Brain Function" Becomes "Mind Control,"* 8 FRONTIERS SYS. NEUROSCIENCE 1 (2014).

ANGELA HOLDER, *LEGAL ISSUES IN PEDIATRICS AND ADOLESCENT MEDICINE* (2d ed. 1985).

ANN RULE, *THE STRANGER BESIDE ME* (Pocket Books 2008) (1981).

ANTHONY BURGESS, *A CLOCKWORK ORANGE* (1962).

Anthony Zador, *The Connectome as a DNA Sequencing Problem*, in *THE FUTURE OF THE BRAIN: ESSAYS BY THE WORLD'S LEADING NEUROSCIENTISTS* (Gary Marcus & Jeremy Freeman eds., 2015).

Antonella Benvenuti et al., *Treatment-Emergent Mania/Hypomania in Unipolar Patients*, 10 BIPOLAR DISORDER 726 (2008).

Atif Kukaswadia, *History of Epidemiology: Jonas Salk and the Eradication of Polio*, PUBLIC HEALTH PERSPECTIVES (June 4, 2013), <http://blogs.plos.org/publichealth/2013/06/04/history-of-epidemiology-jonas-salk-and-the-eradication-of-polio/>.

Benjamin D. Greenberg et al., *Three-Year Outcomes in Deep Brain Stimulation for Highly Resistance Obsessive-Compulsive Disorder*, 31 NEUROPSYCHOPHARMACOLOGY 2384 (2006).

BERNARD WILLIAMS, *MAKING SENSE OF ETHICS* (1995).

BlueCross BlueShield of Alabama, *Deep Brain Stimulation*, BCBSAL.ORG (Nov. 2014), <https://www.bcbsal.org/providers/policies/final/347.pdf>.

Box A: Anti-vaccination arguments, past and present, THE BMJ, <http://www.bmj.com/content/suppl/2002/08/22/325.7361.430.DC1>.

BRET EASTON ELLIS, *AMERICAN PSYCHO* (1991).

Brian Deer, *Exposed: Andrew Wakefield and the MMR-Autism Fraud*, BRIANDEER.COM, <http://briandeer.com/mmr/lancet-summary.htm> (last visited Oct. 21, 2015).

BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, PROBATION AND PAROLE IN THE UNITED STATES, 2014 (2014), <http://www.bjs.gov/content/pub/pdf/ppus14.pdf>.

C. Daniel Batson, Shannon Early, & Giovanni Salvarani, *Perspective Taking: Imagining How Another Feels Versus Imagining How You Would Feel*, 23 PERS. & SOC. PSYCHOL. BULL. 751 (1997).

C. Hamani et al., *Memory Enhancement Inducted by Hypothalamic/Fornix Deep Brain Stimulation*, 63 ANNALS OF NEUROLOGY 119 (2008).

C.A. Davie, *A Review of Parkinson's Disease*, 86 BRIT. MED. BULL. 109 (2008).

Caroline Cassels, *FDA Clears TMS Device for Resistant Depression*, MEDSCAPE (Jan. 9, 2013), <http://www.medscape.com/viewarticle/777357>.

Cass R. Sunstein, *The Paralyzing Principle*, 25 REGULATION MAGAZINE 32 (2003).

Cass Sunstein, *Cost-Benefit Analysis and the Environment*, 115 ETHICS 351 (2005).

Cass Sunstein, *The Limits of Quantification*, 102 CAL. L. REV. 1369 (2014).

Chandra Kiran & Suprakash Chaudhury, *Understanding Delusions*, 18 INDUS. PSYCHOL. 3 (2009).

CHARLES PATRICK EWING & JOSEPH T. MCCANN, *MINDS ON TRIAL: GREAT CASES IN LAW AND PSYCHOLOGY* (2006).

Christiana M. Leonard et al., *Normal Variation in the Frequency and Location of Human Auditory Cortex Landmarks. Heschl's Gyrus: Where Is It?*, 8 CEREBRAL CORTEX 397 (1998).

Christopher J. Ferguson et al., *Defining and Classifying Serial Murder in the Context of Perpetrator Motivation*, 31 J. CRIM. JUST. 287 (2003).

Clyde Haberman, *A Discredited Vaccine Study's Continuing Impact on Public Health*, NY TIMES, Feb. 1, 2015, http://www.nytimes.com/2015/02/02/us/a-discredited-vaccine-study-continuing-impact-on-public-health.html?_r=0.

Consolidated Version of the Treaty of the Functioning of the European Union, art. 191, Sept. 5, 2008 O.J. (C 115) 47.

CONVERGING TECHNOLOGIES FOR IMPROVING HUMAN PERFORMANCE: NANOTECHNOLOGY, BIOTECHNOLOGY, INFORMATION TECHNOLOGY, AND COGNITIVE SCIENCE (NBIC) (Mihail C. Roco & William Sims Bainbridge eds., 2003).

Cristina Cusin & Darin D. Dougherty, *Somatic Therapies for Treatment-Resistance Depression: ECT, TMS, VNS, DBS*, 14 BIOLOGY OF MOOD & ANXIETY DISORDERS 1 (2012).

D. Cobb-Clark & S. Schurer, *The Stability of Big-Five Personality Traits*, 115 ECON. LETTERS 11 (2012).

Daniel C. Murrie et al., *Field Validity of the Psychopathy Checklist-Revised in Sex Offender Risk Assessment*, 24 PSYCHOL. ASSESSMENT 524 (2012).

DANIEL DENNETT, *DARWIN'S DANGEROUS IDEA* (1995).

DAVID A. BAUM ET AL., *THE PRINCETON GUIDE TO EVOLUTION* (2013).

DAVID J. CHALMERS, *THE CHARACTER OF CONSCIOUSNESS* (2010).

DAVID J. CHALMERS, *THE CONSCIOUS MIND: IN SEARCH OF A FUNDAMENTAL THEORY* (1996).

David DeMatteo et al., *The Role and Reliability of the Psychopathy Checklist—Revised in U.S. Sexually Violent Predator Evaluations: A Case Law Survey*, 38 LAW & HUM. BEHAV. 248 (2013).

DAVID HUME, *A TREATISE OF HUMAN NATURE* (Everyman 2003) (1738).

David Post, *Minnesota's Egregious Sex Offender Confinement Statute Held Unconstitutional*, VOLOKH CONSPIRACY (June 18, 2015), <https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/06/18/minnesotas-egregious-sex-offender-confinement-statute-held-unconstitutional/>.

Death With Dignity Act: Frequently Asked Questions, OREGON.GOV, <https://public.health.oregon.gov/ProviderPartnerResources/EvaluationResearch/DeathwithDignityAct/Pages/faqs.aspx#present> (last visited Jan. 12, 2016).

Diane Jeske, *Special Obligations*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Jan. 5, 2014), <http://plato.stanford.edu/entries/special-obligations/>.

Dred Scott v. Sanford, 60 U.S. 393, 407 (1856).

Drug Use Changes the Brain Over Time, U. UTAH HEALTH SCI., <http://learn.genetics.utah.edu/content/addiction/brainchange/> (last visited Jan. 14, 2016).

E. ANN CARSON & WILLIAM J. SABOL, BUREAU JUST. STAT., PRISONERS IN 2011 (2012), <http://www.bjs.gov/content/pub/pdf/p11.pdf>.

E.S. VALENSTEIN, BRAIN CONTROL (1973).

EARL HUNT, HUMAN INTELLIGENCE (2011).

Ed Bullmore & Paul Fletcher, *The Eye's Mind: Brain Mapping and Psychiatry*, 182 BRIT. J. PSYCHIATRY 381 (2003).

EDWARD O. WILSON, SOCIOBIOLOGY (1980).

Elizabeth Allen et al., *Against "Sociobiology,"* NY REVIEW OF BOOKS (Aug. 7, 1975), <http://www.nybooks.com/articles/archives/1975/nov/13/against-sociobiology/>.

Elizabeth Brake, *Parenthood and Procreation*, STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Oct. 28, 2013).

ERIC R. KANDEL, JAMES H. SCHWARTZ & THOMAS M. JESSEL, PRINCIPLES OF NEURAL SCIENCE (4th ed. 2000).

ERICH GOODE, DEVIANT BEHAVIOR (2015).

Fact Sheet: BRAIN Initiative, WHITE HOUSE OFFICE OF THE PRESS SECRETARY (Apr. 2, 2013), <https://www.whitehouse.gov/the-press-office/2013/04/02/fact-sheet-brain-initiative>.

Falling on a Grenade — Notable Examples, WIKIPEDIA, https://en.wikipedia.org/wiki/Falling_on_a_grenade#cite_note-1 (last visited Jan. 21, 2016).

Filippo Brighina et al., *Reduced Cerebellar Inhibition in Migraine with Aura: a TMS Study*, 8 CEREBELLUM 260 (2009).

Frederick L. Coolidge & Daniel L. Segal, *Evolution of Personality Diagnosis in the Diagnostic and Statistical Manual of Mental Disorders*, 18 CLINICAL PSYCHOL. REV. 585, (1998).

Fritz Allhoff et al., *Ethics of Human Enhancement: 25 Questions & Answers*, 4 STUDIES IN ETHICS, LAW, AND TECHNOLOGY 1 (2010).

GENNY BEEMYN, TRANSGENDER HISTORY IN THE UNITED STATES (2016).

George E. Partridge, *Current Conceptions of Psychopathic Personality*, 87 AM. J. PSYCHOL. 53 (1930).

GEORGE GALLUP, THE GALLUP POLL: PUBLIC OPINION 1999 (2000).

Gilberto Pizzoloato & Tomasz Mandat, *Deep Brain Stimulation for Movement Disorders*, 6 FRONTIERS IN INTEGRATIVE NEUROSCIENCE 2 (2012).

Ginny A. Roth, *Celebrating Salk*, NIH (March 26, 2014), <http://circulatingnow.nlm.nih.gov/2014/03/26/the-first-polio-pioneer/>.

Glenn D. Walters, *The Trouble with Psychopathy as a General Theory of Crime*, 48 INT. J. OFFENDER THERAPY COMP. CRIM. 133 (2004).

Heather Mitchell & Michael G. Aamodt, *The Incidence of Child Abuse in Serial Killers*, 20 J. POLICE & CRIM. PSYCHOL. 40 (2005).

Heidi L. Maibom, *(Almost) Everything You Ever Wanted to Know About Empathy*, in EMPATHY AND MORALITY (Heidi L. Maibom ed., 2014).

Heidi Malm, *Killing, Letting Die, and Simply Conflicts*, 18 PHIL. & PUB. AFF. 238 (1989).

Helen S. Mayberg et al., *Deep Brain Stimulation for Treatment-Resistance Depression*, 45 NEURON 651 (2005).

HENRY SIDGWICK, THE METHODS OF ETHICS (Hackett 1981) (1907).

HERVEY M. CLECKLEY, THE MASK OF SANITY (5th ed. 1982).

History of Anti-Vaccination Movements, HISTORYOFVACCINES.ORG (Dec. 18, 2014), <http://www.historyofvaccines.org/content/articles/history-anti-vaccination-movements#Source 8>.

Insanity Defense Reform Act, 18 U.S.C. § 17 (2006).

Irving Kirsch et al., *Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration*, 5 PLOS MED. 260 (2008).

J Leask & S Chapman, *'An Attempt to Swindle Nature': Press Anti-Immunisation Reportage 1993-1997*, 22 AUSTL. & N.Z. J. PUB. HEALTH 17 (1998).

- J.J.C. SMART & BERNARD WILLIAMS, *UTILITARIANISM: FOR AND AGAINST* (1973).
- J.O. Urmson, *Saints and Heroes*, in *ESSAYS IN MORAL PHILOSOPHY* (A.I. Melden ed., 1958).
- Jacobson v. Massachusetts, 197 U.S. 11 (1905).
- JAMES COWLES PRICHARD, *A TREATISE ON INSANITY AND OTHER DISORDERS AFFECTING THE MIND* (1837).
- JAMES F. STEPHEN, *A HISTORY OF THE CRIMINAL LAW OF ENGLAND* (1883).
- Jean-Louis Monestes, *Human Evolution and Contextual Behavioral Science*, in *THE WILEY HANDBOOK OF CONTEXTUAL BEHAVIORAL SCIENCE* (Robert D. Zettle et al., eds., 2016).
- Jeffrey M. Burns & Russell H. Swerdlow, *Right Orbitofrontal Tumor with Pedophilia Symptom and Constructional Apraxia Sign*, 60 *ARCHIVES NEUROLOGY* 437 (2003).
- Jeffrey S. Gerber & Paul A. Offit, *Vaccines and Autism: A Tale of Shifting Hypotheses*, 48 *CLINICAL INFECTIOUS DISEASES* 456 (2009).
- JEFFRIE G. MURPHY, *KANT: THE PHILOSOPHY OF RIGHT* (1994).
- Jeffrie G. Murphy, *Moral Death: A Kantian Essay on Psychopathy*, 82 *ETHICS* 284 (1972).
- JEFFRIE MURPHY, *RETRIBUTION, JUSTICE, AND THERAPY* (1979).
- Jennifer L. Skeem & David J. Cooke, *Is Criminal Behavior a Central Component of Psychopathy? Conceptual Directions for Resolving the Debate*, 22 *PSYCHOL. ASSESSMENT* 433 (2010).
- Jody Norton, *“Brain Says You’re a Girl, But I Think You’re a Sissy Boy”*: *Cultural Origins of Transphobia*, 2 *J. GAY, LESBIAN, & BISEXUAL IDENT.* 139 (1997).
- John A. Bargh & Ezequiel Morsella, *The Unconscious Mind*, 3 *PERSP. PSYCHOL. SCI.* 73, 76 (2008).
- JOHN HARRIS, *ENHANCING EVOLUTION: THE ETHICAL CASE FOR MAKING BETTER PEOPLE* (2010).
- John M. Digman, *Personality Structure: Emergence of the Five-Factor Model*, 41 *ANN. REV. PSYCHOL.* 417 (1990).

JOHN MARTIN FISCHER & MARK RAVIZZA, RESPONSIBILITY AND CONTROL (1998).

JOHN RAWLS, A THEORY OF JUSTICE (1999).

Jones v. United States, 308 F.2d 307 (1962).

Joseph Mercola, *Flu Vaccines: The Horrible “Immune System Mistake” Millions Will Make This Year*, MERCOLA.COM (Jan. 3, 2012), <http://articles.mercola.com/sites/articles/archive/2012/01/03/flu-shot-increase-flu.aspx>.

JOSHUA DRESSLER & STEPHEN P. GARVEY, CASES AND MATERIALS ON CRIMINAL LAW (6th ed. 2012).

JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW (7th ed. 2015).

Julian Lorkin, *Why a Psychopathic Tendency May be Good for Entrepreneurs*, UNIV. N.S.W. AUSTL: BUS. SCHOOL (Aug. 28, 2013), <https://www.business.unsw.edu.au/news-events/news/why-psychopathic-tendency-good-for-entrepreneurs>.

Julian Savulescu, *Justice, Fairness, and Enhancement*, 1093 ANNALS N.Y. ACAD. SCI. 321 (2006).

Julian Savulescu, *Procreative Beneficence: Why We Should Select the Best Children*, 15 BIOETHICS 413 (2001).

Karen Lazo, *Huge Brain Database is Revealing Secrets of the Mind*, LIVE SCIENCE (July 10, 2015), <http://www.livescience.com/51523-massive-brain-scan-database-goes-online.html>.

Karsjens v. Jesson, No. 11-3659, 2015 WL 3755870 (D. Minn. June 15, 2015).

Kate Kelland, *Biology Gives American Psychopaths a Legal Break*, REUTERS (Aug. 16, 2012), <http://www.reuters.com/article/2012/08/16/us-science-psychopaths-biology-idUSBRE87F12M20120816>.

LEONARD WONG ET AL., STRATEGIC STUDIES INSTITUTE, U.S. ARMY WAR COLLEGE, WHY THEY FIGHT: COMBAT MOTIVATION IN THE IRAQ WAR (2003).

LINDA CAMERON, ENVIRONMENTAL RISK MANAGEMENT IN NEW ZEALAND – IS THERE SCOPE TO APPLY A MORE GENERIC FRAMEWORK? (2006), <http://www.treasury.govt.nz/publications/research-policy/ppp/2006/06-06/tpp06-06.pdf>.

LIONEL DAHMER, A FATHER’S STORY (1994).

Lisa G. Aspinwall, Teneille R. Brown, & James Tabery, *The Double-Edged Sword: Does Biomechanism Increase or Decrease Judges' Sentencing of Psychopaths?*, 337 SCIENCE 846 (2012).

LORI HOLYFIELD, VETERANS' JOURNEY HOME: LIFE AFTER AFGHANISTAN AND IRAQ (2011).

Lut Tamam & Nurgul Ozpoyraz, *Selective Serotonin Reuptake Inhibitor Discontinuation Syndrome: A Review*, 19 ADVANCES IN THERAPY 17 (2002).

M. Israel et al., *Deep Brain Stimulation in the Subgenual Cingulate Cortex for an Intractable Eating Disorder*, 67 BIOLOGICAL PSYCHIATRY e53 (2010).

Mark S. George et al., *Daily Repetitive Transcranial Magnetic Stimulation (rTMS) Improves Mood in Depression*, 6 NEUROREPORT 1853 (1995).

Mark S. George et al., *Image-Guided Brain Stimulation*, in NEUROBIOLOGY OF MENTAL ILLNESS (Dennis S. Charney et al., eds., 4th ed. 2013).

MARTHA STOUT, THE SOCIOPATH NEXT DOOR (2006).

MARTIN L. HOFFMAN, EMPATHY AND MORAL DEVELOPMENT (2000).

MARY C. BRINTON & VICTOR NEE, THE NEW INSTITUTIONALISM IN SOCIOLOGY (2001).

Matt Walsh, *Calling Bruce Jenner a Woman is an Insult to Women*, THE BLAZE (June 2, 2015), <http://www.theblaze.com/contributions/calling-bruce-jenner-a-woman-is-an-insult-to-women/>.

Matthew T. Gailliot & Roy F. Baumeister, *The Physiology of Willpower: Linking Blood Glucose to Self-Control*, 11 PERS. & SOC. PSYCHOL. REV. 303 (2007).

MAXWELL J. MEHLMAN, TRANSHUMANIST DREAMS AND DYSTOPIAN NIGHTMARES: THE PROMISE AND PERIL OF GENETIC ENGINEERING (2012).

McFall v. Shimp, 10 Pa. D. & C. 3d 90 (1978).

Megan Heimer, *God Does Not Support Vaccines*, LIVING WHOLE (July 7, 2014), <http://www.livingwhole.org/god-does-not-support-vaccines/>.

Melinda Carrido, Comment, *Revisiting the Insanity Defense: A Case for Resurrecting the Volitional Prong of the Insanity Defense in Light of Neuroscientific Advances*, 41 SW. U. L. REV. 309 (2012).

Michael C. Seto & Howard E. Barbaree, *Psychopathy, Treatment Behavior, and Sex Offender Recidivism*, 14 J. INTERPERSONAL VIOLENCE 1235 (1999).

Michael J. Sandel, *The Case Against Perfection*, THE ATLANTIC (April 2004), <http://www.theatlantic.com/magazine/archive/2004/04/the-case-against-perfection/302927/>.

Michael O. Hardimon, *Role Obligations*, 91 J. PHIL. 333 (1994).

MICHAEL SLOTE, THE ETHICS OF CARE AND EMPATHY (2007).

MINN. STATE. ANN.

MODEL PENAL CODE (1962).

Moriel Zelikowsky et al., *Prefrontal Microcircuit Underlies Contextual Learning After Hippocampal Loss*, 110 PROC. NAT'L ACAD. SCI. U.S. 9938 (2013).

NEL NODDINGS, CARING: A FEMINIST APPROACH TO ETHICS AND MORAL EDUCATION (1984).

NICHOLAS AGAR, HUMANITY'S END (2010).

NICHOLAS AGAR, LIBERAL EUGENICS: IN DEFENCE OF HUMAN ENHANCEMENT (2004).

NIH BRAIN Awards, NATIONAL INSTITUTES OF HEALTH, Project Title: Tracing Brain Circuits by Transneuronal Control of Transcription, <http://braininitiative.nih.gov/nih-brain-awards.htm> (last visited Oct. 13, 2015).

Noah M. Sachs, *Rescuing the Strong Precautionary Principle From Its Critics*, 4 U. ILL. L. REV. 1285 (2011).

Norman Daniels, *Normal Functioning and the Treatment-Enhancement Distinction*, 9 CAMBRIDGE Q. OF HEALTHCARE ETHICS 309 (2000).

OLIVER SACKS, AWAKENINGS (1974).

P. Babiak, C.S. Neuman, & RD Hare, *Corporate Psychopathy: Talking the Walk*, 28 BEHAV. SCI. LAW 174 (2010).

Pamela Paul, *'The Sociopath Next Door': Ruthless People*, NY TIMES: SUNDAY BOOK REV. (Mar. 6, 2005), http://www.nytimes.com/2005/03/06/books/review/the-sociopath-next-door-ruthless-people.html?_r=0.

People v. Beardsley, 113 N.W. 1128 (Mich. 1907).

People's Experiences with Involuntary Treatment, SCHIZOPHRENIA.COM,
<http://schizophrenia.com/family/involuntaryb.htm> (last visited Jan. 19, 2016).

Peter Railton, *Alienation, Consequentialism, and the Demands of Morality*, 13 PHIL. & PUB. AFF. 134 (1984).

PETER SINGER, *ANIMAL LIBERATION: A NEW ETHICS FOR OUR TREATMENT OF ANIMALS* (1975).

Peter Singer, *Famine, Affluence, and Morality*, 1 PHIL. & PUB. AFF. 229 (1972).

PETER SINGER, *THE EXPANDING CIRCLE: ETHICS, EVOLUTION, AND MORAL PROGRESS* (2011).

PETER W. LOW, JOHN CALVIN JEFFRIES JR., & RICHARD J. BONNIE, *THE TRIAL OF JOHN W. HINCKLEY, JR.: A CASE STUDY IN THE INSANITY DEFENSE* (1986).

PHILIP K. DICK, *DO ANDROIDS DREAM OF ELECTRIC SHEEP?* (1968).

PLATO, *REPUBLIC*.

POLLY NELSON, *DEFENDING THE DEVIL: MY STORY AS TED BUNDY'S LAST LAWYER* (1994).

R Chen et al., *Wireless Magnetothermal Deep Brain Stimulation*, 347 SCIENCE 1477 (2015).

R. Peter Hobson & Jessica A. Hobson, *On Empathy*, in *EMPATHY AND MORALITY* (Heidi L. Maibom ed., 2014).

R.I. GEN. LAWS.

R.J.R. Blair, *The Amygdala and Ventromedial Prefrontal Cortex: Functional Contributions and Dysfunction in Psychopathy*, 363 PHIL. TRANSACTIONS ROYAL SOC'Y. LONDON B 2557 (2008).

R.M. Chisholm, *Supererogation and Offence: A Conceptual Schema for Ethics*, 5 RATIO 1 (1963).

R.R. McCRAE & P.T. COSTA, *PERSONALITY IN ADULTHOOD* (1990).

Rajesh Pahwa, Kelly E. Lyons, & Jules M. Nazzaro, *Deep Brain Stimulation for Parkinson's Disease*, U. KAN. MED. CTR. 14,
<http://www.kumed.com/~media/Imported/kumed/parkinson/dbs.ashx> (last visited Jan. 16, 2016).

RAY KURZWEIL, *THE SINGULARITY IS NEAR* (2005).

Rebecca L. Stozler, Violence Against Transgender People: A Review of United States Data, 14 *AGGRESSION & VIOLENT BEHAV.* 170 (2009).

RENÉ DESCARTES, *DISCOURSE ON METHOD* (Richard Kennington trans., Focus Publications 2007) (1637).

RESTATEMENT (SECOND) OF TORTS (1965).

RESTREPO (Nat'l Geographic 2010).

RICHARD DAWKINS, *THE BLIND WATCHMAKER* (1986).

RICHARD DAWKINS, *THE SELFISH GENE* (1976).

RICHARD F. WETZELL, *INVENTION OF THE CRIMINAL: A HISTORY OF GERMAN CRIMINOLOGY, 1880-1945* (2000).

Richard W. Eldwood, Dennis D. Doren, & David Thornton, *Diagnostic and Risk Profiles of Men Detained Under Wisconsin's Sexually Violent Person Law*, 54 *INT'L J. OFFENDER THERAPY & COMP. CRIMINOLOGY* 187 (2010).

Roar Fosse & John Read, *Electroconvulsive Treatment: Hypotheses About Mechanisms of Action*, 4 *FRONTIERS IN PSYCHIATRY* 1 (2013).

Robert Axelrod & William D. Hamilton, *The Evolution of Cooperation*, 211 *SCIENCE* 1390 (1981).

Robert D. Hare et al., *The Revised Psychopathy Checklist: Reliability and Factor Structure*, 2 *PSYCHOL. ASSESSMENT* 338 (1990).

Robert D. Hare, *A Research Scale for the Assessment of Psychopathy in Criminal Populations*, 1 *PERS. & INDIVIDUAL DIFFERENCES* 111 (1980).

Robert D. Hare, *Forward*, in *THE CLINICAL AND FORENSIC ASSESSMENT OF PSYCHOPATHY* (C.B. Gacono ed., 2d ed. 2015).

Robert D. Hare, *Psychopathy: A Clinical Construct Whose Time Has Come*, 23 *CRIM. JUST. BEHAV.* 25 (1996).

ROBERT D. HARE, *WITHOUT CONSCIENCE: THE DISTURBING WORLD OF THE PSYCHOPATHS AMONG US* (1999).

Robert Elliot, *Faking Nature*, 25 *INQUIRY* 381 (1982).

Robert L. Trivers, *The Evolution of Reciprocal Altruism*, 46 QUART. REV. OF BIOL. 35 (1971).

Robert M Wolfe & Lisa K Sharp, *Anti-Vaccinationists Past and Present*, 325 BMJ 430 (2002).

RODNEY DOUGLAS, HENRY MARKRAM, & KEVAN MARTIN, NEOCORTEX (2004).

Ron Berghmans et al., *Scientific, Ethical, and Social Issues in Mood Enhancement*, in ENHANCING HUMAN CAPABILITIES (Julian Savulescu, Ruud ter Meulen, & Guy Kahane eds., 2011).

Sarah F. Brosnan & Frans B.M. de Waal, *Evolution of Responses to (Un)Fairness*, 346 SCIENCE 1251776 (2014).

Sarah F. Brosnan & Frans B.M. de Waal, *Monkeys Reject Unequal Pay*, 425 NATURE 297 (2003).

Sarah F. Brosnan, *Justice-and Fairness-Related Behaviors in Nonhuman Primates*, 110 PNAS 10416 (2013).

Schizophrenia Therapy, THRIVEWORKS, <https://thriveworks.com/austin-schizophrenia/> (last visited Jan. 18, 2016).

Shelley D. Piasecki & James W. Jefferson, *Psychiatric Complications of Deep Brain Stimulation for Parkinson's Disease*, 65 J. CLINICAL PSYCHIATRY 845 (2004).

Sophie Grace Chappell, *Bernard Williams*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Spring 2015), <http://plato.stanford.edu/archives/spr2015/entries/williams-bernard/>.

Souped-Up Remote Control Switches Behavior On-and-Off in Mice, NATIONAL INSTITUTES OF HEALTH (Apr. 30, 2015), <http://www.nih.gov/news/health/apr2015/nimh-30.htm>.

Speech by Martin Luther King, Jr., Montgomery, AL, March. 25, 1965.

State *ex. rel.* Kuntz v. Thirteenth Judicial District, 995 P.2d 951 (Mont. 2000).

State v. Sherman, 266 S.W.3d 395 (Tenn. 2008).

Stefanie C. Altieri et al., *Serotonergic Pathways in Depression*, in NEUROBIOLOGY OF DEPRESSION (Francisco Lopez-Munoz & Cecilio Alamo eds., 2011).

Stephan Hamman, *Deep Brain Stimulation for Depression CNN Presents Special with Dr. Sanjay Gupta*, YOUTUBE (May 6, 2012), <https://www.youtube.com/watch?v=Lq5rIILcVgA>.

Stephanie D. Preston & Frans B.M. de Waal, *Empathy: It's Ultimate and Proximate Bases*, 25 BEHAV. & BRAIN SCI. 1 (2002).

Stephanie D. Preston, *A Perception-Action Model of Empathy*, in EMPATHY IN MENTAL ILLNESS (Tom F.D. Farrow & Peter W.R. Woodruff eds., 2007).

Stephen Gardiner, *A Core Precautionary Principle*, 14 THE JOURNAL OF POLITICAL PHILOSOPHY 33 (2006).

STEPHEN JAY GOULD, FULL HOUSE (1996).

STEPHEN JAY GOULD, WONDERFUL LIFE (1989).

Stephen Morse, *Psychopathy and Criminal Responsibility*, 1 NEUROETHICS 205 (2008).

Susan Page, *Poll: Support for Gay Marriage Hits High After Ruling*, USA TODAY (July 1, 2013), <http://www.usatoday.com/story/news/politics/2013/07/01/poll-supreme-court-gay-marriage-affirmative-action-voting-rights/2479541/>.

T. Millon, E. Simonsen & M. Birket-Smith, *Historical Conceptions of Psychopathy in the Untied States and Europe*, in PSYCHOPATHY: ANTISOCIAL, CRIMINAL, AND VIOLENT BEHAVIOR (Theodore Millon et al., eds., 2002).

Tang et al., *Personality Changes During Depression Treatment: A Placebo-Controlled Trial*, 66 ARCHIVES GEN. PSYCHIATRY 1322 (2009).

Tarynn M. Witten & A. Evan Eyler, *Hate Crimes and Violence Against the Transgendered*, 11 PEACE REV. 461 (1999).

TEMPLE GRANDIN & SEAN BARRON, THE UNWRITTEN RULES OF SOCIAL RELATIONSHIPS (2005).

Terrance McCoy, *Amid Measles Outbreak, Anti-Vaccine Doctor Revels in His Notoriety*, WASH. POST, Jan. 30, 2015, <https://www.washingtonpost.com/news/morning-mix/wp/2015/01/30/amid-measles-outbreak-anti-vaccine-doctor-revels-in-his-notoriety/>.

THE MATRIX (Warner Bros. 1999).

THÉRÈSE DELPECH, NUCLEAR DETERRENCE IN THE 21ST CENTURY (2012).

Thomas Fuchs, *Corporealized and Disembodied Minds: A Phenomenological View of the Body in Melancholia and Schizophrenia*, 12 PHIL., PSYCHIATRY, & PSYCHOL. 95 (2005).

TOM L. BEAUCHAMP & JAMES F. CHILDRESS, PRINCIPLES OF BIOMEDICAL ETHICS (5th ed. 2001).

Tonia L. Nicholls et al., *Psychopathy in Women: A Review of its Clinical Usefulness for Assessing Risk for Aggression and Criminality*, 23 BEHAVIORAL SCIENCES AND THE LAW 779 (2005).

United States Food and Drug Administration, *FDA News Release: FDA Approves Humanitarian Device Exemption for Deep Brain Stimulator for Severe Obsessive-Compulsive Disorder*, FDA.GOV (Feb. 19 2009), <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm149529.htm>.

Vaccines Interfere with your Body's Ability to Naturally Kill Viruses, Bacteria, Parasites and Tumor Cells, PRESS CORE (Jan. 15, 2011), <http://presscore.ca/vaccines-interfere-with-your-body%E2%80%99s-ability-to-naturally-kill-viruses-bacteria-parasites-and-tumor-cells>.

Vaughan Bell, *Don't Touch That Dial! A History of Media Technology Scares, from the Printing Press to Facebook*, SLATE (Feb. 15, 2010), http://www.slate.com/articles/health_and_science/science/2010/02/dont_touch_that_dial_single.html.

Vito Agosti & Patrick J. McGrath, *Comparison of the Effects of Fluoxetine, Imipramine and Placebo on Personality in Atypical Depression*, 71 J. AFFECTIVE DISORDERS 113 (2002).

VT. STAT. ANN.

WALTER GLANNON, BIOETHICS AND THE BRAIN (2007).

WALTER ISSACSON, EINSTEIN: HIS LIFE AND UNIVERSE (2007).

William Cronon, *The Trouble with Wilderness; or, Getting Back to the Wrong Nature*, in UNCOMMON GROUND: RETHINKING THE HUMAN PLACE IN NATURE (ed. William Cronon, 1995).

WINGSPREAD CONFERENCE ON THE PRECAUTIONARY PRINCIPLE, WINGSPREAD STATEMENT ON THE PRECAUTIONARY PRINCIPLE (1998).

WIS. STAT. ANN.

Yaling Yang & Adrian Raine, *Prefrontal Structure and Functional Brain Imaging Findings in Antisocial, Violent, and Psychopathic Individuals: A Meta-Analysis*, 174 PSYCHIATRY RES. 81 (2009).

Yingxu Wang & Vincent Chiew, *On the Cognitive Process of Human Problem Solving*, 11 COGNITIVE SYS. RES. 81 (2010).

Zackary Canepari, Drea Cooper, & Emma Cott, *Prosthetic Limbs, Controlled by Thought*, N.Y. TIMES (May 20, 2015), <http://www.nytimes.com/2015/05/21/technology/a-bionic-approach-to-prosthetics-controlled-by-thought.html>.